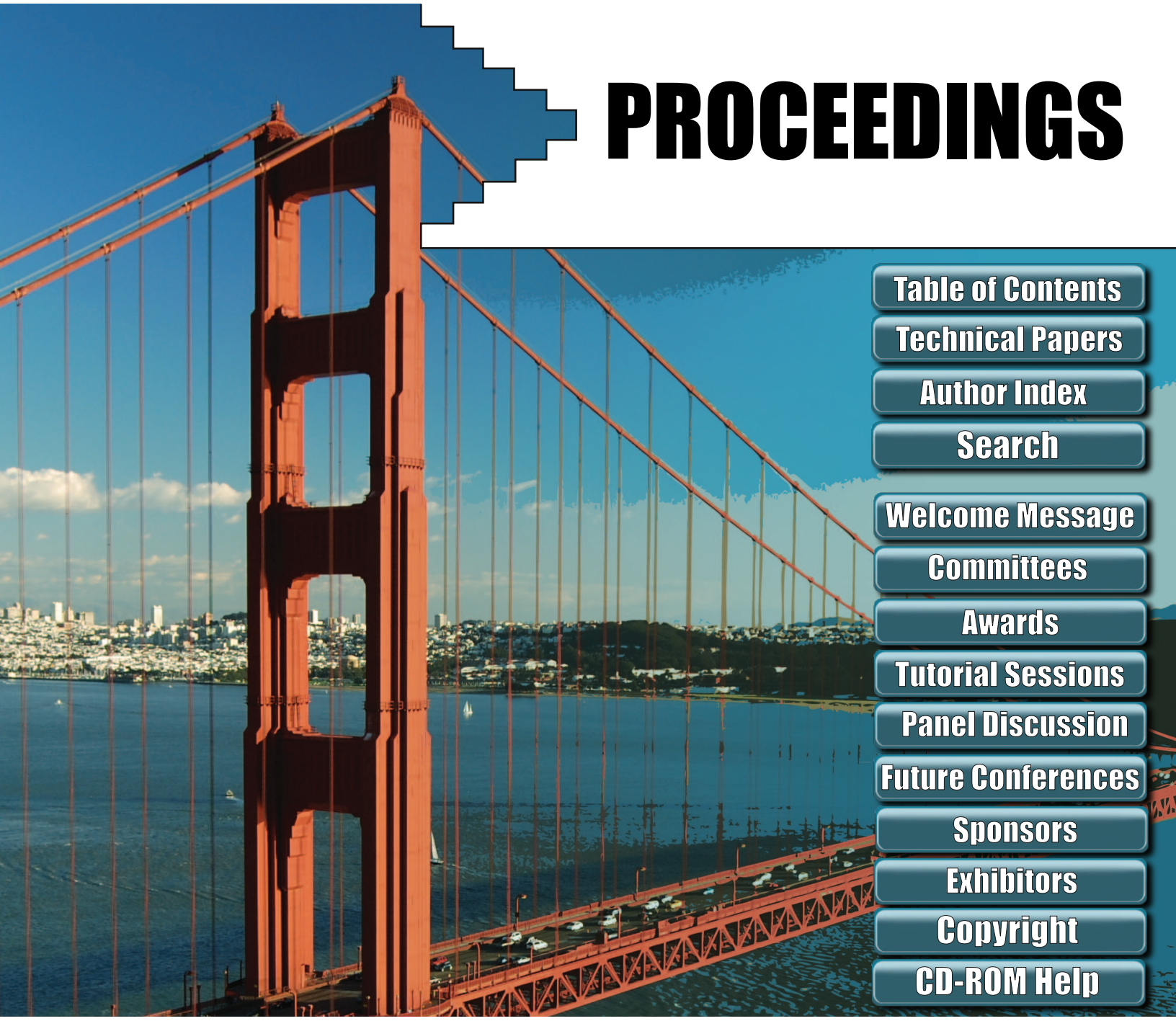
2011 JOINT CONFERENCE OF THE IEEE INTERNATIONAL FREQUENCY CONTROL SYMPOSIUM & EUROPEAN FREQUENCY AND TIME FORUM

Hyatt Regency San Fransisco | MAY 1-5, 2011 | San Fransisco, California, USA

PROCEEDINGS



IEEE UFFC EFTF

Technical support
Conference Catalysts, LLC
Phone: +1 785 341 3583
cdyer@conferencecatalysts.com

IEEE Catalog Number: CFP11FRE-CDR
ISBN: 978-1-61284-110-6

2011 Joint Conference of the IEEE International Frequency Control Symposium & European Frequency and Time Forum Proceedings

Additional copies may be ordered from:
IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08855-1331 USA

+1 800 678 IEEE (+1 800 678 4333)
+1 732 981 1393
+1 732 981 9667 (FAX)
email: customer-service@ieee.org

IEEE Catalog Number: CFP11FRE-CDR
ISBN: 978-1-61284-110-6

TABLE OF CONTENTS

Monday, May 2 10:50 a.m. – 12:30 p.m.

Session: Silicon and CMOS MEMS Resonators
Room: Grand Ballroom C
Session Chair: Sheng-Shiang Li, National Tsing Hua University

Monday, May 2 10:50 a.m. – 12:30 p.m.

Session: Time Transfer and Time Scales
Room: Seacliff Rooms
Session Chair: Samuel Stein, Symmetricom

Monday, May 21:30 p.m.-3:30 p.m.

Session: Fountain Clocks
Room: Grand Ballroom B
Session Chair: Robert Lutwak, Symmetricom

Monday, May 21:30 p.m.-3:30 p.m.

Session: New Material Applications
Room: Grand Ballroom C
Session Chair: Ralf Lucklum, Tech. Univ. of Magdeburg

Monday, May 21:30 p.m.-3:30 p.m.

Session: Phase Noise Measurement and Modeling
Room: Seacliff Rooms
Session Chair: Warren Walls, US Naval Observatory

Monday, May 23:50 p.m. – 5:50 p.m.

Poster Session: Resonators
Room: Pacific Concourse - Area 1
Session Chair: Yoonkee Kim, US Army

Monday, May 2 3:50 p.m. – 5:50 p.m.

Session: Microwave Frequency Standards I
Room: Poster Area 3
Session Chair: Svenja Knappe, NIST

Monday, May 2 3:50 p.m. – 5:50 p.m.

Poster Session: Oscillatorsm Sensors and Transducers I
Room: Pacific Concourse **-** Area 4
Session Chair: Christine Klemenz Rivenbark, Krystal Engineering LLC

Monday, May 2 3:50 p.m. – 5:50 p.m.

Poster Session: Timekeeping, Time and Frequency Transfer, GNSS Applications I
Room: Pacific Concourse - Area 5
Session Chair: Patrizia Tavella, INRIM

Monday, May 2 3:50 p.m. – 5:50 p.m.

Poster Session: Optical Frequency Standards & Applications I
Room: Pacific Concourse **-** Area 6
Session Chair: Pierre Thomann, Universite de Neuchatel

Tuesday, May 3 8:30 a.m. – 11:00 a.m.

Session: Ion Clocks
Room: Grand Ballroom B
Session Chair: Eric Burt, JPL/NASA

Tuesday, May 3 8:30 a.m. – 11:00 a.m.

Session: Novel structures for MEMS Resonators
Room: Grand Ballroom C
Session Chair: Gianluca Piazza, University of Pennsylvania

Tuesday, May 3 8:30 a.m. – 11:00 a.m.

Session: Two Way Optical and Microwave Links
Room: Seacliff Rooms
Session Chair: Mizuhiko Hosokawa, NICT

Tuesday, May 3 1:30 p.m. – 3:30 p.m.

Session: Combs
Room: Grand Ballroom B
Session Chair: Scott Crane, USNO

Tuesday, May 3 1:30 p.m. – 3:30 p.m.

Session: Liquid Sensors
Room: Grand Ballroom C
Session Chair: Fabien Josse, Marquette University

Tuesday, May 3 1:30 p.m. – 3:30 p.m.

Session: Frequency Synthesis & CMOS Oscillators
Room: Seacliff Rooms
Session Chair: Jeremy Everard, University of York

Tuesday, May 3 3:50 p.m. – 5:50 p.m.

Poster Session: Materials and Filters
Room: Pacific Concourse - Area 1
Session Chair: Yoonkee Kim, US Army

Tuesday, May 3 3:50 p.m. – 5:50 p.m.

Poster Session: Oscillators, Synthesizers, Noise & Circuit Techniques II
Room: Pacific Concourse - Area 2
Session Chair: Jeremy Everard, University of York

Tuesday, May 3 3:50 p.m. – 5:50 p.m.

Poster Session: Microwave Frequency Standards II
Room: Pacific Concourse - Area 3
Session Chair: Pierre Thomann, Universite de Neuchatel

Tuesday, May 3 3:50 p.m. – 5:50 p.m.

Poster Session: Timekeeping, Time and Frequency Transfer, GNSS Applications II
Room: Pacific Concourse - Area 5
Session Chair: Patrizia Tavella, INRIM

Tuesday, May 3 3:50 p.m. – 5:50 p.m.

Poster Session: Optical Frequency Standards & Applications II
Room: Pacific Concourse - Area 6
Session Chair: Pierre Thomann, Universite de Neuchatel

Wednesday, May 4 8:30 a.m. – 10:30 a.m.

Session: Optimized and Tunable MEMS Resonators
Room: Grand Ballroom C
Session Chair: Yoonkee Kim, US Army

Wednesday, May 4 8:30 a.m. – 10:30 a.m.

Session: Celebrating Kalman Filter 50th Anniversary
Room: Seacliff Rooms
Session Chair: Demetrios Matsakis, USNO

Wednesday, May 4 **10:50 a.m. – 12:30 p.m.**

Session: Microclocks & Novel Concepts
Room: Grand Ballroom B
Session Chair: John Kitching, NIST

Wednesday, May 4 **10:50 a.m. – 12:30 p.m.**

Session: Quartz Based Resonators
Room: Grand Ballroom C
Session Chair: Ji Wang, Ningbo University

Wednesday, May 4 **10:50 a.m. – 12:30 p.m.**

Session: Algorithms for Clock and Time Scale Estimations
Room: Seacliff Rooms
Session Chair: Ilaria Sesia, INRIM

Wednesday, May 4 **1:50 p.m.-3:50 p.m.**

Session: Lattice Clocks II
Room: Grand Ballroom B
Session Chair: Ekkehard Peik, PTB

Wednesday, May 4 **1:50 p.m.-3:50 p.m.**

Session: Physical Sensors
Room: Grand Ballroom C
Session Chair: Don Malocha, University of Central Florida

Wednesday, May 4 1:50 p.m.-3:50 p.m.

Session: Acoustic Oscillators and Micromechanical Resonators
Room: Seacliff Rooms
Session Chair: Michael Driscoll, Northrop Grumman

Wednesday, May 4 **4:10 p.m.-6:10 p.m.**

Session: Space Clocks & Novel Oscillators
Room: Grand Ballroom B
Session Chair: Louis Marmet, NRC Canada

Wednesday, May 4 **4:10 p.m.-6:10 p.m.**

Session: Mass Sensors
Room: Grand Ballroom C
Session Chair: Leo Reindl, University of Freibourg

Wednesday, May 4 **4:10 p.m.-6:10 p.m.**

Session: Optical Fiber Time and Frequency Transfer
Room: Seacliff Rooms
Session Chair: Jan Johansson, SP

Thursday, May 5 **8:30 a.m.-10:30 a.m.**

Session: Vapor Cell Clocks
Room: Grand Ballroom B
Session Chair: Svenja Knappe, NIST

Thursday, May 5 **8:30 a.m.-10:30 a.m.**

Session: Filters
Room: Grand Ballroom C
Session Chair: Rick Puccio, Quartzdyne

Thursday, May 5 **8:30 a.m.-10:30 a.m.**

Session: Optical Oscillators & Components
Room: Seacliff Rooms
Session Chair: Michael Tobar, University of Western Australia

Thursday, May 5 **10:50 a.m.-12:50 p.m.**

Session: Fundamental Measurement & Sensors
Room: Grand Ballroom B
Session Chair: Giorgio Santarelli, LNE-SYTRE

Thursday, May 5 **10:50 a.m.-12:50 p.m.**

Session: Materials
Room: Grand Ballroom C
Session Chair: Bernard Dulmet, ENS2M

Thursday, May 5 10:50 a.m.-12:50 p.m.

Session: GPS Carrier Phase Time Comparison and Calibration
Room: Seacliff Rooms
Session Chair: Pascale Defraigne, ORB

Welcome from the General Co-Chairs, Prof. Clark T.-C. Nguyen and Dr. Giorgio Santarelli

It is our great pleasure to welcome you to the San Francisco Bay Area and to the Joint Conference of the 65th IEEE International Frequency Control Symposium and 25th European Frequency and Time Forum. This event marks the 5th joint meeting to combine the two leading international technical conferences for research, development, and applications of timing and frequency control.

The 2011 Organizing and Joint Technical Program Committees have strived to provide a stimulating and educational program for all attendees. With a larger than usual number of abstracts to choose from, the conference program easily coalesced to one that encompasses all traditional topics, to include resonators, oscillators, atomic standards, optical frequency standards, timekeeping, and resonant sensors, with a noticeable increase in contributions from MEMS technologists.

On 1 May, the day prior to the meeting, twelve educational and informative Tutorial Sessions are offered. These sessions will cover topics of interest to attendees, and will be presented by internationally recognized experts in their field.

From May 2-5, the conference program includes three parallel sessions of technical presentations and two poster sessions, as well as a panel session on Tuesday morning that is sure to provoke thoughtful discourse among conference participants. Monday and Tuesday evenings offer a Welcome Reception and an Exhibitor's Reception, respectively, giving you ample opportunity to socialize with colleagues and perhaps stimulate that next big idea over a few drinks. Throughout the Symposium an exhibit area of leading frequency control and timekeeping manufacturers will be open and will serve as the site for session breaks, as well as the aforementioned Exhibitor's Reception.

Since the conference venue is situated near the bay at the edge of beautiful downtown San Francisco, it is very close to many of San Francisco's bay front attractions that you may wish to enjoy before or after the conference. Among these attractions are the San Francisco Museum of Modern Art, Pier 39, the Exploratorium, the Aquarium of the Bay, Coit Tower, the Palace of Fine Arts Theatre, the Presidio, and a plethora of world class restaurants. Beyond the bay front lie countless other attractions in the greater San Francisco Bay Area, e.g., the wine country up north, the Golden Gate Bridge, Silicon Valley down south, etc. Needless to say, there is much to do and see at this location, so hopefully you've planned a few extra days before or after the conference in the area.

We hope the Joint Conference will provide you an opportunity to network with colleagues, reconnect or make new friends, and share advancements in the frequency and timing field.

Have a great stay in San Francisco and enjoy the 2011 Joint Conference!

Clark T.-C. Nguyen and Giorgio Santarelli
2011 Joint Conference General Co-Chairs

Welcome from the Joint Program Co-Chairs, Eric Burt and Patrick Gill

It is our pleasure to welcome you to the 2011 Joint IFCS/EFTF. The first joint conference was held in 1999 in Besancon, France with an initial agreement to repeat the joint configuration every 4 years. However, with the success of these joint conferences, in 2007 it was decided to move to a biennial format. 2011 marks the 5th joint conference. With 3 parallel tracks over 4 days covering the entire spectrum of time and frequency topics, the 2011 conference promises to be as technically successful as its predecessors, providing attendees with the latest scientific and engineering results from the international time and frequency community

The field of time and frequency is currently experiencing rapid development in several areas. Pushing the limits of accuracy, optical frequency metrology has emerged as one of the fastest growing sub-disciplines. Developments over the past decade now include optical clocks with accuracy in the eighteenth digit as well as all-optical means of transferring that accuracy both in-house and over trans-national distances via optical fiber and femtosecond combs. The most recent developments in optical frequency metrology will be thoroughly represented at the 2011 conference.

Pushing in the direction of the ultra-small, MEMS technology is poised to revolutionize applications of time and frequency that require low-cost, miniature, ubiquitous sources with high performance. MEMS technology was featured in the 2010 IFCS and a successful panel discussion was held to discuss if and how MEMS technology might replace quartz. In 2011 we will continue to emphasize MEMS technology through sessions devoted to this topic.

In between the two extremes of highly accurate laboratory frequency standards and ultra-miniature devices are small or even chip-scale atomic clocks with impressive performance. With size reduction and increasing performance of such devices, a wide spectrum of applications are arising ranging from clocks in space to sensors that are based on atomic clock technology. With their decreasing mass/power/volume footprint comes the question of whether chip-scale atomic clocks also might someday replace quartz oscillators in certain applications. On Tuesday, May 3, there will be a panel discussion on this topic.

The SI second remains defined in terms of a microwave transition in cesium so that all primary frequency standards are based on this transition. The performance of cesium primary microwave frequency standards is 2 orders of magnitude better than two decades ago. The latest generation of these frequency standards will be discussed in a session devoted to fountain clocks.

In 1960 Prof. Rudolf Kalman described what came to be known as Kalman Filtering. This method of estimating the values of variables in the presence of multiple noise types is now a standard tool in time scale formation. To mark the 50th anniversary of Kalman's work, which occurred last year, we will have an entire session and several tutorials devoted to Kalman Filtering.

The student poster competition has become a tradition at this conference. Started in 2004, the competition has grown steadily each year and the quality of the papers has evolved to be extremely high. In the 2010 IFCS there were over 50 entries. In 2011 there were 76 entries submitted. 24 finalists representing all 6 topical groups of the conference will be on display and judged during the first poster session on Monday, May 2. Winners will be announced during the reception on Monday night.

In addition to the regular technical sessions, we will have a day dedicated to tutorials. There will also be an extensive exhibition with representatives from leading manufacturers and suppliers of frequency control products and equipment from around the world.

We hope that you will have an enjoyable and technically rewarding 2011 conference!

Eric Burt and Patrick Gill
2011 IFCS/EFTF JPC Co-Chairs

Organizing Committee

2011 Joint Conference of the IEEE International Frequency Control Symposium & European Frequency and Time Forum

General Co-Chair
Clark T.-C. Nguyen
University of California at Berkeley
Berkeley, CA
ctnguyen@eecs.berkeley.edu

General Co-Chair
Giorgio Santarelli
LNE-SYRTE/OBSERVATOIRE DE PARIS
Paris, France
giorgio.santarelli@obspm.fr

Technical Program Co-Chair
Eric Burt
Jet Propulsion Laboratory
Pasadena, CA
Eric.A.Burt@jpl.nasa.gov

Technical Program Co-Chair
Patrick Gill
National Physical Laboratory
Middlesex, UK
patrick.gill@npl.co.uk

Panel Session Chair
John R. Vig
Consultant, USA
j.vig@ieee.org

Academic Chair
Kurt Gibble
Penn State University
University Park, PA
kgibble@psu.edu

Editorial Chair
Debra Coler
OEWaves, Inc.
Pasadena, CA
debra@OEWaves.com

Finance Chair
Wan-Thai Hsu
Discera, Inc.
whsu@discera.com

Tutorial Chair
Judah Levine
NIST
Boulder, CO
jlevine@boulder.nist.gov

Awards Chair
Andreas Bauch
Physikalisch-Technische Bundesanstalt
Braunschweig, Germany
andreas.bauch@ptb.de

Exhibits Co-Chair
Doug Lowrie
Symmetricom
DLowrie@symmetricom.com

Exhibits Co-Chair
Joel Petetin
Service Temps-Fréquence-Chronométrie
Institut UTINAM - Observatoire de Besançon
joel.petetin@obs-besancon.fr

Exhibits Coordinator
Sue Kingston
Conference Catalysts, LLC
skingston@roadrunner.com

Conference Management
Chris Dyer
Conference Catalysts, LLC
cdyer@conferencecatalysts.com

2011 IFCS/EFTF Joint Program Committee

Group 1: Materials, Resonators, & Resonator Circuits

Yoonkee Kim, US Army (Vice Chair)

Jean-Pierre Aubry, Oscilloquartz
Sunil Bhave, Cornell University
Shih Chuang, Statek
Bernard Dulmet, Ecole Nat. Superior of Mech. & Micro (Besancon)
Raymond Filler, US Army
Vladimir Klipov, Sawyer
George Mansfeld, Inst of Radio Engrg
Sheng-Shian Li, Nat. Tsing Hua Univ.
Bernd W. Neubig, Adv. Crystal Products
Moorthi Palaniapan, Nat. Univ. of Singapore
Gianluca Piazza, Univ. of Pennsylvania
Derek (Rick) Puccio, Quartzdyne
Dan Stevens, Vectron
Ji Wang, Ningbo Univ.
Yasuaki Watanabe, Tokyo Metro Univ.
Dana Weinstein, MIT
Yook-Kong Yong, Rutgers University

Group 2: Oscillators, Synthesizers, Noise, & Circuit Techniques

Jeremy Everard, University of York (Vice Chair)

Farrokh Ayazi, Georgia Tech.
Martin Bloch, Frequency Electronics
Rodolphe Boudot, FEMTO-ST
Pierre-Yves Bourgeois, FEMTO-ST
Michael Driscoll, Northrop Grumman
Gary Fedder, CMU
Marvin Frerking, Innovative Tech. Products
Patrick Green, Northrop Grumman
Philippe Guillemot, CNES
John Hartnett, Univ. of Western Australia
Dave Howe, NIST
Wan-Thai Hsu, Discera
Eugene Ivanov, Uni Western Australia
Eun Sok Kim, USC
Takeo Oita, NDK
Enrico Rubiola, Femto-ST
Jesse Searls, Poseidon Scientific Inst.
Michael Tobar, Univ. of Western Australia
Mike Underhill, Underhill Research
Warren Walls, US Naval Observatory

Group 3 and 6: Microwave and Optical Frequency Standards and Applications

Pierre Thomann, University Neuchatel (Vice Chair)

Sebastien Bize, SYRTE
Eric Burt, JPL/NASA
Luigi Cacciapuotti, ESA
Scott Crane, USNO
Anne Curtis, NPL
Scott Diddams, NIST
Chris Ekstrom, USNO
Stephen Feltham, ESA
Mike Garvey, Symmetricom
Patrick, Gill NPL
Kurt Gibble, Penn State Univ.
John Kitching, NIST
Steve LeComte, CSEM
Pierre Lemonde, SYRTE
Filippo Levi, INRIM
Robert Lutwak, Symmetricom
Andrew Ludlow, NIST
Long-Sheng Ma, East China Normal Univ.
Louis Marmet, Nat. Rsch. Council Canada
Eamonn Murphy, ESA
Ekkehard Peik, PTB
Nicola Poli, LENS
John Prestage, JPL/NASA
Fritz Riehle, PTB
Christophe Salomon, LKB/ENS
Giorgio Santarelli, SYRTE
Bob Tjoelker, JPL/NASA
Peter Wolf, SYRTE
Jun Ye, JILA

Group 4: Sensors & Transducers

Don Malocha, Univ. of Central Florida (Vice Chair)

Jeff Andle
Sylvain Ballandras, Femto-St
Mauricio Pereira da Cunha, Univ. of Maine
Philip X.L. Feng, Cal Tech
Diethelm Johannsmann, TU Clausthal
Fabien Josse, Marquette Univ.
Shigeru Kurosawa, Adv. Industrial Science & Tech.
A.K. Sarin Kumar, Honeywell
Svenja Knappe, NIST
Ryszard Lec, Drexel Univ.
Ralf Lucklum, Tech. Univ. of Magdeburg
Paul Muralt, EPFL
Glen McHale, Nottingham Trent University
Leo Reindl, Univ. of Freiburg
Clemens Ruppel, EPCOS AG
Ashwin Seshia, University of Cambridge
Isao Shimoyama, Univ. of Tokyo

Group 5: Timekeeping, Time and Frequency Transfer, GNSS Applications

Patrizia Tavella, INRIM (Vice Chair)

Andreas Bauch, PTB
Laurent Guy Bernier, METAS
Pascale Defraigne, Obs. Royal de Belgique
Jerome Delporte, CNES
Javier De Vicente, ESA
Gesine Grosche, PTB
Jorg Hahn, ESA
Michito Imae, AIST
Jan Johansson, SP
Judah Levine, JILA
Demetrios Matsakis, USNO
Dirk Piester, PTB
Wolfgang Schaefer, Timetech
Samuel Stein, Symmetricom
Philip Tuckey, SYRTE
Pierre Uhrich, SYRTE
Pierre Waller, ESA
Bruce Warrington, CSIRO
Victor Zhang, NIST

2011 IFCS Awards

IFCS I.I. Rabi Award

This award recognizes outstanding contributions related to the fields of atomic and molecular frequency standards, time scale realization, and time comparison and dissemination.

2011 IFCS I.I. Rabi Award:

Prof. Fritz Riehle, Physikalisch-Technische Budesanstalt (PTB)

Citation: “For outstanding contributions to the field of atomic frequency standards, including pioneering research and development of optical frequency standards and their measurement.”

IFCS W.G. Cady Award

This award recognizes outstanding contributions related to the fields of crystal resonators, frequency control, frequency synthesis, noise measurement and sensor devices.

2011 IFCS W.G. Cady Award:

Prof. Georgy D. Mansfeld, Russian Academy of Sciences

Citation: “For pioneering research in physical acoustoelectronics and acoustics, including acoustic resonance spectroscopy, attenuation mechanisms, piezoelectric properties of thin layers and films, and new acoustic materials.”

IFCS C.B. Sawyer Memorial Award

This award recognizes outstanding contributions in the development, production or characterization of piezoelectric materials of interest to the Symposium Technical Program Committee, or to recognize entrepreneurship or leadership within profit or non profit organizations in the frequency control community (including all parts of the community).

2011 IFCS C.B. Sawyer Memorial Award:

Prof. Achim Wixford, University of Augsburg

Citation: “For the development and successful commercialization of surface acoustic wave driven microfluidic systems.”

2012 IFCS Award Nominations

Nominations are sought for the 2012 IFCS awards. Information is available at the Registration Desk and is also available on the IEEE Frequency Control Conference website at: http://www.ieee-uffc.org/fc

2011 EFTF Awards

EFTF Award

The award, which is open to anyone working in the fields traditionally associated with the European Frequency and Time Forum (EFTF), recognizes exceptional contributions in all fields covered by the EFTF either for fundamental advances or important applications. The significance of contributions includes the degree of initiative and creativity, the quality of work, the degree of success obtained as well as the worldwide scientific impact on the Time and Frequency Community. This award is sponsored by the Société Française des Microtechniques et de Chronométrie. The recipients are selected by the Executive Committee of the EFTF.

2011 EFTF Award:

Dr. James C. Bergquist, NIST Time and Frequency Division

Citation: “For seminal contributions to high-resolution laser spectroscopy and frequency standards for over 30 years.”

EFTF Young Scientist Award:

The EFTF Young Scientist Award is conferred in recognition of a personal contribution that demonstrated a high degree of initiative and creativity and lead to already established or easily foreseeable outstanding advances in the field of time and frequency metrology. A prerequisite is that the work to be taken under consideration has been published in the Proceedings of the European Frequency and Time Forum or the IEEE Frequency Control Symposium. The award honors a person under the age of 40 at the date of the opening session of the 2011 EFTF-IFCS joint conference. The goal is to encourage scientific endeavor and competition and to help young scientists along their career paths. This award is sponsored by the Société Française des Microtechniques et de Chronométrie. The recipients are selected by the Executive Committee of the EFTF.

2011 EFTF Young Scientist Award:

Dr. Tobias Kippenberg, Ecole Polytechnique Fédérale de Lausanne and Max Planck Institute for Quantum Optics

Citation: “For contributions to optical frequency metrology by the demonstration of monolithic micro resonator frequency comb generators.”

EFTF Marcel Ecabert Award:

The Marcel Ecabert Award of the European Frequency and Time Forum (EFTF) is a lifetime award and honors the excellent achievements of the recipient in the field of time and frequency. It is named after the late Marcel Ecabert, founding member of the EFTF and member of its Executive Committee.

2011 EFTF Marcel Ecabert Award:

Dr. Giovanni Busca, KYTIME

Citation: “In recognition of a career successfully devoted to the development of atomic frequency standards, and for his major contributions to the establishment of an industrial infrastructure for ground and space clocks at European level.”

Student Paper Competition Awards

Group 1: Materials, Resonators, & Resonator Circuits

"Mechanically-Coupled CMOS-MEMS Free-Free Beam Resonator Arrays with Two-Port Configuration"
Ming-Huang Li, National Tsing Hua University

Group 2: Oscillators, Synthesizers, Noise, & Circuit Techniques

"Effects of Volume and Frequency Scaling in AlN Contour Mode NEMS Resonators on Oscillator Phase Noise"
Matteo Rinaldi, University of Pennsylvania

Group 3: Microwave Frequency Standards and Applications

"CPT Pump-Probe Measurement of the Cs Clock Transition DC Stark Shift"
Jean-Luc Robyr, University of Fribourg

Group 4: Sensors & Transducers

"SAW-Thin-Film Acoustoelectric Effect in-Situ Observation and Measurement"
Brian Fisher, University of Central Florida

Group 5: Timekeeping, Time and Frequency Transfer, GNSS Applications

"Absolute Calibration of GNSS Time Transfer Systems : NRL and CNES Techniques Comparison"
Amandine Proia, CNES/BIPM

Group 6: Optical Frequency Standards and Applications

"Absolute Frequency Measurement of the 1S-2S Transition in Atomic Hydrogen Using a 900 km Remote Frequency Reference"
Katharina Predehl, Max-Planck-Institue of Quantum Optics

Tutorials

	Seacliff A	Seacliff B	Seacliff C
8:30 a.m. - 10:30 a.m.	**Use of Kalman Filters in Time and Frequency Analysis** John Davis, National Physical Laboratory, Teddington, UK	**The Leeson Effect, Phase Noise and Frequency Stability in Oscillators** Enrico Rubiola, Femto-St, CNRS	**Theory and Analysis of MEMS Resonators** Ville Kaajakari, VTI Technologies
10:30 a.m.- 10:45 a.m.	Break – Seacliff Foyer		
10:45 a.m. - 12:45 p.m.	**Fundamentals of Kalman Filtering and Applications to GNSS** Mohinder Grewal, California State University, Fullerton	**Phase Noise Measurements** Craig Nelson, NIST Boulder	**MEMS Resonators for Frequency Control and Sensing Applications** Gianluca Piazza, Univ. of Pennsylvania
12:45 a.m. - 1:45 p.m.	Lunch – Hospitality Room		
1:45 p.m. - 3:45 p.m.	**Principles of Atomic Clocks** Robert Lutwak, Symmetricom	**Variance Measurements – Practical Use – Statistics – Long Term Prediction** Francois Vernotte, Besancon Observatory	**Nanoscale Electromechanical Resonators and Oscillators** Philip X. L. Feng, Case Western Reserve University
3:45 p.m. - 4:00 p.m.	Break – Seacliff Foyer		
4:00 p.m. - 6:00 p.m.	**Next-Generation Frequency Standards** Bruce Warrington, National Measurement Institute	**The Kalman Filter** Lorenzo Galleani, Politecnico di Torino	**Theory and Analysis of Quartz Crystal Resonators** Yook Kong Yong, Rutgers University and Ji Wang, Ningbo University, China

8:30 a.m. – 10:30 a.m.
Room: Seacliff A

Use of Kalman Filters in Time and Frequency Analysis
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Davis.pdf)
John Davis, National Physical Laboratory

Kalman filters have many applications both in time and frequency and in related topics, for example GNSS analysis.

In this tutorial an overview will be given of the Kalman filter and its relationship to other filters and smoothing algorithms. A detailed description will be given of each element of the Kalman filter, with emphasis on the underlying physical meanings and limited use of mathematics. A step by step description of the filtering process will then be provided.

There are several key applications of Kalman filters within time and frequency analysis. These include clock ensemble algorithms, clock predictors and steering algorithms. The use within the Kalman filter of each of the well-known power law noise processes will be examined. Models of clock and time transfer noise will be discussed in detail, along with the treatment of periodic instabilities and linear frequency drift. Methods of testing Kalman filters using simulated data sets will be demonstrated. Finally the concept of observability, which is of particular importance to clock ensemble algorithms, will be examined.

8:30 a.m. – 10:30 a.m.
Room: Seacliff B

The Leeson Effect, Phase Noise and Frequency Stability in Oscillators
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Rubiola.pdf)
Enrico Rubiola, CNRS FEMTO-ST Institute

Simply stated, an oscillator consists of a loop in which a resonator sets the oscillation frequency and an amplifier compensates for the resonator loss. The oscillation amplitude is set by clipping or other gain-saturation mechanism, usually in the amplifier. When phase noise is introduced in the loop, the oscillator converts it to frequency noise through a process of time-domain integration. The consequence is that the oscillator phase fluctuation diverges in the long run. This phenomenon was originally referred as the “Leeson model” after a short article published by D. B. Leeson [1]. On my side, I prefer the term “Leeson effect” in order to emphasize that it is far more general than a simple model [2].

The first part of this tutorial explains the phase-to-frequency conversion mechanism as a general phenomenon inherent in the feedback, following an heuristic approach based on physical insight. There follow the relationships between the noise of the internal components (sustaining amplifier, resonator, etc.) and the phase noise at the oscillator output, or equivalently the frequency stability.

The second part is the analysis of the phase noise spectra found in the data-sheet of commercial oscillators: dielectric-resonator oscillator (DRO), whispering gallery oscillator (WGO), 5–100 MHz quartz crystal oscillators, opto-electronic oscillator (OEO). The analysis gives information on the most relevant design parameters, like the quality factor Q and the driving power of the resonator, and the flicker noise of the sustaining amplifier.

The last part shows the derivation of the oscillator phase noise formulae from the elementary properties of the resonator. Interestingly, the amplitude non-linearity, necessary for the oscillation amplitude to be stable, splits the resonator relaxation time into two time constants. The approach shown in this last part is general. It applies to all oscillators, including quartz, RLC, microwave cavity, delay-line, laser, etc.

This tutorial updates on the earlier editions, and also on the book [2] by including the AM noise and the impact of AM noise on phase noise [3].

References
[1] D. B. Leeson, A simple model for feed back oscillator noise spectrum, Proc. IEEE 54(2) pp. 329–330, Feb. 1966.
[2] E. Rubiola, *Phase noise and frequency stability in oscillators*, Cambridge 2008 & 2010.
[3] E. Rubiola, R. Brendel, A generalization of the Leeson effect, April 2010 (48 pages).
arXiv:1004.5539 [physics.ins-det].

8:30 a.m. – 10:30 a.m.
Room: Seacliff C

Theory and Analysis of MEMS Resonators
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Kaajakari.pdf)
Ville Kaajakari, VTI Technologies

Microresonators offer promise to integrate bulky analog and RF components on-chip enabling smaller form and reduced system cost. MEMS resonators have been proposed as one-to-one replacements of macroscopic components such as quartz crystals and SAW filters. Microresonator based oscillators have been demonstrated to meet the stringent phase noise requirements of reference oscillators for wireless communication systems, and resonators have been demonstrated to reach frequencies of over 2 GHz. While small may be desirable, scaling down the resonator also has unavoidable physical limits. The
research work of past years allows critical evaluation of these limits and provides tools to analyze feasibility of different microresonator applications.

In this course, we will cover the microresonator basics with the goal of understanding what are the practical applications and limitations of the resonators. We start by modelling of the mechanical resonator with the goal of representing the mechanical resonator as a circuit element (LRC resonator). First, the distributed resonator vibrations are represented with a lumped element. Next, electromechanical coupling mechanisms (capacitive and piezoelectric) are introduced and compared. The theory is expanded to include second order effects such as electrostatic spring softening and nonlinear spring effects. As nonlinear effects effectively set the resonator power handling capacity, the estimation of the nonlinear limit is of fundamental importance for estimating the fundamental limits for miniaturization. Finally, the mechanical resonator and electromechanical transduction models are combined to an electrical equivalent circuit suitable for system design.

In the second part of the tutorial, we compare the model parameters (e.g. motional resistance Rm, electromechanical coupling coefficient kt2) for selected microresonators and commercial devices. These characteristics are compared to application specific criteria.

10:45 – 12:45
Room: Seacliff A

Fundamentals of Kalman Filtering and Applications to GNSS
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Kaajakari.pdf)
Mohinder Grewal, University of California, Fullerton

This tutorial presents the fundamentals of Kalman filtering with application to GNSS. It addresses the subtleties, problems, and limitations of estimation theory as applied to real world situations encountered in GNSS and navigation, and provides application examples.

10:45 a.m. – 12:45 p.m.
Room: Seacliff B

Phase Noise Measurements
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Nelson.pdf)
Craig Nelson, NIST Boulder

Noise is everywhere. Its ubiquitous nature interferes with or masks desired signals and fundamentally limits all electronic measurements. Noise in the presence of a carrier is experienced as amplitude and phase modulation noise. Modulation noise will be covered from its theory, to its origins and consequences. The effects of signal manipulation such as amplification, frequency translation and multiplication on spectral purity are examined. Practical techniques for measuring AM and PM noise, from the simple to complex will be discussed. Calibration of measurements and common problems and pitfall will also be covered.

Craig Nelson is an electrical engineer at the Time and Frequency Division of the National Institute of Standards and Technology. He received his BSEE from the University of Colorado in Boulder in 1990. After co-founding SpectraDynamics, he joined the staff at the NIST. He has worked on the synthesis and control electronics, as well as software for both the NIST-7 and F1 primary frequency standards. He is presently involved in research and development of ultra-stable synthesizers, low phase noise electronics, and phase noise metrology. Current areas of research include optical oscillators, high-speed pulsed phase noise measurements and phase noise metrology in the 100 GHz range. He has published over 35 papers and teaches classes, tutorials, and workshops at NIST, the IEEE Frequency Control Symposium, and several sponsoring agencies on the practical aspects of high-resolution phase noise metrology.

10:45 a.m. – 12:45 p.m.
Room: Seacliff C

MEMS Resonators for Frequency Control and Sensing Applications
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Piazza.pdf)
Gianluca Piazza, University of Pennsylvania

This tutorial will review the state-of-the-art in MEMS resonator technology that uses piezoelectric, electrostatic and thermal transduction. Examples of some of the latest micromechanical resonators under development at Universities, National Labs and industry will be provided. Device design, fabrication and performances will be discussed and compared to commercially available resonators such as quartz crystals and surface acoustic wave devices. The utilization of these resonators for frequency control applications such as filtering and high precision signal generation will be presented. In addition, the ability of arraying these devices, fabricating them in large volumes and integrating them with electronics will be emphasized in regards to new opportunities in the area of low power front-ends, channelizers and reconfigurable circuits. Furthermore, the continuous miniaturization and frequency scaling of the MEMS and NEMS resonator technology has enabled the utilization of these devices for the making of very sensitive, portable and low-power sensor arrays. The tutorial will briefly review some examples of chem/bio resonant detectors from academia and industry, and highlight the advantages that MEMS will bring to this space with respect to currently available solutions.

10:45 a.m. – 12:45 p.m.
Room: Seacliff A

Principles of Atomic Clocks
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Lutwak.pdf)
Robert Lutwak - Symmetricom Technology Realization Center

Atomic frequency standards provide the ultimate source of accuracy and stability for all modern communications, navigation, and time-keeping systems. Commercially-available “Industrial” atomic clocks, including cesium beam frequency standards, rubidium oscillators, and hydrogen masers, are based on technology originally developed in the 1950’s. Since that time, technology evolution and field experience have led to a level of performance and reliability that atomic clocks are now deployed throughout critical infrastructure applications.
Recently, for the first time in 50 years, a new type of atomic clock has transitioned from the laboratory into practical application. Chip-Scale Atomic Clocks (CSAC) employ novel techniques for atomic confinement and interrogation, combined with micro-electromechanical systems (MEMS) and modern electronics, to enable atomic timing accuracy in portable battery-powered devices. While CSAC technology is yet relatively immature, already a new family of applications has emerged in handheld communication systems, enhanced personal navigation systems, and remote sensing.
The tutorial will provide an introduction to industrial atomic clocks, focusing on mature technologies: rubidium oscillators, cesium beam frequency standards, and hydrogen masers, as well as the emerging technology of CSAC.

1:45 p.m. – 3:45 p.m.
Room: Seacliff B

Variance Measurements – Practical Use – Statistics – Long Term Prediction
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Vernotte.pdf)
Francois Vernotte, Besancon Observatory

The seminar addresses the estimation of the error bars in the Allan variance, and explains why the upper half-bar in the log-log ADEV plot is bigger than the lower half-bar (while it is a common belief that the lower half-bar is bigger).

Outline
- Practical Allan variance algorithms (with / without overlapping)
- Allan variance versus Allan deviation
- Statistics of the Allan variance and the Allan deviation: chi-squared and Rayleigh distribution
- How to estimate the number of equivalent degrees of freedom (edf) of an Allan variance/deviation measure?
- Probability Density Function and Cumulative Distribution Function of the Allan variance/deviation measures
- Confidence interval over the Allan variance/deviation measures
- Fitting curve over Allan variance measurement
- Link between noise levels and Allan variance response
- Estimation of the noise levels and their uncertainty from the fitting curve
- Extrapolation to very long term time stability
- How to increase the number of edf: the Total variance
- Generalization to the other variances.

1:45 p.m. – 3:45 p.m.
Room: Seacliff C

Nanoscale Electromechanical Resonators and Oscillators
Philip Feng**,** Case Western Reserve University

Nanoscale devices with mechanical degrees of freedom offer compelling characteristics that make them not only interesting tools for fundamental studies, but also intriguing candidates for technological applications. In particular, nanoelectromechanical systems (NEMS) *vibrating* in their *resonant* modes provide promising opportunities and advantages for developing novel transducers and sensors, in the previously inaccessible regimes. In fundamental physics, recently there have been considerable interests and intensive efforts in exploiting such NEMS devices for quantum measurements. In technology development, significant milestones have been achieved using prototypical devices and systems, and rapid progresses are currently being made in resonant sensing; plus other interesting applications are emerging. At the juncture from understanding fundamentals to creating practical technologies, there are great challenges and also enormous opportunities for the *engineering* of NEMS resonators and oscillators.

This tutorial will give an introduction to fundamentals of NEMS and an overview of recent efforts in developing and engineering nanoscale electromechanical resonators and oscillators. After discussing the basic ideas and fundamental aspects of such devices, I shall introduce prototypes of devices enabled by the state-of-the-art nanofabrication techniques, including both top-down and bottom-up approaches. Related and enabling materials, along with their interesting properties and relevant challenges, will also be discussed. I will then discuss device characterization techniques and introduce recent advances in very-high and ultra-high frequency (VHF/UHF) NEMS resonators and oscillators, along with important challenges such as noise and integration. Finally, I will review and discuss some representative areas of applications, latest highlights and new possibilities in this field, and their interesting and promising interactions with a few of today's mainstream technologies.

4:00 p.m. – 6:00 p.m.
Room: Seacliff A

Next-Generation Frequency Standards
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Warrington.pdf)
Bruce Warrington, National Measurement Institute

Optical frequency standards have reached systematic uncertainties of a few parts in 10^17, equivalent to an error of a few seconds accumulated over the lifetime of the universe, and are able to detect time dilation caused by a change in height of less than one metre in the Earth's gravitational potential. Next-generation microwave frequency standards are set to fly aboard the International Space Station, testing relativistic red-shift to a few parts per million and transferring time to ground stations with picosecond-level precision. How are these astonishing feats of ultra-precise measurement achieved? The presentation will cover a short history, some core techniques, and the two main approaches to realising an optical frequency standard: single trapped ions and neutral atoms in optical lattices. Space-borne microwave frequency standards and linear ion traps will also be briefly described, as well
as the challenges in comparing standards across large distances. To conclude, we will
speculate on a redefinition of the SI second, currently defined with reference to cesium,
and on the frontier for the measurement of time and frequency.

4:00 p.m. – 6:00 p.m.
Room: Seacliff B

The Kalman Filter
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/Slides_Galleani.pdf)
Lorenzo Galleani, Politecnico di Torino

The motion of a satellite, the trajectory of an airplane, the position of a boat: all these quantities can be predicted accurately and efficiently with a Kalman filter. Since its introduction half a century ago, the Kalman filter has revolutionized the field of applied optimal estimation. Literally hundreds of applications have been made and the number is still growing. Also atomic timing benefits greatly from the Kalman filter. GPS time, for instance, is generated by the Composite Clock, an algorithm based on the
Kalman filter. In this tutorial we do two things. First, we define the Kalman filter by using a simple example, namely, the motion of a boat. Second, we discuss a few applications of the Kalman filter to atomic timing, including the GPS Composite Clock.

4:00 p.m. – 6:00 p.m.
Room: Seacliff C

Theory and Analysis of Quartz Crystal Resonators
(http://ifcs-eftf2011.org/sites/ifcs-eftf2011.org/files/editor-files/FCSTutorial_2011_Resonator.pdf)
Yook-Kong Yong, Rutgers University
Ji Wang, Ningbo University

Quartz crystal resonators as a key element for frequency control and detection have wide applications in modern electronics. The manufacturing technology of quartz crystal resonators have under gone tremendous changes in last few decades with focused efforts on accuracy, stability, miniaturization and novel applications. These changes required extensive investigations on improving the theory and analytical techniques as represented by improvements in the Mindlin plate equations and the development of finite element analysis aimed at the analysis and design of novel structures of resonators. The critical roles of traditional quartz crystal resonators have been targeted to be replaced by newer technologies such as the surface acoustic wave (SAW) resonators, acoustic wave MEMS, and lately the film bulk acoustic wave resonators (FBAR), but the sophistication of quartz crystal resonator technology itself has demonstrated the continuing and sustainable presence of the traditional quartz resonator with improved performance and refined structure for unique roles in many critical applications.

This tutorial will provide an overview of the needs of future development, design, and research of the quartz resonator technology through an introduction of the material, basic theory, approximate equations, practical methods, and design of the traditional resonator. The lecture will be presented to cover the following: 1) History and trends of quartz crystal resonator technology, 2) Basic theory of wave propagation, 3) Quartz crystal material, 4) Mindlin plate equations, 5) Analytical considerations, 6) Finite element methods.

Panel Discussion

Subject: To What Extent Will Quartz be Squeezed Out By CSAC and MEMS?

Tuesday, May 3rd, 10:30 a.m. – 12:30 p.m.
Grand Ballroom B/C

Panelists:

Chair: John Vig, Consultant, USA
Jean-Pierre Aubry, Oscilloquartz, Switzerland
Martin Bloch, Frequency Electronics, Inc., USA
John Kitching, NIST, USA
Amit Lal, Cornell U., USA
Robert Lutwak, Symmetricom, Inc., USA
Gaetano Mileti, The LTF laboratory - University of Neuchâtel, Switzerland
Clark Nguyen, UC Berkeley, USA

During the past several decades, many billions of quartz crystal oscillators have been produced. The applications have ranged from inexpensive clock oscillators to high stability devices for military and space applications. No technology has been able to seriously compete with quartz - until recently. Now, competition to quartz is emerging. MEMS oscillators are now competing with quartz at the low end, and, recently, the Chip Scale Atomic Clock (CSAC) has become available to compete with some applications at the high end.

At last year's Symposium, a panel of experts, together with an enthusiastic audience, debated the question, "To What Extent Will MEMS Replace Conventional Quartz Resonators?" This year, a panel of experts will continue and extend the debate to the question, "To What Extent Will Quartz Be Squeezed Out by CSAC and MEMS?" The panelists, who include experts in atomic clocks, MEMS, and conventional quartz technologies, will compare the advantages and disadvantages of the three technologies. Brief updates on the state of MEMS oscillator technologies and markets will be provided at the start of the panel discussion.

Future Conferences

2012 IEEE International Frequency Control Symposium
19 - 24 May 2011 – Baltimore, Maryland, USA
www.ifcs2012.org

EFTF 2012 European Frequency & Time Forum
23-27 April 2012
Chalmers University of Technology, Gothenburg, Sweden
http://www.eftf2012.org/EFTF2012Home.htm

Special Thanks

The 2011 Joint Conference of the IEEE International Frequency Control Symposium and European Frequency and Time Forum is made possible with help from:

Sponsors

IEEE Ultrasonics, Ferroelectrics and Frequency Control Society

European Frequency and Time Forum

Diamond Patrons

National Institute of Standards and Technology (NIST)

Jet Propulsion Laboratory (JPL)

Sandia

Gold Patron

Symmetricom

Tote Bag Patron

Frequency Electronics

Special Assistance

University of California at Berkeley

Exhibitors

ACS is the proud producer of Lattice Etch, the only non-aqueous HF solution specifically formulated for the etching (chemical milling and/or polishing) of crystalline quartz available on the market. Lattice Etch provides superior etching performance with consistent and predictable results. Used by leaders in the frequency control market for years to produce high frequency fundamental quartz crystals.

http://adv-chem.com/

AMSystems designs and manufactures thin film processing equipment for the industry, research labs and universities. Our equipment ranges from single wafer R&D modules to production cluster tools. We provide a complete solution for FBAR/SAW technologies, including thin film deposition and thickness uniformity trimming. AMSystems can help build early prototypes utilizing our dedicated equipment.

http://www.amssb.com

Anapico Ltd

Switzerland based Anapico Ltd is a innovative manufacturer of test equipment such as signal source analyzers and low phase noise RF and microwave signal generators. The displayed APPH6000 signal source analyzer measures phase noise down to -190 dBc/Hz.

http://www.anapico.com/

Brandywine Communications is the premier supplier of GPS and time code based precision time and frequency products including wireline and wireless telecommunication networks. Brandywine offers a full suite of Next Generation synchronization, timing and monitoring products to maximize a telecom network's performance and flexibility to the network. Better timing and synchronization means better quality for voice, video and data services.

http://www.brandywinecomm.com/

Brilliant Instruments

Time interval analyzers with 3 ps resolution and 4 million measurements per second in modular instrument formats.

http://www.b-i-inc.com/

Discovery Semiconductors, Inc. is an industry leader in manufacturing ultrafast, high optical power handling InGaAs photodetectors, RF over fiber optical receivers, balanced optical receivers and several custom products for applications ranging from analog RF links to ultrafast digital communications. Discovery's instrumentation includes their Lab Buddy, Highly Linear Photodiode (HLPD), and Optical Coherent Receiver System.

http://www.discoverysemi.com

EndRun TECHNOLOGIES

"Smarter Timing Solutions"

EndRun Technologies is the only provider of indirect-GPS disciplined oscillators. Designed and manufactured in California, EndRun synchronizes time and frequency systems using GPS or indirect-GPS (indoor antenna with IS-95 CDMA). A wide variety of products are available spanning from GPS fiber optic links, GPS inline pre-amplifiers, NTP time servers, IEEE-1588 grandmaster clocks, primary time & frequency references and distribution. The gem of the product line includes the ultra-low-noise oscillator that provides the most precise and stable GPS-derived frequency reference in the world. Guaranteed.

http://www.endruntechnologies.com/

Frequency Electronics, Inc. is a world leader in the design, development and manufacture of high precision timing, frequency control and synchronization products for space and terrestrial applications. Frequency Electronics' products are used in commercial, government and military systems, including satellite payloads, missiles, UAVs, aircraft, GPS, secure radios, SCADA, energy exploration and wireline and wireless communication networks.

http://www.freqelec.com/

GuideTech is a leading provider of precision test and measurement instruments for scientific laboratories, development and volume manufacturing applications based on its patented "Continuous Timing Interval Analyzer" (CTIA) technology. GuideTech's multi-channel timing test systems substantially improve test throughput of Automated Test Equipment (ATE) systems through fast measurement rates, greater accuracy, increased parallelism, and through critical high-speed timing test coverage. The Company sells its PC-based timing instruments since 1988 into the scientific lab market to customers such as NIST, NASA, CERN, US Naval Observatory, Northrop, Rolls Royce and other respected science organizations worldwide. GuideTech's PC products and test systems are available in ISA, PCI, PCIe, PXI, and PXIe. With very strong R&D, the company is scheduled to launch 9 new products in 2011.

http://www.guidetech.com

Jackson Labs Technologies, Inc. designs and manufactures Precision Timing and Frequency References including ground-breaking Chip Scale Atomic Oscillators (CSACs) and ultra-miniature GPS Disciplined Oscillators. Our smallest product is less than 1 inch square and consumes less than 0.5W, and our highest performance product is slightly larger than a pack of cigarettes, and is based on a Cesium Vapor Atomic Reference for ultra-fast warm-up, very low power, and very good holdover performance.

http://www.jackson-labs.com/

Noise XT is a high spectral purity instrumentation provider offering Phase and Amplitude Noise Test Systems. Our products measure Absolute, Residual, Added noise on CW and Pulsed signals with offsets up to 500MHz and on input signals up to 140GHz. Our new product DCNTS exhibits the lowest noise floor of the industry with down-to -195dBc/Hz performance. It is the only complete dual phase noise test set with cross-correlation feature. Its unique signal processing software allows very fast high performance measurements.

http://noisext.com/

Pascall Electronics Ltd based on the Isle of Wight in England (UK) is a leading supplier of RF & Microwave components, subsystems and custom power supplies for aerospace, defence, marine and civil markets. Pascall has special expertise in providing industry leading ultra low noise oscillators and frequency generators solutions covering VHF to X-band. Such devices are particularly suited for radar, communication and high end test systems

http://www.pascall.co.uk/

Piktime Systems Ltd specializes in: *time transfer solutions for time scales purpose, *high precision measurements of frequency and time, *distribution of time signals and reference frequency. Our core product is time transfer system TTS-4 which allows comparison of time scales and clocks on long distance. TTS-4 receiver is capable to carry observations on all available frequencies for GPS, GLONASS and GALILEO. Our product offers high precision along with comfortable and stable operation. Piktime Systems Ltd closely co-operates with The Astrogeodynamical Observatory (AOS) of the Space Research Centre.

http://www.piktime.com/

Roth & Rau MicroSystems is a leading supplier of plasma and ion beam processing systems for the semiconductor, MEMS, and sensor industries. Our “IonScan” frequency trimming systems are widely used in SAW, BAW, and FBAR manufacturing for yield maximization. We also offer a full range of precision thin film deposition and etching systems for a broad range of processes, including Reactive Ion Etching (RIE), Plasma Enhanced Chemical Vapor Deposition (PECVD), Ion Beam Etching (IBE) and Ion Beam Sputter Deposition (IBSD), using our own in-house developed plasma and ion beam sources.

http://www.roth-rau.de/konzern/microsystems/de/index.php

Sansei-Showa Co., Ltd produces production and test equipment for the Quartz Crystal and Semiconductor industry.

http://www.sanseishowa.com/

"Industry leader of test and manufacturing equipment for the frequency control industry." We make measurement, temperature test, frequency adjustment, sorting and various other production equipment for the frequency control industry.

http://saunders-assoc.com

Spectracom is a leading provider of time and frequency technology. Our application-specific products are designed for synchronization and timing of networks, systems and devices; calibration and test of precise frequency instruments, and measurement of frequency signals. Products include sync tester/analyzers, GPS simulators, time and frequency analyzers/counters, secure master clock synchronization systems, and more.

http://www.spectracomcorp.com/

SpectraDynamics, Inc. is a leading supplier of Time and Frequency Products with exceptionally low noise and high stability. Our main product lines are Time and Frequency Distribution Amplifiers, Frequency Synthesizers and Noise Measurement Systems. SpectraDynamics, Inc. is proud to serve the Time and Frequency community around the world.

http://www.spectradynamics.com/

Spectratime is a world's leader in high-performance crystal, rubidium, maser and integrated GPS/GNSS reference clocks and related test instruments, by offering smart, cost-effective and high-quality products for the telecom, defense, broadcasting, instrumentation, and space industries.

http://www.spectratime.com/

Symmetricom's precision frequency standards deliver exceptional performance and unsurpassed accuracy, stability and reliability in the most challenging of environments. Our hydrogen, cesium and rubidium standards and quartz oscillators have supported more military communications, satellite ground stations, metrology laboratories and test & measurement applications than any other precision frequency references in the world.

http://www.symmetricom.com/

T4Science is a world's leader in high-performance hydrogen maser clocks, by offering smart, cost-effective and high-quality products and total support services for the multiple scientific and GNSS space industries.

http://www.t4science.com/

XECO, Inc. is a manufacturer of quartz inverted mesa blanks offering many styles and sizes of blanks both electroded and non-electroded from frequencies of 20 to over 600 MHz fundamentals. For unique applications requiring special electrode patterns or mesa shapes, XECO can produce blanks to your specifications with minimal tooling fees. XECO exclusively uses swept quartz and a unique photolithographic process to produce blanks and can offer the highest quality blank available on the market from an independent supplier.

http://www.xeco.net/

Technology-leader Zurich Instruments (ZI) designs and manufactures high performance dynamic signal analysis instruments for advanced scientific research and leading industrial applications. ZI products include lock-in amplifiers, phase-locked loops, instruments for electrical impedance spectroscopy, and application specific pre-amplifiers. Headquartered in Zurich, Switzerland, ZI was established in 2008. ZI customers are scientists and engineers in leading research labs and organizations worldwide.

http://www.zhinst.com/

Plenary Session Schedule

Welcome, Acknowledgements, & Information
(Clark Nguyen, Giorgio Santarelli, General Co-Chairs)

Technical Program Information
(Eric Burt, Patrick Gill, JPC Co-Chairs)

2012 IEEE-IFCS Announcement

2012 EFTF Announcement

IFCS IEEE Awards
(Andreas Bauch, Awards Chair)

Rabi Award

Cady Award

Sawyer Award

EFTF Award

EFTF Young Scientist Award

EFTF Marcel Ecabert Award

Acknowledgement of Recent IEEE Fellows
(Mike Garvey)

Introduction of Plenary Speaker and Plenary Presentation
(Eric Burt, Dave Wineland)

Logistical Conference Information
(Clark Nguyen)

Adjourn

Plenary Session Invited Talk

Quantum Information Processing and Atomic Clocks*

D. J. Wineland, NIST, Boulder, CO

Abstract:
For approximately the last fifteen years, the subject of quantum information processing (QIP) has occupied the efforts of many laboratories, leading to logic gate implementations and demonstrations of basic protocols for universal quantum computation. The potential strength of quantum computing derives in part from the exponentially-increasing size of state space with the number of bits - when such "qubits" exist in superposition states. Primary drivers for this work include the potential realization of efficient factoring (P. Shor) and unstructured searches (L. Grover), however these applications are probably at least decades away. In the near term, simple elements of QIP can be applied to simulation of interesting quantum systems or to metrology. For atomic clocks, simple logic gate protocols have been used to realize "quantum logic spectroscopy" which has led to an Al+ optical clock with systematic uncertainty less than $1x10^{-17}$ (T. Rosenband et al.). Particular entangled states created with QIP techniques can provide improved clock stability in certain circumstances. Extensions of these techniques may also have more general spectroscopic applications.

* NIST work supported by ONR, AFOSR, IARPA, NSA, DARPA, and the NIST Quantum Information Program.

Bio:
DAVID WINELAND received a bachelor's degree from Berkeley in 1965 and his Ph.D. from Harvard in 1970, under Norman Ramsey. After a postdoctoral appointment at the University of Washington, under Hans Dehmelt, he joined NBS (now NIST), where he is the leader of the Ion-Storage Group in the Time and Frequency Division at Boulder. The group's research has focused on laser cooling and spectroscopy of trapped atomic ions with applications to atomic clocks, quantum-limited metrology, and quantum state control.

Parallel Oral Session Schedule

Monday, May 2 10:50 a.m. - 12:30 p.m.			
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	A1L-A - Lattice Clocks I	**A1L-B - Silicon and CMOS MEMS Resonators**	**A1L-B - Time Transfer and Time Scales**
	Session Chairs:		
	Scott Diddams, NIST	**Sheng-Shiang Li, National Tsing Hua University**	**Samuel Stein, Symmetricom**
10:50a.m. - 11:10a.m.	**INVITED: Trapping Induced Frequency shifts by Comparison of Two Sr optical lattice clocks at the 10-17Level** *Jérôme Lodewyck, et al.* *LNE-SYRTE* *INRIM* *Nicolaus Copernicus University* *JPL*	**INVITED: Hybrid MEMS Resonators and Oscillators** *Sunil Bhave, Cornell University*	**Aces MWL Status and Test Results** *Marc Peter Hess, et al.* *Astrium GmbH* *Timetech GmbH* *ESA/ESTEC*
11:10a.m. - 11:30a.m.			**Single-Frequency time and frequency transfer with Galileo E5** *María Del Carmen Martínez-Belda, et al.* *University of Alicante* *Royal Observatory of Belgium*
11:30a.m. - 11:50a.m.	**Higher-Order Effects on the Clock Transition 1S0 – 3P0 in Strontium Atoms** *Vitaly Pal'chikov, et al.*	**A CMOS-compatible 20MHz Poly-SiGe MEMS oscillator with low-power heating for stabilization over temperature.** *Roelof Jansen, et al.* *IMEC*	**New realization strategy for UTC(PTB** *Andreas Bauch, et al.* *PTB* *NUDT*
11:50a.m. - 12:10p.m.	**Reducing the Uncertainty of the Yb Optical Lattice Clock** *Andrew Ludlow* *NIST* *ECNU*	**Mechanically-Coupled CMOS-MEMS Free-Free Beam Resonator Arrays with Two-Port Configuration** *Ming-Huang Li, et al.* *Institute of NanoEngineering and MicroSystems* *National Tsing Hua University*	**Contribution to TAI frequency by a travelling Primary Frequency Standard** *Daniele Rovera, et al.* *LNE-SYRTE*
12:10p.m. - 12:30p.m.	**Transition Frequency of the 698 nm Clock Transition of ^{87}Sr** *Stephan Falke, et al.* *Physikalisch-Technische Bundesanstalt*	**Bulk Modes in Single-Crystal Silicon** *Gavin Ho, et al.* *ParibX, Inc* *University of Denver* *Georgia Institute of Technology*	**Direct comparison of optical clocks Using a telecommunication fiber link in Tokyo** *Miho Fujieda, et al.* *National Institute of Information and Communications Technology* *The University of Tokyo*

Parallel Oral Session Schedule (continued)

Monday, May 2 1:30 p.m. - 3:30 p.m.			
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	A2L-A - Fountain Clocks	**A2L-B - New Material Applications**	**A2L-C - Phase Noise Measurement and Modeling**
	Session Chairs:		
	Robert Lutwak, Symmetricom	**Ralf Lucklum, Tech. Univ. of Magdeburg**	**Warren Walls, US Naval Observatory**
1:30p.m. - 1:50p.m	**INVITED: Measuring the Blackbody Radiation Shift – Comparisons between Room Temperature and Cryogenic (77K) Cs Fountains at INRIM and NIST** *T.P. Heavner, et al.* *NIST* *INRIM* *Politecnico di Torino*	**Piston Temperature Measurements with SAW Sensors***Thomas Plum, et al.* *Deutz AG* *SENSeOR*	**INVITED: Modeling Phase Noise in Multifunction Subassemblies** *Michael Driscoll, et al.* *Northrop Grumman Electronic Systems*
1:50p.m. - 2:10p.m.		**Characterization of WO_3 layers deposited on Quartz and Lithium Niobate SAW Resonators for the design of gas sensors** *M. David, et al.* *Université deToulon var* *Université Femto Besancon*	
2:10p.m. - 2:30p.m.	**Uncertainty evaluation of the continuous cesium fountain frequency standard FOCS-2** *Gianni Di Domenico, et al.* *Université de Neuchâtel* *METAS*	**Langasite SAW Tmperature and oxygen multi-sensor** *P. Zheng, et al.* *National Energy Technology Laboratory* *Carnegie Mellon University*	**Vibration Isolation of Acceleration-Sensitive Devices** *Joseph Donovan, et al.* *Northrop Grumman Corporation*
2:30p.m. - 2:50p.m.	**Improved accuracy evaluation of the NPL-CsF2 primary frequency standard** *Ruoxin Li, et al.* *Pennsylvania State University* *National Physical Laboratory*	**SAW Sensor Correlator System Performance Parameters** *Nikolai Kozlovski, et al.* *University of Central Florida*	**Limited Live-Time Measurements of Frequency Spectra** *David Howe, et al.* *NIST*
2:50p.m. - 3:10p.m.	**Measurements with Multiple Operational Fountain Clocks** *Steven Peil, et al.* *U.S. Naval Observatory*	**Thin Film Plate Acoustic Wave Based Resonant (FPAR) Pressure Sensors** *Emil Anderås, et al.* *Uppsala University*	**Simplified Phase Noise Model for Negative Resistance Oscillators** *Jeremy Everard, et al.* *University of York*
3:10p.m. - 3:30p.m.	**Quantitative Evaluation of Distributed Cavity Phase Shifts to Improve the Accuracy of SYRTE FO2** *Jocelyne Guéna, et al.* *LNE-SYRTE* *The Pennsylvania State University*	**Elasto-Acoustic Properties of ReCa4O(BO3)3 (Re=La, Pr, Nd, Y. Gd) Piezoelectric Crystals** *Thomas Shrout, et al.* *Pennsylvania State University* *Crystal Photonics Inc.*	**An integrated phase noise measurement bench for the on-chip characterization of resonators and VCOs** *Sylvain Godet, et al.* *Axess Europe* *LAAS-CNRS* *ST Microelectronics*

Parallel Oral Session Schedule (continued)

Tuesday, May 3 8:30 a.m. - 11:00 a.m.			
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	B1L-A - Ion Clocks	**B1L-B - Novel structures for MEMS Resonators**	**B1L-C- Two Way Optical and Microwave Links**
	Session Chairs:		
	Eric Burt, JPL/NASA	**Gianluca Piazza, University of Pennsylvania**	**Mizuhiko Hosokawa, NICT**
8:30a.m. - 8:50a.m.	**Precision Calculation of Blackbody Radiation shifts for Metrology at the 18th Decimal Place** *Marianna Safronova, et al.* *University of Delaware* *PNPI* *Joint Quantum Institute*	**INVITED: Ferroelectric PZT RF MEMS Resonators** *Jeffrey Pulskamp, et al.* *US Army Research Laboratory* *Cornell University*	**INVITED: Time transfer by laser link - T2L2: Current Status and Future Experiments** *Etienne Samain, et al.* Observatoire de la Côte d'Azur *CNES*
8:50a.m. - 9:10a.m.	**Precision Calculation of Blackbody Radiation shifts for Metrology at the 18th Decimal Place** *Nils Huntemann, et al.* *Physikalisch-Technische Bundesanstalt*		
9:10a.m. - 9:30a.m.	**Search for time-evolution of the fine structure constant using two optical transitions in a single Ion of 171Yb+** *Steven King, et al.* *National Physical Laboratory*	**Ultra High Frequency Air/Aluminum Nitride Fractal Phononic Crystals** *Nai-Kuei Kuo, et al.* *University of Pennsylvania*	**Construction Progress in the Photon Counting Detector for the European Laser Timing Experiment** *Jan Kodet, et al.* *Czech Technical University in Prague* *CSRC*
9:30a.m. - 9:50a.m.	**Experiments with trapped thorium ions: Towards an optical nuclear clock with Th-229** *Ekkehard Peik, et al.* *Physikalisch-Technische Bundesanstalt*	**A new acoustic resonator concept based on acoustic waveguides using Silicon/Periodically Poled Transducer/Silicon structures for RF applications** *Florent Bassignot, et al.* *FEMTO-ST* *DGA*	**TWSTFT Experiment by Employing Dual Pseudorandom Noise (DPN) Codes between NICT and TL** *Shinn-Yan Lin, et al* *Chunghwa Telecom* *National Tsing Hua University*
9:50a.m. - 10:10a.m.	[redacted]	**High-Q Factor in electrostrictive thin Fflms resonators** *Mathieu Pijolat, et al.* *CEA LETI* *FEMTO-ST*	**Transatlantic 2.5 MChip/S Two-Way Satellite Time and Frequency Transfer with Surface Acoustic Wave Filters** *Victor Zhang, et al.* *NIST* *LNE-SYTRE*
10:10a.m. - 10:30a.m.		**Motional Impedance of Resonators in the Nonlinear Regime** *Hyung Kyu Lee, et al.* *Stanford University*	**Dual-Stage Quad-Mixer Satellite Simulator for TWSTFT** *Faisal Ali Mubarak, et al.* *Dutch Metrology Institute*

Parallel Oral Session Schedule (continued)

Tuesday, May 3 1:30 p.m. - 3:30 p.m.			
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	B2L-A - Combs	**B2L-B - Liquid Sensors**	**B2L-C- Frequency Synthesis & CMOS Oscillators**
	Session Chairs:		
	Scott Crane, USNO	**Fabien Josse, Marquette University**	**Jeremy Everard, University of York**
1:30p.m. - 1:50p.m.	**INVITED: The evolving optical frequency comb: Expanding applications for a precision optical frequency synthesizer** *Scott Diddams, et al.* *NIST*	**Surface Functionalization and Monolayer Formation on Silicon Resonant Nanoballances** *Babak Tousifar, et al.* *University of Denver*	**INVITED: A History of the Development of CMOS Oscillators: the Dark Horse in Frequency Control** *Michael McCorquodale, et al.* *Integrated Device Technology*
1:50p.m. - 2:10p.m.		**Damping and Mass Sensitivity of Laterally Vibrating Resonant Microcantilevers in Viscous Liquid Media** *Russell Cox, et al.* *Marquette University* *Georgia Institute of Technology* *Universite de Bordeaux*	
2:10p.m. - 2:30p.m.	**Carrier-envelope phase stabilization with as little As 120 pJ Laser Pulse Energy (10 mW at 80 MHz) with a new method** *Peter Hommelhoff, et al.* *Max-Planck-Institut für Quantenoptik*	**Liquid Sensor Utilizing a Regular Phononic Crystal with Normal Incidence of Sound** *Ralf Lucklum, et al.* *Otto-von-Guericke-University* *Wuhan University*	**A Highly Stable CMOS Self-Compensated Oscillator (SCO) Based on an LC Tank Temperature Null Concept** *Ayman Ahmed, et al.* *Si-Ware Systems*
2:30p.m. - 2:50p.m.	**INVITED: Chip Scale Optical Frequency Combs** *Tobias Kippenberg, et al.* *EPFL* *MPQ*	**Influence of Ambient Parameters on the Response of Polymer-Coated Sh-Surface Acoustic Wave Sensors to Aromatic Analytes in Liquid-Phase Detection** *Florian Bender, et al.* *Marquette University* *Stanford University*	**Design and Performance of the RF Synthesiser for the ELISA cryocooled Sapphire Oscillator** *Wolfgang Schaefer, et al.* *TimeTech GmbH* *FEMTO-ST* *European Space Agency*
2:50p.m. - 3:10p.m.		**Surface Acoustic Wave Flow Sensor** *Yizhong Wang, et al.* *University of Pittsburgh* *Peking University*	**A 4.6 GHz, 15 mW Frequency Synthesizer CMOS ASIC with Milli-Hertz Frequency Resolution for Miniature Atomic Clocks** *Yazhou Zhao, et al.* *Ecole Polytechnique Fédérale de Lausanne* *ASULAB*
3:10p.m. - 3:30p.m.	**Low-Noise Near-Infrared Optical Frequency Comb from a Solid-State Femtosecond Laser** *Stéphane Schilt, et al.* *Universite de Neuchatel* *ETH Zurich*	**Solidly Mounted Resonators with Carbon Nanotube Electrodes for Biosensing Applications** *L. Garcia-Gancedo, et al.* *University of Cambridge* *Universidad Politecnica de Madrid* *University of Bolton* *University of Cambridge*	**A Novel High Frequency Synthesizer Using Injection-Mode Coupled VCSO for Low Jitter and Low Noise Applications** *Ulrich Rohde, et al.* *University of Cottbus* *Synergy Microwave Corp*

Parallel Oral Session Schedule (continued)

	Wednesday, May 4 8:30 a.m. - 10:30 a.m.		
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	C1L-A - : Local Oscillators & Resonators	**C1L-B - Optimized and Tunable MEMS Resonators**	**C1L-C- Celebrating Kalman Filter 50th Anniversary**
	Session Chairs:		
	Patrick Gill, NPL	**Yoonkee Kim, US Army**	**Demetrios Matsakis, USNO**
8:30a.m. - 8:50a.m.	**INVITED: Sensitivity Training for the Laser Interferometer Space Antenna (LISA) Mission** *William Klipstein, et al.* *California Institute of Technology*	**3.2 GHz AlN Lateral Overtone Bulk Acoustic Resonator** *Songbin Gong, et al.* *University of Pennsylvania*	**INVITED: Applications of Kalman Filtering from 1960 to Present** *Mohinder Grewal, California State University*
8:50a.m. - 9:10a.m.		**Tunable Silicon Bulk Acoustic Resonators with Multi-Face AlN Transduction** *Roozbeh Tabrizian, et al.* *Georgia Institute of Technology*	
9:10a.m. - 9:30a.m.	**Laser stabilization at the 10-16 level** *Yanyi Jiang, et al.* *East China National University* *NIST* *University of Colorado*	**Voltage-Controlled Phase Matching to Optimize Micromechanical Array-Composite Resonator Output Power** *Mehmet Akgul, et al.* *University of California at Berkeley*	**Reduced Kalman Filters for Clock Ensembles** *Charles Greenhall, California Institute of Technology*
9:30a.m. - 9:50a.m.	**Robust, portable ultra-stable lasers for non-laboratory applications** *David Leibrandt, et al.* *NIST*	**Acoustic Bragg Reflectors for Q-Enhancement of Unreleased MEMS Resonators** *Wentao Wang, et al.* *MIT*	**A Kalman Filter UTC(k) prediction and steering algorithm** *John Davis, et al.* *National Physical Laboratory*
9:50a.m. - 10:10a.m.	**Force-insensitive Fabry-Perot etalon** *Stephen Webster, et al.* *National Physical Laboratory*	**Tunable Piezoelectric MEMS Resonators for Real-Time Clock** *Diego Emilio Serrano, et al.* *Georgia Institute of Technology*	**Synchronizing Computer Clocks using Kalman Filters** *Judah Levine, NIST*
10:10a.m. - 10:30a.m.	**Frequency-stabilization to $6x10^{-16}$ via spectral-hole burning** *Michael Thorpe, et al.* *NIST*	**Linear Acoustic Bandgap Arrays for Spurious Mode Suppression in Piezoelectric MEMS Resonators** *Logan Sorenson, et al.* *Georgia Institute of Technology*	**Using the Kalman filter to detect time and frequency jumps in atomic clocks** *Lorenzo Galleani, et al.* *Politecnico di Torino* *INRIM*

Parallel Oral Session Schedule (continued)

	Wednesday, May 4 10:50 a.m. - 12:30 p.m.		
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	C2L-A - : Microclocks & Novel Concepts	**C2L-B - Quartz Based Resonators**	**C2L-C- Algorithms for Clock and Time Scale Estimations**
	Session Chairs:		
	John Kitching, NIST	**Ji Wang, Ningbo University**	**Ilaria Sesia, INRIM**
10:50a.m. - 11:10a.m.	**Dark Line Resonances in Cs-Ne microcells for Chip Scale Atomic Clocks** *Rodolphe Boudot, et al.* *FEMTO-ST*	**An Efficient AT-Cut Quartz Crystal Resonator Design Tool for Activity Dip in Working Temperature Range** *Shih-Yung Pao, et al.* *TXC Corporation*	**Straightforward estimations of GNSS on-board clocks** *Jerome Delporte, et al.* *CNES*
11:10a.m. - 11:30a.m.	**All-Optical Integrated Rubidium Atomic Clock** *Lute Maleki, et al.* *OEwaves Inc* *The Aerospace Corp*	**Recent Investigations on BAW Resonators at Cryogenic Temperatures** *Maxim Goryachev, et al.* *FEMTO-ST*	**Statistical biases and very long term time stability analysis** *Francois Vernotte, et al.* *University of Franche-Comte*
11:30a.m. - 11:50a.m.	**Low-power chip-scale Rubidium Plasma Light Source for Miniature Atomic Clocks** *Vinu Venkatraman, et al.* *École Polytechnique Fédérale de Lausanne* *University of Neuchâtel*	**A Parametrically Driven Quartz UHF Oscillator** *Randall Kubena, et al,* *HRL Laboratories* *Rutgers University*	**Multi-Scale Clock Ensembling Using Wavelets** *Ken Senior, et al.* *U.S. Naval Research Laboratory* *University of Washington*
11:50a.m. - 12:10p.m.	**MOT Loading Enhancement with Stimulated Light Forces** *Tara Liebisch, et al.* *NIST*	**Results on 40MHz Quartz Resonators Realized by Deep Reactive Ion Etching….267** *Jean-Jacques Boy, et al.* *FEMTO-ST*	**A new prediction algorithm for EAL** *Gianna Panfilo, et al.* *International Bureau for Weights and Measures*
12:10p.m. - 12:30p.m.	**CPT pump-probe measurement of the Cs clock transition DC Stark shift** *Jean-Luc Robyr, et al.* *University of Fribourg*	**Quartz Crystal Industry of China in the Crossroads** *Ji Wang, et al.* *Ningbo University* *Piezoelectric Crystal Association of China* *TXC (Ningbo) Corporation* *East Crystal Electronic Co., Ltd* *Ningbo Hiking Electronics Tech. Co., Ltd* *Timemaker Crystal Technology Co., Ltd* *203 Institute of the Second Academy* *Zhejiang University*	**Performance Comparison of Composite Clock Algorithms Based on Future GPS Clock Scenarios** *Matthias Suess, et al.* *DLR* *JPL* *USNO*

Parallel Oral Session Schedule (continued)

	Wednesday, May 4 1:50 p.m. - 3:50 p.m.		
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	C3L-A - : Lattice Clocks II	**C3L-B - Physical Sensors**	**C3L-C- Acoustic Oscillators and Micromechanical Resonators**
	Session Chairs:		
	Ekkehard Peik, PTB	**Don Malocha, University of Central Florida**	**Michael Driscoll, Northrop Grumman**
1:50p.m. - 2:10p.m.	**INVITED: Suppression of collisional frequency shifts in an optical lattice clock** *Matthew Swallows, et al.* *NIST*	**INVITED: Two Types of Micromachined Vibratory Gyroscopes** *Andrei Shkel, et al.* *DARPA/MTO*	**RF oscillators stabilized by temperature compensated HBARs based on $LiNbO_3$/Quartz Combination** *Thomas Baron, et al.* *FEMTO-ST*
2:10p.m. - 2:30p.m.			**Passive Tuning in Lateral-Mode Thin-Film Piezoelectric Oscillators** *Mohsen Shahmohammadi, et al.* *Oklahoma State University*
2:30p.m. - 2:50p.m.	**Evidence of a fermionic collisional shift** *Wilfried Maineult, et al.* *LNE-SYRTE* *Laboratoire Kassler Brossel* *The Pennsylvania State University*	**Ultrasonic Microparticle Trapping by Multi-Foci Fresnel Lens** *Youngki Choe, et al.* *University of Southern California* *Palos Verdes Peninsular High School*	**High Frequency Dual-Mode Thermal-Piezoresistive Oscillators** *Amir Rahafrooz, et al.* *University of Denver*
2:50p.m. - 3:10p.m.	**INVITED: Development of Optical Clocks for Future Applications in Space** *Uwe Sterr, et al.* *Physikalisch-Technische Bundesanstalt* *Heinrich-Heine-Universität Düsseldorf*	**Resonant Piezoelectric Energy Converters and Cardiological Health** *Rui Zhang, et al.* *Case Western Reserve University*	**Wafer-scale packaging for FBAR-based Oscillators** *Martha Small, et al.* *Avago Technologies* *University of Washington*
3:10p.m. - 3:30p.m.		**Influence of non-ideal clamping in microcantilever resonant frequency estimation** *Ludivine Fadel-Taris, et al.* *Université de Bordeaux* *Marquette University Milwaukee* *Georgia Institute of Technology Atlanta* *CNRS, LAAS Laboratory*	**Low Phase Noise Quartz BAW Oscillator for Space Applications** *T. McClelland, et al.* *Frequency Electronics, Inc.*
3:30p.m. - 3:50p.m.	**Measurement of the magic wavelength for the ^{199}Hg optical clock transition** *Sinda Mejri, et al.* *LNE-SYRTE*		

Parallel Oral Session Schedule (continued)

	Wednesday, May 4 4:10 p.m. - 6:10 p.m.		
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	C4L-A - : Space Clocks & Novel Oscillators	**C4L-B -Mass Sensors**	**C4L-C- Optical Fiber Time and Frequency Transfer**
	Session Chairs:		
	Louis Marmet, NRC Canada	**Leo Reindl, University of Freibourg**	**Michael Driscoll, Northrop Grumman**
4:10p.m. - 4:30p.m.	**Study of Fe^{3+}-Sapphire Maser Above 4 Kelvin** *Karim Benmessai, et al.* *University of Western Australia* *FEMTO-ST*	**INVITED: Vibratory MEMS for room temperature electrometry** *Ashwin Seshia, et al.* *University of Cambridge* *University of Hong Kong*	**INVITED: A 900 km long optical fiber link for remote comparison of frequency standards** *Harald Schnatz, et al.* *Physikalisch-Technische Bundesanstalt* *Max-Planck Institute for Quantum Optics*
4:30p.m. - 4:50p.m.	**Space Mini Passive Hydrogen Maser - a Compact Passive Hydrogen Maser for Space Applications** *Marco Belloni, et al.* *Selex Galileo* *Spectratime* *European Space Agency* *Kytime*		
4:50p.m. - 5:10p.m.	**Cavity pulling study and long-term stability evaluation of HORACE compact cold atom clock** *Nicolas Rossetto, et al.* *LNE-SYRTE* *NIST* *CNES*	**Nanoscale Thin Plate and Membrane Resonators for Sensing** *Philip Feng, Case Western Reserve University*	**Phase Noise of Integrated Optical Phase Locked Loops** *Amir Nejadmalayeri, et al.* *MIT* *Sandia National Labs*
5:10p.m. - 5:30p.m.	**Photonically generated 10 GHz microwaves with close-to-carrier phase noise < -100 dBc/Hz** *Tara Fortier, et al.* *NIST*	**High Frequency Thermal-Piezoresistive MEMS Resonators for Detection of Organic Gases** *Arash Hajjam, et al.* *University of Denver*	**Progress on an Optical Link for Ultra-Stable Frequency Dissemination using a Public Telecommunication Network** *Olivier Lopez, et al.* *Laboratoire de Physique des Lasers* *LNE-SYRTE*
5:30p.m. - 5:50p.m.		**Receptor-Coated Porous Silicon Resonators for Enhanced Sensitivity of Vapor Detection** *Yongha Hwang, et al.* *University of California, Los Angeles*	**Active Detection of Propagation Delay Variations in Single Way Time Transfer Utilizing Dual Wavelengths in an Optical Fiber Network** *Sven-Christian Ebenhag, et al.* *SP Technical Research Institute of Sweden*
5:50p.m. - 6:10p.m.		**Effect of a mass layer on SH Waves in Piezomagnetic/Piezoelectric Material Structures** *Jing Cui, et al.* *Ningbo University*	**One-Way Temperature Compensated Fiber Link** *James Hanssen, et al.* *US Naval Observatory*

Parallel Oral Session Schedule (continued)

	Thursday, May 5 8:30 a.m. - 10:30 a.m.		
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	D1L-A - Vapor Cell Clocks	**D1L-B -Filters**	**D1L-C- Optical Oscillators & Components**
	Session Chairs:		
	Svenja Knappe, NIST	**Rick Puccio, Quartzdyne**	**Michael Tobar, University of Western Australia**
8:30a.m. - 8:50a.m.	**INVITED: Super-coatings enabling M\minute-long Zeeman coherence times in alkali-vapor Cells** *Micah Ledbetter, et al.* *University of California at Berkeley* *State Optical Institute*	**INVITED: Positioning FBAR Technology in the Frequency and Timing Domain** *Rich Ruby, Wireless Semiconductor Division*	**Experimental Demonstration of Phase-Modulated OEO Using Balance Detection** *Patrick Callahan, et al.* *Johns Hopkins University Applied Physics Laboratory*
8:50a.m. - 9:10a.m.			**Characterization of Surface Acoustic Wave Optomechanical Oscillators** *Gaurav Bahl, et al.* *University of Michigan*
9:10a.m. - 9:30a.m.	**A compact laser-pumped Rb clock with < $5x10^{-13}$ $\tau^{-1/2}$ frequency stability** *Christoph Affolderbach, et al.* *University of Neuchatel*	**INVITED: Tunable RF SAW/BAW Filters: Dream or Reality** *Ken-Ya Hashimoto, Chiba University*	**New strategies for fiber-based femtosecond lasers low-noise microwave generation** *Wei Zhang, et al.* *LNE-SYRTE* *University of Western Australia* *MenloSystems GmbH*
9:30a.m. - 9:50a.m.	**Pulsed Optically Pumped Rb Clock with High Frequency Stability Performances** *Salvatore Micalizio, et al.* *INRIM* *Laboratoire Temps Fréquence*		**Optical Scattering Induced Noise in RF-Photonic Systems** *Olukayode Okusaga, et al.* *U.S. Army Research Laboratory* *University of Maryland Baltimore County*
9:50a.m. - 10:10a.m.	**The Influence of Laser Polarization Noise on CPT Atomic Clock Signals** *James Camparo, et al.* *The Aerospace Corporation*	**Cascaded Channel-Select Filter Array Architecture Using High-K Transducers for Spectrum Analysis** *Eugene Hwang, Cornell University* *U.S. Army Research Laboratory*	**Theoretical Investigation of Optical Fiber-Length-Dependent Phase Noise in Opto-Electronic Oscillators** *Andrew Docherty, et al.* *University of Maryland Baltimore County* *U.S. Army Research Laboratory*
10:10a.m. - 10:30a.m.		**Laterally Coupled Narrow-Band High Overtone Bulk Wave Filters Using Thinned Single Crystal Lithium Niobate Layers….** *Dorian Gachon, et al.* *PROMES* *FEMTO-ST*	**Generation of Kerr Combs in MgF_2 Microresonators** *Wei Liang, et al.* *OEwaves Inc*

Parallel Oral Session Schedule (continued)

1sp

	Thursday, May 5 10:50 a.m. - 12:50 p.m.		
	Grand Ballroom B	**Grand Ballroom C**	**Seacliff Rooms**
	D2L-A - Fundamental Measurement & Sensors	**D2L-B - Materials**	**D2L-C- GPS Carrier Phase Time Comparison and Calibration**
	Session Chairs:		
	Giorgio Santarelli, LNE-SYTRE	**Bernard Dulmet, ENS2M**	**Pascale Defraigne, ORB**
10:50a.m. - 11:10a.m.	**Recommended standard frequency values for applications including the practical realisation of the metre and secondary representations of the second** *Riehle Fritz, et al.* *Physikalisch-Technische Bundesanstalt* *National Physical Laboratory* *BIPM*	**INVITED: GaN: a multifunctional material enabling MEMS resonators based on amplified piezoelectric detection** *Marc Faucher, et al.* *IEMN* *CRHEA* *MC2-technologies*	**The time stability of PPP links** *Gérard Petit, et al.* *Bureau International des Poids et Mesures*
11:10a.m. - 11:30a.m.	**Absolute Frequency Measurements of the 2S-6S,6D Transitions in Atomic Hydrogen** *Jeff Flowers, et al.* *National Physical Laboratory* *LNE-Cnam* *University of Oxford* *DK Research*		**Precision and Accuracy of USNO GPS Carrier Phase Time Transfer: Further Studies** *Christine Hackman, et al.* *United States Naval Observatory*
11:30a.m. - 11:50a.m.	**Testing the gravitational redshift with atomic clocks and atomic gravimeters?** *Peter Wolf, et al.* *LNE-SYTRE* *GRECO* *LKB*	**Hot Filament CVD Conductive Microcrystalline Diamond for High Q, High Acoustic Velocity Micromechanical Resonators** *Mehmet Akgul, et al.* *University of California at Berkeley* *sp3 Inc.*	**PPP Using NRCan Ultra Rapid products (EMU): near real-time comparison and monitoring of time scales generated in Time and Frequency laboratories** *Giancarlo Cerretto, et al.* *INRIM* *Natural Resources Canada* *LNE-SYRTE*
11:50a.m. - 12:10p.m.	**Continous g monitoring with atom interferometry** *Sébastien Merlet, et al.* *LNE-SYRTE*	**Very Low-Loss High Frequency Lateral-Mode Resonators on Polished Ultrananocrystalline Diamond** *Hediyeh Fatemi, et al.* *Oklahoma State University* *Advanced Diamond Technologies Inc.*	**Absolute Calibration of GNSS Time Transfer Systems : NRL and CNES Techniques Comparison** *Amandine Proia, et al.* *CNES* *Naval Research Laboratory*
12:10p.m. - 12:30p.m.	**Atom interferometry for inertial sensing** *Ernst M. Rasel, University of Hanover*	**$LiNbO_3$ thin single crystal layer for RF applications** *Bruno Imbert, et al.* *CEA-LETI*	**BIPM Calibration Scheme for UTC Time Links** *Zhiheng Jiang, et al.* *Bureau International des Poids et Mesures*
12:30p.m. - 12:50p.m.		**16 GHz acoustic resonator with ultra thin epitaxial Aluminum Nitride films24** *Mathieu Pijolat, et al.* *CEA-LETI* *CNRS-CRHEA* *FEMTO-ST*	**Coordinating GPS Calibrations Among NIST, NRL, USNO, PTB, and OP** *Marc Weiss, et al.* *NIST* *Naval Research Laboratory* *Observatoire de Paris* *BIPM* *Physikalisch-Technische Bundesanstalt*

Trapping induced frequency shifts by comparison of two Sr optical lattice clocks at the 10^{-17} level

J. Lodewyck, P. G. Westergaard, L. Lorini, M. Zawada, E. A. Burt, M. Gurov and P. Lemonde
LNE-SYRTE
Observatoire de Paris, CNRS and UPMC
Paris, France

Abstract— We present a comprehensive study of the frequency shifts associated with the lattice potential in Sr lattice clocks. By comparing two such clocks with a frequency stability better than 5.10^{-17} after one hour of averaging time, and varying the lattice depth up to U_0 = 900 E_r with E_r being the recoil energy, we evaluate lattice related shifts with an unprecedented accuracy.

Three different types of frequency shifts induced by the lattice have been identified as potential limitations to the accuracy of optical lattice clocks [1]. The first one is related to the non scalar feature of the atom-lattice interaction, which results from the atomic hyperfine structure [1,2]. It induces a small vector and tensor component in the confinement potential. In addition to the intensity dependence, this makes the light shift slightly dependent on the lattice polarization and geometry. We report the first observation of these effects in a Sr lattice clock and measure them with percent accuracy. We also show that the tensor effect can be used to operate a clock with ^{87}Sr in an optimal 3D lattice.

The second effect, Hyperpolarizability, refers to a shift due to higher order coupling [1]. The frequency shift scales as U_0^2 and therefore cannot be compensated for by a change of the lattice frequency. A first experimental study showed that this effect was small enough to not alter Sr clock accuracy at the 10^{-18} level for U_0=10 Er [3]. We improve on this limit by almost two orders of magnitude.

Finally, a subtle effect of higher order multipolar terms has recently been predicted [4]. Dipole magnetic (M1) and quadrupole electric (E2) interactions indeed lead to a shift of the internal atomic energy levels that is linear in lattice laser intensity. However, their spatial dependence in a lattice formed by the interference of several traveling waves does not match the one of the main dipole electric term. This in turn leads to a change of the oscillation frequency in the lattice potential wells and thereby of the spacing of the atomic external energy levels. This effect is expected to differ for both clock states. The net M1/E2 frequency shift of the clock frequency scales as $U_0^{1/2}$ and again cannot be compensated for by tuning the lattice frequency. In Ref. it is argued that this effect could be as large as 10^{-16} for U_0 = 50 E_r,! We perform the first experimental study of the M1/E2 effect showing that it is controllable to within 10^{-17} even for U_0 as large as 700 E_r.

REFERENCES

[1] H. Katori, M. Takamoto, V. Pal'chikov, and V. Ovsiannikov "Ultrastable optical clock with neutral atoms in an Engineered light shift Trap", Phys. Rev. Lett., *vol. 91*, p. 173005, 2003.

[2] M. Boyd, T. Zelevinsky, A. Ludlow, S. Blatt, T. Zanon-Willette, S. Foreman, and J. Ye, "Nuclear spin effects in optical lattice clocks" Phys. Rev. A, vol. 76, p. 022510, 2007.

[3] A. Brusch, R. Le Targat, X. Baillard, M. Fouché, and P. Lemonde, "Hyperpolarizability Effects in a Sr Optical Lattice Clock" Phys. Rev. Lett., vol. 96, p. 103003, 2006.

[4] A. Taichenachev, V. Yudin, V. Ovsiannikov, V. Pal'chikov, and C. Oates, "Frequency Shifts in an Optical Lattice Clock Due to Magnetic-Dipole and Electric-Quadrupole Transitions" Phys. Rev. Lett., vol. 101, p. 193601, 2008.

L. Lorini, M. Zawada and E.A. Burt are also affiliated to INRIM (Torino, Italy), Nicoalus Coprnicus University (Torun, Poland) and JPL (Celtech, Pasadena, USA) respectively. SYRTE is a member of IFRAF. This work has received funding from the European Community'sFP7,ERANETPlus, under Grant Agreement No. 217257, as well as from CNES and ESA (SOC project).

978-1-61284-110-6/11/$26.00 ©2011 IEEE

Improving the stability and accuracy of the Yb optical lattice clock

Y. Y. Jiang[1,2], A. D. Ludlow[1], N. D. Lemke[1], J. A. Sherman[1],
J. von Stecher[3], R. W. Fox[1], L. S. Ma[2], A. M. Rey[3], C. W. Oates[1]
[1]NIST, 325 Broadway, Boulder, CO 80305, USA
[2]East China Normal University, 3663 North Zhongshan Road, Shanghai, 200062, China
[3]JILA, NIST and University of Colorado, Boulder, CO 80309, USA
Email: ludlow@boulder.nist.gov

***Abstract*—We report results for improving the stability and uncertainty of the NIST ^{171}Yb lattice clock. The stability improvements derive from a significant reduction of the optical Dick effect, while the uncertainty improvements focus on improved understanding and constraint of the cold collision shift.**

I. Introduction

Optical lattice clocks have potential for achieving time and frequency measurement at unprecedented levels of accuracy and stability. However, to date these young systems are far from reaching this potential. Here, we describe efforts to mitigate effects which, left unresolved, pose significant obstacles to achieving these levels of performance. To improve the clock stability, we reduce the optical Dick effect. To improve the eventual accuracy of this system, we probe cold collisions between the lattice-confined atoms.

II. Reducing the optical Dick effect

The fundamental limit to the instability of an atomic frequency standard is given by the quantum projection noise (QPN). For an optical atomic clock that probes a large ensemble of quantum absorbers, the QPN limit can be quite low. For example, for a Yb or Sr lattice clock with 10^5 atoms, a 1 Hz transition linewidth, and a cycle time of $T_c = 1$ second, the QPN instability limit is below 10^{-17} at 1 s. However, while such clocks have resolved optical transition linewidths approaching 1 Hz, the instability of these systems is much higher than indicated above. Instead, the fractional frequency instability is such that it would require significantly more than 10^4 seconds to reach the 10^{-17} level. This is because these systems are usually limited by the Dick effect [1], i.e. as the atoms are periodically interrogated by the local oscillator (LO) with a period of T_c and for a probe duration of T_p, the frequency noise of the LO is downsampled onto the transition spectrum and contaminates the clock stability. In this way, the frequency noise, the probe time T_p, and the cycle time T_c all influence the magnitude and downconverted Fourier frequency of the aliased noise and thus the stability degradation. The previous operating conditions for our Yb lattice clock were $T_p = 0.08$ s and $T_c > 0.5$ s. With a frequency noise of 0.5 Hz/$\sqrt{\text{Hz}}$ at a Fourier frequency of 1 Hz, this led to a Dick-limited instability of 1-2 $\times 10^{-15}$ at 1 s.

In order to reduce this limitation, several improvements could be made: reducing the LO frequency noise that is aliased, achieving a higher duty cycle with longer probe times or reduced dead time [2], or choosing a form of spectroscopy which is less sensitive to the aliasing process (e.g. short-pulse Ramsey spectroscopy). All of these are useful from a practical perspective. However, by improving the LO laser coherence and stability, we can achieve both reduction of the LO frequency noise and longer probe times. In order to improve the LO stability, the level of Brownian thermal noise in the optical cavities used for laser stabilization must be reduced. When stabilizing a laser to an optical cavity, the thermal-noise limited laser instability dominated by the cavity mirrors [3] is given by:

$$\sigma_{therm} = \sqrt{\ln 2 \frac{4k_bT}{\pi^{3/2}} \frac{1-\sigma^2}{Ew_0L^2} \left(\phi_{sub} + \phi_{coat} \frac{2}{\sqrt{\pi}} \frac{1-2\sigma}{1-\sigma} \frac{d}{w_0} \right)} \quad (1)$$

Here, σ, E and ϕ_{sub} are Poisson's ratio, Young's modulus and the mechanical loss for the mirror substrate, and ϕ_{coat} and d denote the mechanical loss and thickness of the thin-film reflective coating. w_0 is the laser beam size on the mirror, T is the mirror temperature (K), k_B is Boltzmann's constant and L is the cavity length. In order to improve laser stabilization by reducing the thermal noise, we designed a cavity with length $L = 29$ cm, fused silica substrates with low mechanical loss ($\phi_{sub} \approx 10^{-6}$) and a somewhat large beam size using mirrors with a somewhat longer radius of curvature ($R = 1$ m). For these parameters, the thermal-noise-limited fractional frequency instability is 1.4×10^{-16}.

To reduce acceleration-induced cavity length changes, the optical cavity sits horizontally on four symmetrically placed Viton hemispheres [4], [5]. The precise support position is optimized to reduce the vertical acceleration sensitivity. Figure 1 shows the vertical acceleration sensitivity as a function of support position, indicating both finite-element simulation results as well as experimentally measured sensitivities. Simulation results are also shown for cavity acceleration sensitivity offset from the cavity optical axis. The measured sensitivity reaches as low as 7 kHz/(ms^{-2}) at the laser wavelength of 578 nm. The experimental measurement also includes weak, incidental acceleration in the horizontal dimensions. During

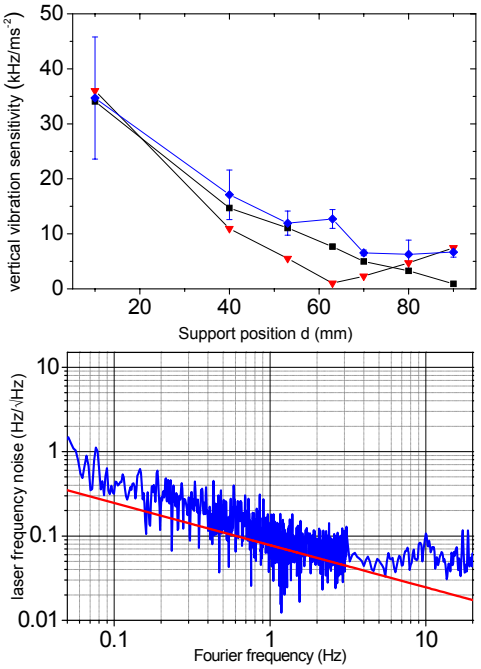


Fig. 1. top: Vertical acceleration sensitivity of the optical cavity. Experimental measurement (blue diamonds), simulation results (red triangles), simulation results for 250 μm removed from the cavity optical axis (black squares). The support position, d, is measured from where the cavity taper begins. bottom: Measured frequency noise spectrum for the laser locked to the fully-isolated cavity. The solid red line gives the theoretical prediction of the Brownian thermal noise.

normal operation, the reference cavities sat on vibration isolators.

While the use of fused silica mirrors helps reduce Brownian thermal-mechanical noise, it makes the cavity more susceptible to thermal expansion. However, a balance between the thermal expansion of the cavity components can be achieved, and we have been able to demonstrate operation of these cavities at the zero crossing of the coefficient of thermal expansion. This is done at a temperature conveniently just above ambient [6].

By comparing two independent cavity systems, we measured the frequency noise spectrum of the cavity-stabilized laser as shown in Figure 1. The red line is the thermal noise limit for each reference cavity. Up to several Hz, the laser frequency noise is close to the thermal noise limit. This low level of frequency noise in turn reduces the aliased noise contributing to the Dick effect. Furthermore, the reduced frequency noise spectrum yields laser coherence times which allow us to extend our clock probe time up to 1 second. For a more conservative operating condition of $T_p = 0.3$ s, the Dick instability is at $1.5 \times 10^{-16}/\sqrt{\tau}$, for measurement time τ. This constitutes an order of magnitude improvement over our previous Dick limit, and enables a significant improvement of the clock stability. By using the narrow atomic transition as a frequency discriminator, we have made measurements consistent with a clock instability at the $5 \times 10^{-16}/\sqrt{\tau}$ fractional frequency level [6].

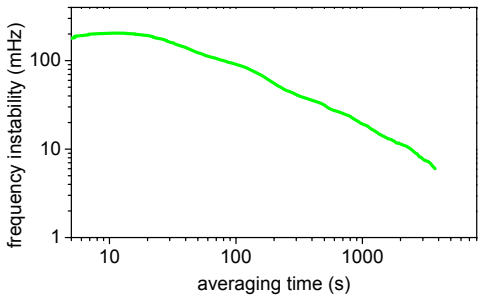


Fig. 2. Frequency instability from measurement of the cold collision shift, when the shift is cancelled to zero (see text).

III. Cold collision shift

While the presence of many atoms in the optical lattice is desirable for low clock instability due to quantum projection noise, another consequence of large atom number is high number density which can lead to significant interactions between the atoms. These interactions can shift the clock transition frequency and in so doing can potentially compromise the absolute uncertainty of the optical standard. Non-zero interactions in a ^{171}Yb optical lattice clock were first observed using one-pulse Rabi spectroscopy [7]. More recently, we have studied the resulting cold collision shift of the clock transition using two-pulse Ramsey spectroscopy, for atoms confined both in a one- and two-dimensional optical lattice. The choice of Ramsey spectroscopy was driven by simplification: the interactions yielding the collision shift occur primarily during the Ramsey dark time, when the atomic population is not simultaneously being driven by the probing laser field. By studying the collision shift as a function of excitation fraction, sample inhomogeneity, and Ramsey dark time, we have determined that the collisions responsible for the shift are dominated by p-wave interactions between a ground and excited state atom pair [8].

We have been able to identify regimes where the cold collision shift can be canceled, which is metrologically interesting for clock operation. One such case is given by controlling the excitation fraction just above 50%, where the shift on each clock state is the same. By operating at these conditions, the collision shift was measured to be consistent with zero at or below the 10^{-17} fractional frequency level. Figure 2 highlights such a measurement, showing the frequency instability of the measured collision shift. This data is taken by stabilizing the probe laser to interleaved samples of high and low atomic density, and measuring the frequency shift between the two cases. Due to the clock stability improvements described in section II, we are able to measure the shift at the 10^{-17} level with only several thousand seconds of measurement time. This measurement emphasizes the potential of the Yb lattice clock for very high measurement stability, as well as high accuracy through control of the cold collision shift.

ACKNOWLEDGMENT

The authors would like to thank T. Fortier, M. Kirchner, S. Diddams, and D. Hume.

REFERENCES

[1] G. J. Dick, "Local oscillator induced instabilities in trapped ion frequency standards," *Proc. Precise Time and Time Interval Meeting*, p. 133–147, 1987.

[2] P. G. Westergaard, J. Lodewyck, and P. Lemonde, "Minimizing the Dick effect in an optical lattice clock," *IEEE Trans. Ultra. Ferro. Freq. Cont.* **57** 623628, 2010.

[3] K. Numata, A. Kemery, and J. Camp, "Thermal-noise limit in the frequency stabilization of lasers with rigid cavities," *Phys. Rev. Lett.* **93** 250602, 2004.

[4] J. Millo *et al.*, "Ultrastable lasers based on vibration insensitive cavities," *Phys. Rev. A* **79** 053829, 2009.

[5] S. A. Webster, M. Oxborrow, and P. Gill, "Vibration insensitive optical cavity," *Phys. Rev. A* **75** 011801, 2007.

[6] Y. Y. Jiang *et al.*, "Making optical clocks more stable with 10^{-16}-level laser stabilization," *Nature Photon.* **5** 158-161, 2011.

[7] N. D. Lemke *et al.*, "Spin-1/2 optical lattice clock," *Phys. Rev. Lett.* **103** 063001, 2009.

[8] N. D. Lemke *et al.*, "p-Wave cold collisions in a Yb lattice clock," arXiv:1105.2014, 2011.

Hybrid MEMS Resonators and Oscillators

Sunil A. Bhave
OxideMEMS Laboratory, Cornell University
Ithaca, NY, USA
sunil@ece.cornell.edu

Abstract— **With quality factors (Q) often-exceeding 10,000, vibrating micromechanical resonators have emerged as leading candidates for on-chip versions of high-Q resonators used in wireless communications systems. However, as in the case for transistors, extending the frequency of MEMS resonators generally entails scaling of resonator dimensions. Unfortunately, smaller size often coincides with lower-power handling capability and increased motional impedance. In this paper we introduce novel transduction techniques which can improve the motional impedance of MEMS resonators by 1000× over traditional 'air-gap' transduced resonators, present latest results on narrow-bandwidth parametric filters for frequency-agile radio receivers, and discuss performance scaling of NEMS resonators to X-band frequencies.**

I. Introduction

Maturity in micro-electro-mechanical systems (MEMS) fabrication technology has opened the opportunity to design high quality factor (Q) mechanically-coupled micromechanical structures capable of implementing ultra-low-power frequency-domain signal processing functions. We endeavor to demonstrate new and high-impact micro-electro-mechanical circuits using such mechanically-linked networks. In the OxideMEMS Lab we have invented new mechanical designs and efficient transducers to manufacture MEMS resonators with high resonance frequency, low motional impedance, strong transducer coupling coefficient, low bias drift and large tuning range. We characterized and cataloged these resonators, developed and verified butterworth van-dyke (BVD) resonator models, investigated their coupling mechanisms to demonstrate channel-select filters, and explored inter-domain coupling to design merged resonant body transistors.

II. Dielectrically Transduced Resonators

Modern wireless transceivers are designed under a mandate to minimize the use of off-chip passive devices. However, the increasingly-crowded radio spectrum and the impending arrival of next generation 7-band cellular phone and the joint task force radio system (JTRS) has necessitated front-end filters capable of eliminating both out-of-band and out-of-channel interferers. Such channel-select filters will also relax the dynamic range requirements of the LNA and the phase noise requirements of the local oscillators, leading to a 10× reduction in power consumption of direct-conversion radio receivers.

Building of initial pioneering work at Berkeley and Delft, we demonstrated solid-dielectric transduced thickness shear mode MEMS resonators with 807 MHz center frequency, motional impedance $R_X < 100\Omega$ and Q's > 7,000 [1]. We have developed 'orthogonal frequency tuning' technique for independently changing the series and parallel resonant frequencies of dielectrically transduced thickness shear mode resonators. Using this technique our channel-select filter array is capable of dynamically tuning center frequency and filter bandwidth, thereby reducing capacitive loading at the filter input. This enables reception of multiple waveforms, and leads to substantial reduction in the number of filters in next-generation frequency-agile transceivers [2].

The thickness shear mode resonators have high Q and low R_X. However, their second-order dependence on CAD-defined dimensions severely limits the frequency design space. Contour-mode resonators can be designed across a much wider frequency range but have minimal tuning capability [3]. We have invented, designed and demonstrated a digital tuning technique for a filter consisting of mechanically coupled contour-mode resonators. By altering the polarization voltage applied to each resonator in the array we can digitally change the filter center frequency and alter the bandwidth from 1.4 MHz to 700 kHz (Figure 1) [4]. Unlike the analog voltage tunable shear-mode filters, the digitally-tunable filters are significantly less sensitive to variations and noise on the DC polarization voltage. We also proposed and demonstrated a 2D filter arraying technique to make the filters less susceptible to process variations [5]. We worked with Professor Tom Kenny's group at Stanford and developed an epi-silicon encapsulation process for vacuum packaging our resonators and filters in milli-Torr, clean micro-environment (Figure 2) [6].

Dielectric-transduction enables us to design high-Q, low impedance resonators and filters. However, dielectric transducers have very small bandwidth (< 0.2%), preventing them from being used for wide bandwidth IF filters. Aluminum Nitride (AlN) has been used to fabricate IF filters with 2% band-width, however these are fixed-frequency filters. We collaborated with Army Research Labs and co-developed voltage tunable ferroelectric contour-mode resonators and filters using lead zirconate titanate (PZT)-on-Silicon technology [7]. We have demonstrated resonators with 5% center frequency tuning and impedance < 100Ω. At MEMS 2009 we demonstrated the first monolithic, frequency

978-1-61284-110-6/11/$26.00 ©2011 IEEE

agile, channel-select filter bank integrating RF MEMS PZT switches with PZT-on-Silicon contour-mode filters (Fig 3) [8].

III. Aqueous Transduction

The frequencies of contour mode resonators are defined by CAD-defined lateral dimensions and can cover frequencies from 40 MHz – 2.4 GHz. But unlike shear mode resonators, the frequency expressions for contour modes and flexural vibration modes do not directly couple. It is therefore more difficult to perform orthogonal frequency tuning of contour mode resonators. BioMEMS groups have recently demonstrated that by using an LO signal that is faster than the response time of a polar fluid, it is possible to prevent electrode polarization and double-layer formation, enabling electrostatic transduction in liquid media. To determine whether the same approach enhances the performance of contour-mode RF MEMS resonators, we used an air-gap poly-SiGe disk resonator and submerged it under a water droplet. However, immersing the resonator in water caused excessive mass-loading and Q losses resulting from viscous drag. To eliminate mass-loading and viscosity effects of the water-droplet on the resonator, we coated the resonators with a hydrophobic, non-conformal self-assemble monolayer (SAM). We rolled a water droplet over the resonator such that it had enough time to wick the nano-gaps, but greatly reduced the chance of water seeping under the resonator. We were surprised to measure Q > 3,000 and motional impedance R_X < 4.2 kΩ, reduced from 510 kΩ with air-gap transduction. Replacing air in the sub-micron transducer gap with DI water enhanced the electrostatic spring by > 50× and extended the tuning range to 3%, the highest tuning-range reported for contour-mode resonators [9] (Figure 4).

IV. Internal Dielectric Transduction

Placing the dielectric or ferroelectric transducer on top of the resonator avoids the dielectric from hindering the contour-mode resonator vibration. However, such transducers suffer from two inefficiencies – the Poisson's effect and improper placement of the dielectric above regions of minimal strain. Both affect the effective coupling coefficient, and fractional bandwidth of the resonator. In the OxideMEMS lab we developed the concept of internal dielectric transduction, in which dielectric transducers are incorporated directly into the resonator body. With dielectric films positioned at points of maximum strain in the resonator, this transduction improves in efficiency with increasing frequency, enabling resonator scaling to previously unattainable frequencies. Using internal dielectric transduction in longitudinal-mode resonators, we have demonstrated a frequency-quality factor (f.Q) product of 5.1×10^{13} [10].

We measured these resonators by capacitively driving and sensing acoustic vibrations in the device. However, capacitive detection often requires 3-port scalar mixer measurement, complicating monolithic integration of the resonators with CMOS circuits. The internal dielectric bulk-mode resonators can be utilized in a 2-port configuration with capacitive drive and piezo-resistive detection, in which carrier mobility is dynamically modulated by elastic waves in the resonator. We demonstrated piezo-resistive sensing of silicon-based dielectrically transduced resonators with 1.6% frequency tuning and control of piezo-resistive trans-conductance gm by varying the current flowing through the device [11]. Furthermore at Transducers'09, we demonstrated selective excitation of targeted harmonics in poly-silicon dielectrically transduced bar resonators by varying dielectric position in the resonator body (Figure 5). We also showed frequency scaling behavior, with improved transduction efficiency at higher frequency. RF resonators up to 6.2 GHz were demonstrated, with f·Q products up to 3.1×10^{13} [12]. The high Qs measured in these resonators indicates the ability to scale resonators to high frequency without compromising f·Q product.

The initial calculations for optimizing internal dielectric transduction assumed the silicon and dielectric transducers had similar acoustic velocity and Young's modulus. Recently, we refined and published a complete theoretical treatment without such assumptions and provided correction factors for optimally designing the internal dielectric transduced resonators [13].

V. Resonant Body Transistor

The Resonant Body Transistor is a NEMS resonator integrating a sense transistor directly into the resonator body, geometrically similar to a suspended (HF undercut and released) split-gate FinFET, as shown in Figure 6. The region in light grey in Figure 6 represents the active region of the resonator, while the dark grey region is highly doped. The active region near the drive electrode is biased into accumulation (red), so that a large capacitive force acts across the thin dielectric film (yellow), driving longitudinal resonant motion in the body. A gate voltage is applied to the opposing electrode, generating an inversion channel (blue) which results in a DC drain current. At resonance, elastic waves formed in the resonator modulate the drain current both by physically changing the gate capacitance and by piezo-resistive modulation of carrier mobility. The internally amplified RBT has significantly lower output impedance than capacitive detection mechanisms, simplifying impedance matching with active circuits.

The RBT was designed and fabricated at CNF. Using internal dielectric transduction for driving the resonator into acoustic resonance, the RBT uses a field-effect transistor embedded into the resonator to detect resonance. We demonstrated an 11.7 GHz RBT with a Q > 1,800 in vacuum (Figure 6) [14]. This is by far the highest frequency ever measured in silicon. Moreover, due to its use of internal dielectric transduction, the RBT promises to scale favorably to 60GHz.

Acknowledgment

The author would like to thank NSF, DARPA, ARL, Rockwell Collins, Draper Labs, Analog Devices, Lockheed Martin and Intel for supporting various aspects of the research conducted in the OxideMEMS Lab.

REFERENCES

[1] Hengky Chandrahalim, Dana Weinstein, Lih Feng Cheow and Sunil A. Bhave, "Channel-select micromechanical filters using high-K dielectrically transduced MEMS resonators," IEEE MEMS 2006, Istanbul, Turkey, January 22-26, 2006, pp. 894-897.

[2] Lih Feng Cheow, Hengky Chandrahalim and Sunil A. Bhave, "MEMS filter with voltage-tunable center frequency and bandwidth," Hilton Head 2006 Workshop, Hilton Head Island, South Carolina, June 4-8, 2006, pp. 304-307.

[3] Dana Weinstein, Hengky Chandrahalim, Lih Feng Cheow and Sunil A. Bhave, "Dielectrically transduced single-ended to differential MEMS filter," ISSCC 2006, San Francisco, California, February 4-8, 2006, pp. 318-319.

[4] Hengky Chandrahalim and Sunil A. Bhave, "Digitally-tunable MEMS filter using mechanically-coupled resonator array," IEEE MEMS 2008, Tucson, Arizona, Jan 13-17, 2008, pp. 1020-1023.

[5] Dana Weinstein, Sunil A. Bhave, Masahiro Tada, Shun Mitarai, Shinya Morita and Koichi Ikeda, "Mechanical coupling of 2D resonator arrays for MEMS filter applications," IEEE Frequency Control Symposium (FCS 2007), Geneva, Switzerland, May 29 - June 1, 2007, pp. 1362-1365.

[6] Kuan-Lin Chen, Hengky Chandrahalim, Andrew Graham, Sunil A. Bhave, Roger T. Howe and Thomas W. Kenny, "Epitaxial silicon microshell vacuum-encapsulated CMOS-compatible 200 MHz bulk-mode resonator," IEEE MEMS 2009, Sorrento, Italy, Jan 25-29, 2009, pp. 23-26.

[7] Hengky Chandrahalim, Sunil A. Bhave, Ronald Polcawich, Jeff Pulskamp, Daniel Judy, Roger Kaul and Madan Dubey, "Performance comparison of Pb(Zr0.52Ti0.48)O3-only and Pb(Zr0.52Ti0.48)O3-on-silicon resonators," Applied Physics Letters 93 233504 (2008).

[8] Jeff Pulskamp, Daniel Judy, Roger Kaul, Ronald Polcawich, Hengky Chandrahalim and Sunil A. Bhave, "Monolithically integrated piezo-MEMS SP2T switch and contour-mode filters," IEEE MEMS 2009, Sorrento, Italy, Jan 25-29, 2009, pp. 900-903.

[9] Hengky Chandrahalim, Sunil A. Bhave, Emmanuel Quévy and Roger T. Howe, "Aqueous transduction of poly-SiGe disk resonators," Transducers'07, Lyon, France, June 10-14, 2007, pp. 313-316.

[10] Dana Weinstein and Sunil A. Bhave, "Internal dielectric transduction of a 4.5 GHz silicon bar resonator," IEDM 2007, Washington DC, Dec. 10-12, 2007, pp. 415-418.

[11] Dana Weinstein and Sunil A. Bhave, "Piezoresistive sensing of a dielectrically actuated silicon bar resonator," Hilton Head 2008 Workshop, Hilton Head Island, South Carolina, June 1-5, 2008, pp. 368-371.

[12] Dana Weinstein, Sunil A. Bhave, Shun Mitarai, Shinya Morita and Koichi Ikeda, "Frequency scaling and transducer efficiency in internal dielectrically transduced silicon bar resonators," Transducers'09, Denver, Colorado, June 21-25, 2009, pp. 708-711.

[13] Eugene Hwang and Sunil A. Bhave, "Acoustic mismatch and its effects on internally transduced micromechanical resonators," IEEE Frequency Control Symposium (FCS 2009), Besançon, France, April 20-24, 2009, pp. 460-465.

[14] Dana Weinstein and Sunil A. Bhave, "The resonant body transistor," Nano Letters (2010).

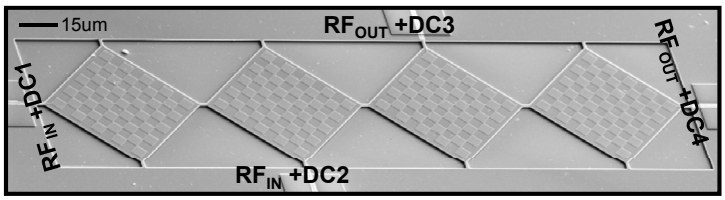


Figure 1.a: *Four series mechanically-coupled checkered-electrode resonators. Each resonator has individually addressable RF+DC lines. The silicon device layer is an RF ground to prevent substrate losses. All RF paths are metalized, routed on an elevated oxide bridge to reduce capacitance and de-embedded to the edge of DRIE trench.*

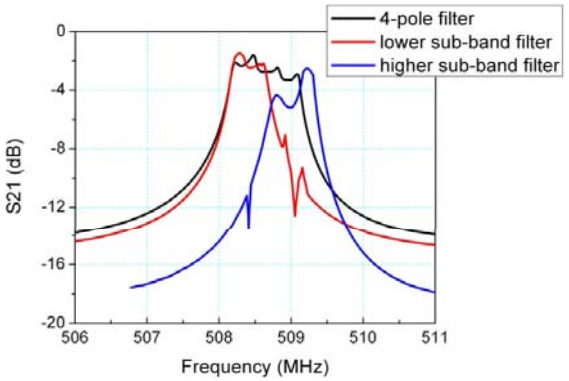


Figure 1.b: *Transmission response of three filters after termination using the PNA's pole-Z conversion function.*

	4-pole filter	Lower sub-band filter	Higher sub-band filter
IL	-2.4 dB	-2 dB	-3.1 dB
3dB BW	1.4 MHz	720 kHz	660 kHz
f_{CENTER}	508.7 MHz	508.3 MHz.	509.1 MHz
ripple	< 2 dB	<1 dB	< 3 dB
stop band rejection	-14 dB	- 14 dB	-20 dB
DC Bias	50V on first and last resonators	40V and -40V on alternate resonators.	40V on all resonators
$R_{TERMINATION}$	4.1 kΩ	2.8 kΩ	2.9kΩ

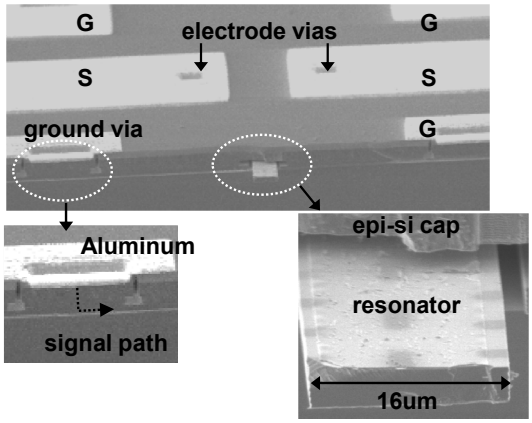


Fig. 2.a: *SEM of the fabricated resonator. The zoom-in picture of middle part shows the thin poly electrode on top of the resonator bar. The zoom-in picture of left part shows electrical contact between epi-silicon and device layer.*

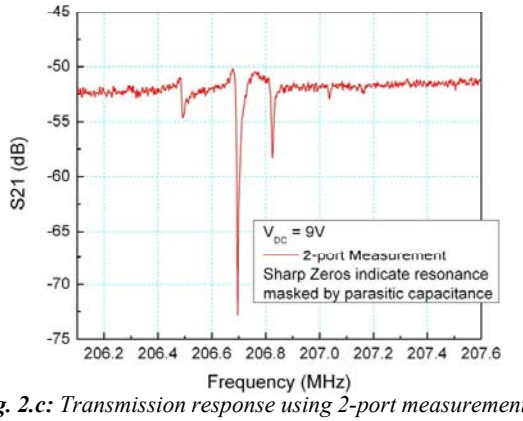


Fig. 2.c: *Transmission response using 2-port measurement. Capacitive feed-through masks the electromechanical resonance.*

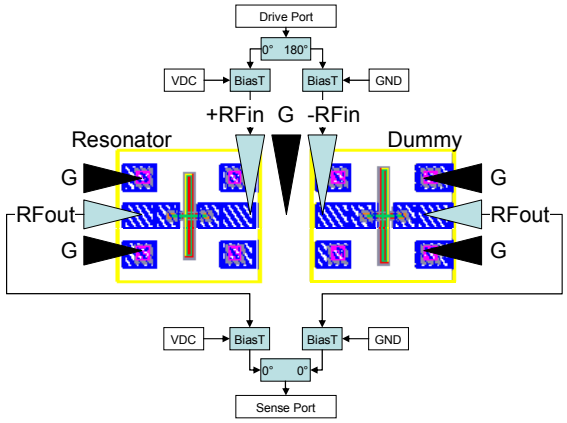


Fig. 2.b: *Schematic of the pseudo-differential measurement.*

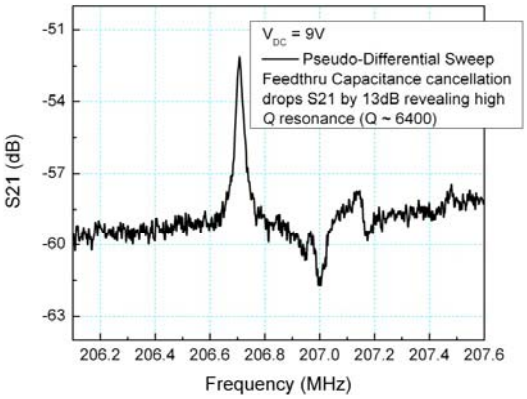


Fig. 2.d: *Transmission response using pseudo-differential measurement technique to cancel capacitive feed-through.*

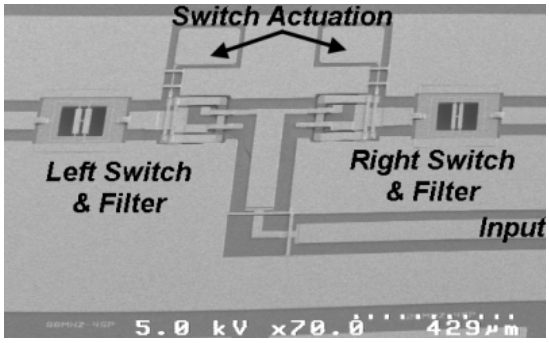


Fig. 3.a: *SEM of a monolithically integrated single pole dual throw PZT MEMS switch and two PZT contour mode filters.*

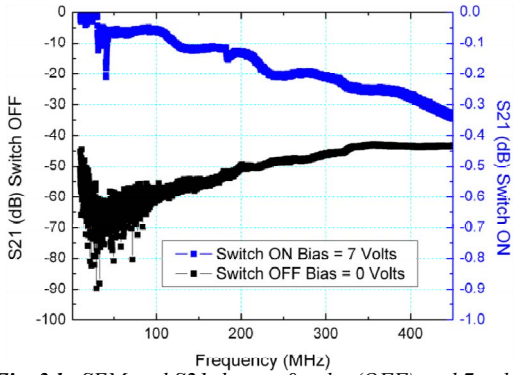


Fig. 3.b: *SEM and S21 data at 0 volts (OFF) and 7 volts (ON) for the PZT RF MEMS switch.*

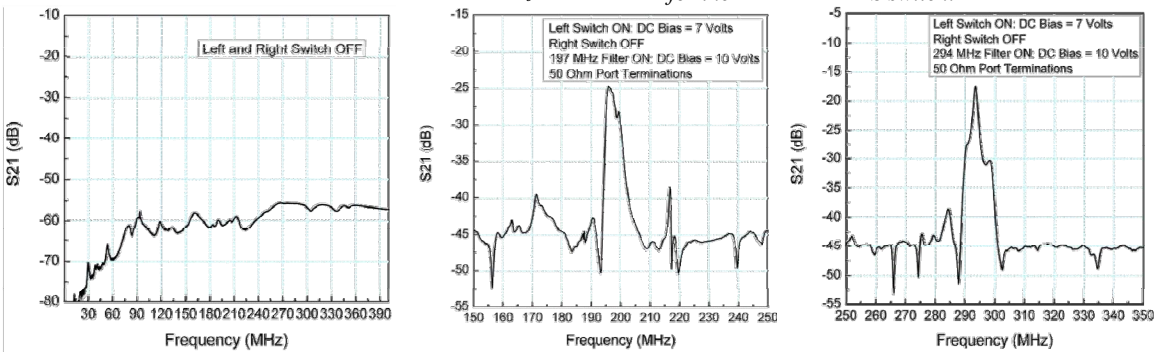


Fig. 3.c: *(a) S21 response for an integrated SP2T switch and filter with the switch in the off-state (0 V), (b) S21 response for the left switch and filter with the switch on (7 V) and with 10 VDC applied to the filter, and (c) S21 response for the right switch and filter with the switch on (7 V) and with 10 VDC applied to the filter.*

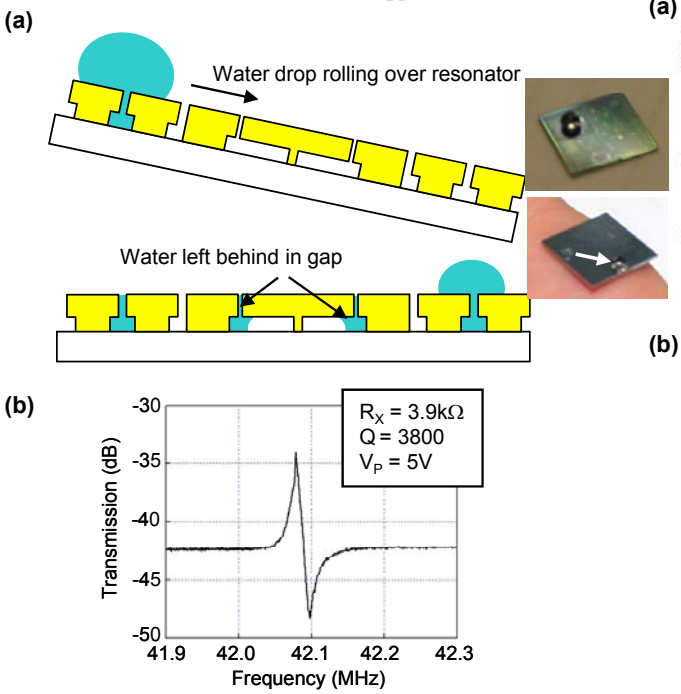


Figure 4.a. *(a) Schematic of water droplet rolled across the resonator, wicking the transducer gap, but minimizing the water left underneath the resonator (b) transmission response: Q of 3,800 at 42 MHz with R_x = 3.9 kΩ*

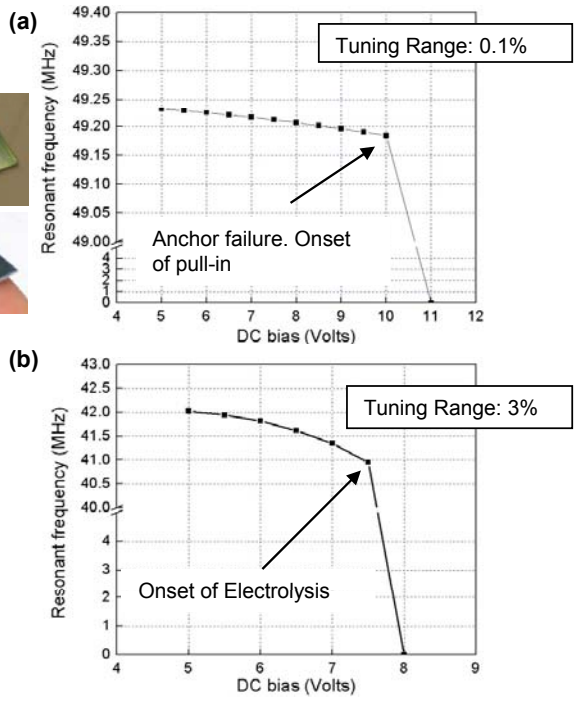


Figure 4.b. *(a) 0.1% resonant frequency tuning of the air-gap transduced disk resonator before anchor failure, (b) In DI water, the resonator is tuned up to 3%, before the onset of electrolysis.*

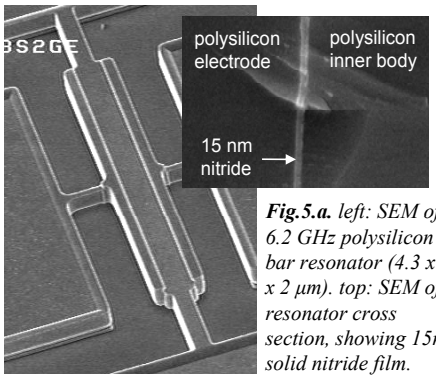


Fig.5.a. *left: SEM of 6.2 GHz polysilicon bar resonator (4.3 x 40 x 2 μm). top: SEM of resonator cross section, showing 15nm solid nitride film.*

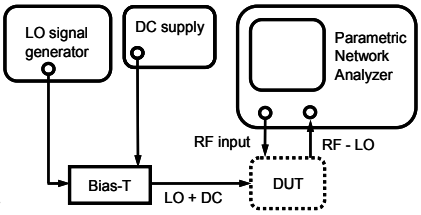


Fig.5.b. *Scalar mixer measurement of the 3-port MEMS resonator. The resonator (DUT) acts as a mixer for the input RF and LO signals. The resonance is detected at RF – LO.*

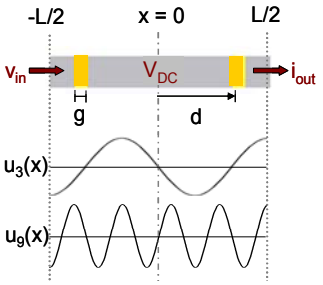


Fig.5.c. *Cross section of resonator showing dielectric (yellow), with normalized mode shapes of 3rd and 9th harmonics of longitudinal resonance (1D in-plane*

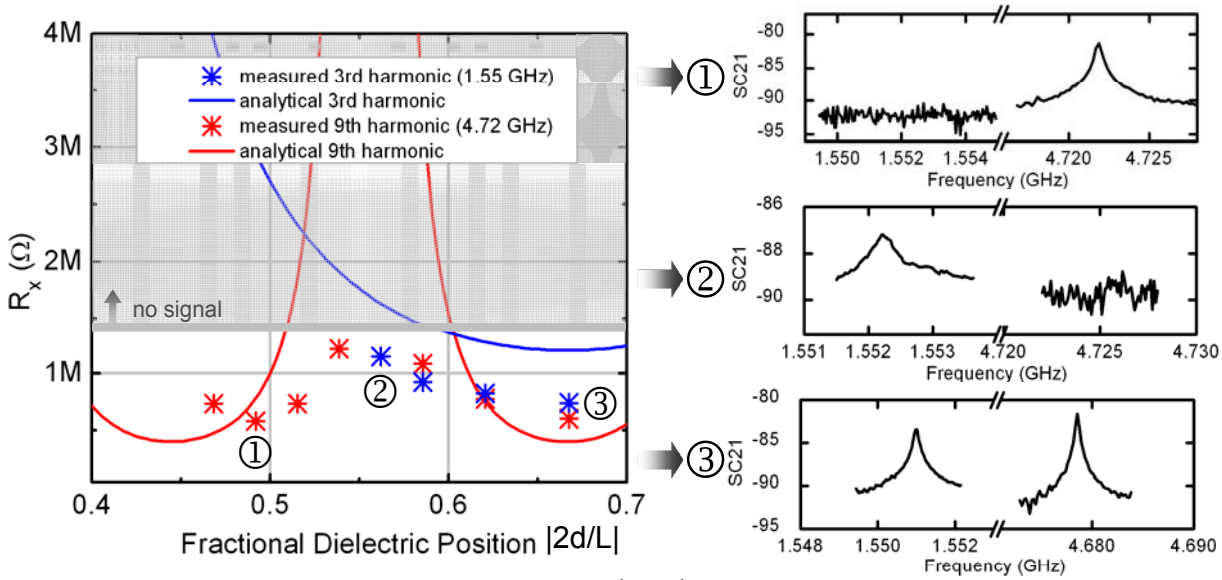
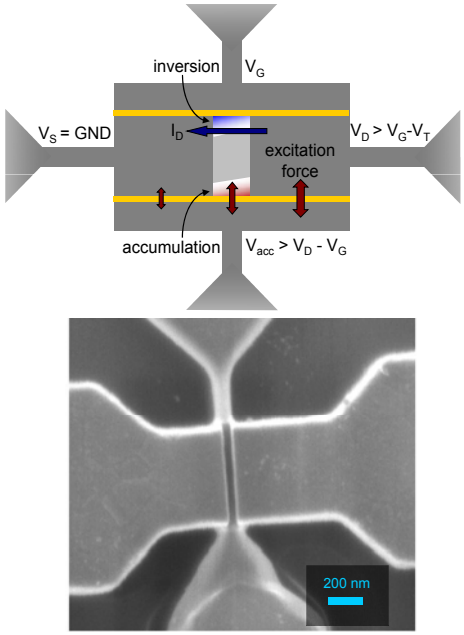


Fig.5.d. *Motional impedance vs. position of dielectric transducer in 3rd and 9th harmonics of 8 different resonators. Selective positioning of dielectric can excite ① only 9th harmonic at 4.72 GHz, ② only 3rd harmonic at 1.55 GHz, and ③ both 3rd and 9th harmonics.*



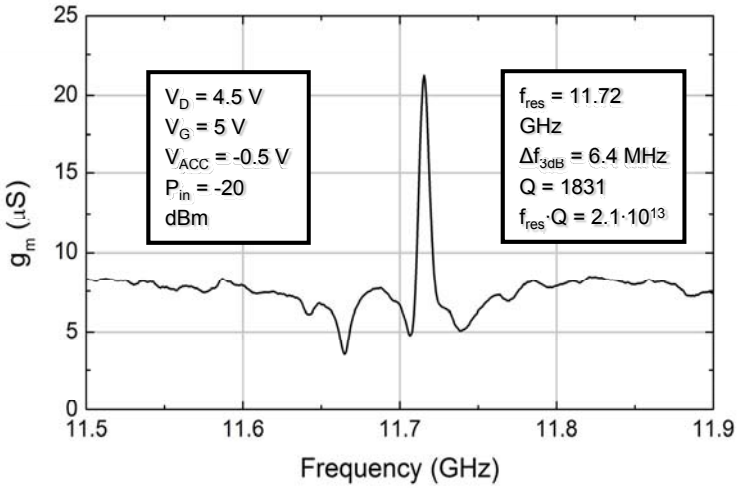


Fig.6. *Schematic, SEM and electrical transfer characteristics of an 11.7GHz Resonant Body Transistor.*

A CMOS-compatible 24MHz poly-SiGe MEMS oscillator with low-power heating for frequency stabilization over temperature

R. Jansen, M. Libois, X. Rottenberg, M. Lofrano, J. De Coster, R. Van Hoof, S. Severi, G. Van der Plas, W. de Raedt, H.A.C. Tilmans, S. Donnay and J. Borremans
imec
Leuven, Belgium
Jonathan.Borremans@imec.be

***Abstract*—MEMS based timing devices have been proposed as an alternative to Quartz systems for certain applications. Through using an oven-controlled system it is possible to stabilize the frequency response of such a MEMS system over a large ambient temperature range. This work presents a BAW MEMS resonator in poly-SiGe which achieves significantly lower power consumption for frequency stabilization over temperature through Joule heating than for a similar system in SOI, while showing promising phase noise performance in an oscillator setup. Since the poly-SiGe resonator can be processed on top of standard CMOS, this enables the possibility of full integration of an Oven-Controlled MEMS Oscillator.**

I. INTRODUCTION

Silicon-based resonators provide an attractive alternative to replace conventional frequency control and timing devices that are based on high quality factor (*Q*) resonators such as quartz crystal, piezoelectric ceramic and surface acoustic wave (SAW) [1,2]. The bar-type bulk-acoustic (BAW) poly-SiGe resonators exemplify this by offering the possibility to integrate MEMS and CMOS in a "MEMS-last" approach [3,4]. Optimizing the quality factor Q is crucial, as it is a critical figure of merit for resonators and has a large impact both on the motional resistance and on the phase noise for circuits employing these resonators.

A further important consideration for timing devices is the frequency stability over temperature (typically ±10ppm down to ±0.1ppm over temperature range -40°C to +80°C). Due to the relatively large frequency drift of MEMS devices over temperature (in the order of tens of ppm/°C) in comparison to quartz, a compensation scheme is required to stabilize frequency over temperature. Several frequency-correcting mechanisms can be employed, typically steered by an ambient temperature sensor. A well-known solution is to embed the MEMS oscillator in a fractional PLL, whose fractional frequency control word is steered by a temperature sensor [5]. In such a system, the PLL provides a clock output, stable over temperature. However the output is polluted with fractional-PLL noise. Moreover, the feed-forward temperature compensation requires multipoint calibration for ppm-level stability over a wide temperature range, and requires a high-resolution, sufficiently linear temperature sensor.

A second possibility involves keeping the resonator at a fixed temperature by placing the device in a micro-oven [6]. This approach is followed in this work, creating an Oven-Controlled MEMS Oscillator (OCMO), which potentially yields good clock stability over ambient temperature. The target temperature for the resonator should be set above the specified ambient temperature range, for example at 95°C. Hence, a power-efficient micro-oven is needed that can keep the resonator at 95°C. A bar-type BAW resonator system is presented that proposes to heat the resonator, keeping it at a constant temperature above the maximum operating range and therefore minimizing the effect of ambient temperature changes on the resonance frequency. This design is made possible through the use of T-type supports which allow for Joule heating through the supports, while simultaneously optimizing the design in terms of anchor losses to the substrate, therefore maximizing the Q [7]. The T-type design is rigid in the direction of actuation, allowing the relatively high bias voltages needed to lower the motional resistance and ease the design of an oscillator system.

In this paper such an oscillator system is demonstrated using commercially available off-the-shelf components (COTS), showing sufficient phase noise performance for many applications, with the promise of full integration of the poly-SiGE resonator with CMOS to produce an Oven-Controlled MEMS Oscillator (OCMO). Furthermore we show that the lower thermal conductivity of poly-SiGe compared to SOI provides a significant advantage in the amount of power needed to heat the resonator for similar devices manufactured in both technologies.

This work is organized as follows: Section II introduces the poly-SiGe technology, Section III describes the resonator design using T-shaped supports, while section IV discusses the implementation and of an oscillator based on the SiGe resonator using COTS. We demonstrate the proposed poly-SiGe OCMO is thus adequate for analog performance and for temperature stabilization through Joule heating.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

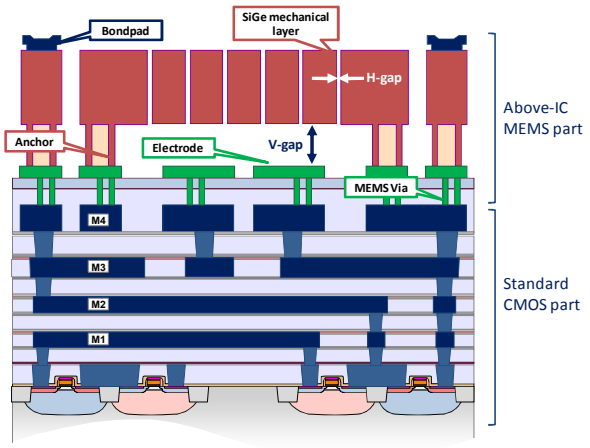


Figure 1, Cross-section layer build-up of the SiGe MEMS above-IC

II. POLY-SIGE MEMS TECHNOLOGY

Poly-SiGe IMEC MEMS technology allows for the fabrication of MEMS devices directly above a planarized CMOS wafer, on top of the metal interconnection lines, as described in [3,8]. This integration approach enables cost and area reduction as several CMOS functionalities, such as sustaining amplifier and temperature-compensation, can be placed just below the resonator. Ultimately, other functionalities such as a real time clock, gyros, accelerometers or pressure sensors can be integrated on the same die. For explanatory purposes, Figure 1 illustrates the cross-section of a resonator above the complete CMOS process. In this work, however, the MEMS device is defined on top of a simple AlCu interconnection that mimics the top metal of a CMOS backend stack. A SiC layer is deposited to protect the CMOS metallization during the vHF MEMS release process [3]. Vias are opened into this layer to realize the connection with the IC. A chemical vapor deposited (CVD) SiGe layer fills in the vias and form simultaneously the electrode. After electrode patterning, the wafer is planarized to ensure optimal definition of the 3µm high density plasma (HDP) oxide sacrificial gap. Anchors are defined in the oxide layer and then filled during the deposition of the SiGe MEMS structural layer by a combination of CVD and plasma-enhanced CVD (PECVD) processes [9].

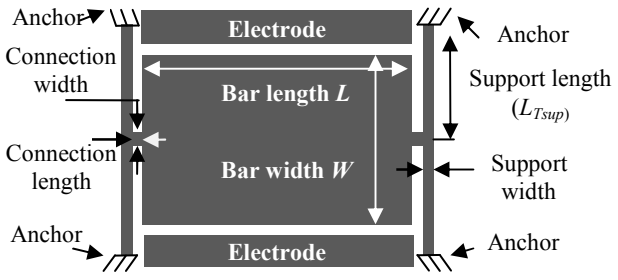


Figure 2, Schematic top view of a bar-type BAW resonator with "acoustically long" T supports. [7]

III. BAR RESONATOR DESIGN

The bar-type BAW resonators studied in this paper consist of right parallelepipeds suspended in free space by means of T-supports, *e.g.*, as in Figure 2. Figure 3 shows a SEM image of such a bar-type resonator as realized in imec's poly-SiGe MEMS platform [3,8]. The bar vibrates along its width in an extensional mode. A bar targeting a resonance frequency f_{res} is designed with a bar width W close to $\lambda_{long}/2$, where λ_{long} is the acoustic longitudinal wavelength at f_{res}. The resonance frequency, for the first width extensional mode, is approximately given by:

$$f_{res} \approx \frac{1}{2W}\sqrt{\frac{E}{\rho}} \qquad (1)$$

where E and ρ denote the Young's modulus and specific mass of the resonator material, respectively. In this work an E of 120 GPa, ρ of 4575 kg/m^3 and W of 110 µm result in a resonance frequency of 24MHz. The other design parameters are presented in Table 1. The bar resonator is electro-statically actuated via two electrodes separated from the resonator by transduction gaps (250 nm in this work), e.g., as shown in Fig. 2.

A. T-type support

The T-type support comprises of a support placed parallel to the side of the resonator connecting to the resonator at the symmetry line with a short, narrow connection. Figure 4 contains SEM images of a manufactured resonator with a

Table: Design parameters considered in this work.

Variable	Symbol	Value
Design Frequency (SiGe)	f_{res}	24 MHz /48 MHz
Bar Width	W	110µm /55µm
Bar Length	L	220µm/110µm
Support Width	W_{sup}	1 µm
Support Length	L_{sup}	Design dependent
Transduction Gap	d_{gap}	250 nm
Connection Width	$W_{connect}$	2 µm
Connection Length	$L_{connect}$	2 µm

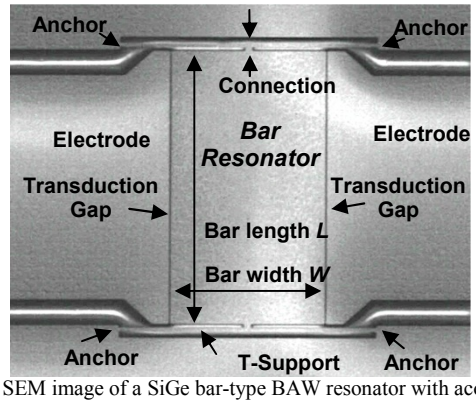


Figure 3, SEM image of a SiGe bar-type BAW resonator with acoustically long T-type supports. Design parameters are as in Table 1 for a 48MHz device, with a support length of 38µm.

T-type support. A λ/4 transformer technique can be applied to obtain an optimum support length $L_{Tsup,opt}$ to minimize energy losses to the anchors, with $L_{Tsup,opt}$ given by:

$$L_{Tsup,opt} \approx \frac{2i\lambda_{flex}}{4} + \frac{L_{cl-cl,1}}{2}, \qquad i = 0,1,2,3... \tag{4}$$

where for a given clamped-clamped beam, the flexural wavelength λ_{flex} is given by:

$$\lambda_{flex} = \lim_{n\to\infty} \left(L_{(n+2)} - L_n \right) \tag{6}$$

where $L_{(n+2)}$ and L_n are the lengths of beams with respectively $(n+2)^{th}$ and n^{th} flexural harmonics occurring at the frequency f_{res} [7]. We can compute λ_{flex} with eq. (6), after which $L_{Tsup,opt}$ is obtained from eq. (4). $L_{cl-cl,1}$ is the length of a clamped-clamped beam with first flexural resonance frequency equal to f_{res}, approximated by [7]:

$$L_{cl-cl,1} \approx \sqrt{\frac{1.028 W_{sup} \sqrt{E/\rho}}{f_{res}}} \tag{5}$$

Figure 5 demonstrates the benefits of using this technique. It shows the measured Q values against support length for a bar type resonator with design frequency at 48 MHz, along with simulations of the Q factor due to anchor losses. It can be seen that maxima for the Q values do indeed exist at multiples of $\lambda_{flex}/2$ given a certain offset as described by eq. (4). Therefore the Q factor can be optimized for longer support lengths, allowing us to design supports with higher thermal isolation between the resonator and the anchor.

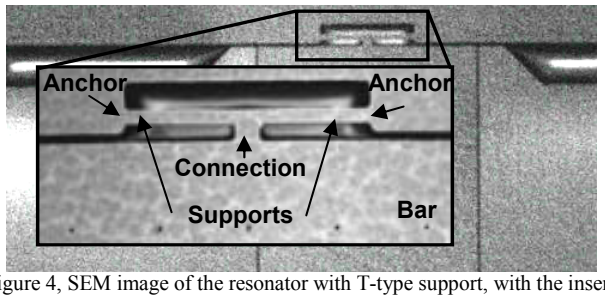


Figure 4, SEM image of the resonator with T-type support, with the insert showing a closer view of the support itself.

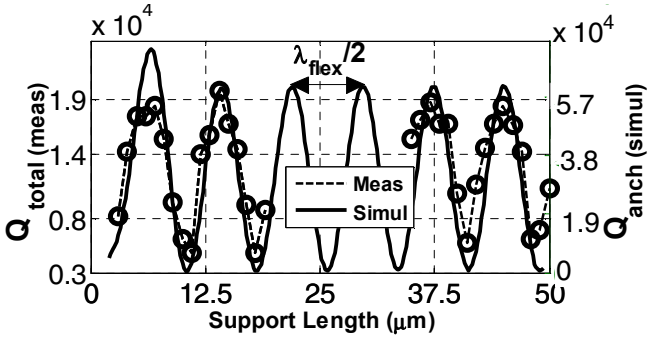


Figure 5, Measurement of Q and COMSOL simulation of the anchor loss Q_{anch} versus support length for a T-support, with $\lambda_{flex}/2$ shown to be 7.5μm. Devices were designed to operate at 48 MHz.

IV. Oven-Controlled MEMS Oscillator

An Oven-Controlled MEMS Oscillator (OCMO) potentially yields good clock stability over ambient temperature. The target temperature for the resonator should be set above the specified ambient temperature range, for example at 95˚C. Hence, a power-efficient micro-oven is needed that can keep the resonator at 95˚C.

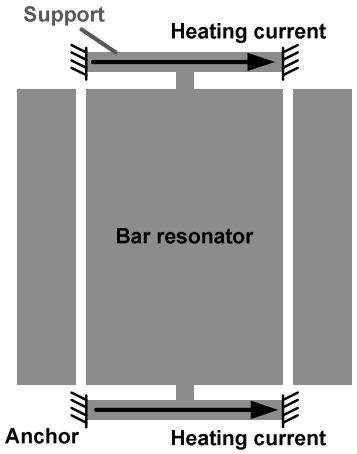


Figure 6, Schematic layout of the resonator, top view, showing the current used for Joule heating to keep the resonator at a fixed temperature.

A. Resonator design for use in OCMO

We use the T-type support described in the previous section for the BAW type bar resonator targeted at 24 MHz. This T-type support design provides several advantages for an OCMO design. Firstly the design is very stable in the direction of actuation, which allows the application of large bias voltages [7]. The supports also provide both a thermal resistance for Joule heating, as well as thermal isolation between the bar and the substrate. Figure 6 demonstrates how the current for Joule heating flows through the supports. Since the resonator is intended to operate in a vacuum packaged environment, the predominant thermal losses occur through the anchors which are connected to the substrate. The thermal isolation provided by the support length is therefore

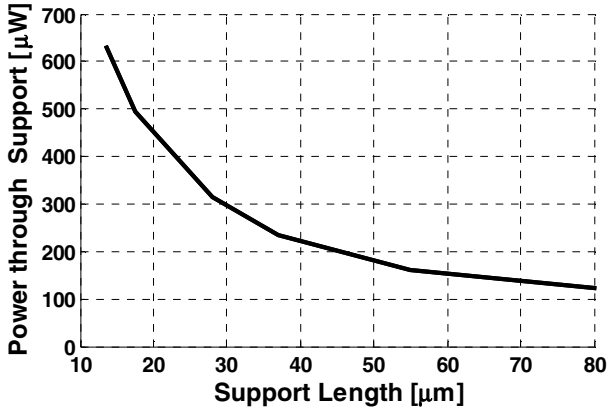


Figure 7, Simulated power required to heat SiGe resonator (with design parameters as in Table 1 for 24MHz and support length 38μm) to a temperature of 95˚C degrees with ambient at -40˚C.

crucial in determining the amount of power needed to keep the resonator at a fixed temperature. In Fig. 7 FEM simulation results can be seen showing that a minimum support length of 35µm is required (with support width at 1µm) to keep the heating power under 250µW. Using eq. (4) the support was designed to have a width and length of 1µm and 38µm respectively, maximizing the quality factor while meeting the required length specification provided by the thermal limitations.

B. Oscilator using COTS components

To demonstrate the feasibility of the system, a 24MHz oscillator has been built around the SiGe resonator using COTS components. Figure 8 shows a diagram depicting the system and contains an amplifier, a tunable phase shifter, a band-pass filter, a 20 dB coupler, a programmable attenuator and a limiter. The resonator is probed on wafer inside a vacuum chamber. Figure 9 depicts the measured phase noise of the 24MHz oscillator. For reference, the phase noise of the same setup using a commercial 26MHz quartz crystal is also shown. Both achieve the GSM specifications (referred to 24MHz and also indicated on Figure 9) at lower offset frequencies. At higher offset frequencies, the SiGe oscillator exhibits a higher noise floor. This is partly due the larger amount of gain required due to the relatively larger motional resistance of SiGe resonator as opposed to the crystal. However the SiGe performance is also limited by the 50Ω external loop components and is expected to improve through full CMOS co-integration of the sustaining amplifier. Indeed, a trans-impedance amplifier can be designed with far lower input-referred noise than an off-the-shelf 50Ω amplifier. Also, the tunable attenuator used to control the amount of gain in the system is one source of noise that can be removed when using a custom designed amplifier with the correct gain specifications. When considering the performance achieved despite the limitations of the probing setup and the COTS components used in this design, it can be seen that the technology shows potential to achieve GSM phase noise performance.

C. Power consumption for heating in SiGe and SOI

The OCMO micro-oven functionality is implemented by heating the resonator through use of Joule heating. Currents are applied through the T-type supports as demonstrated in Figure 6. The amount of power required to reach 95˚C is of crucial importance to the feasibility of the system. Poly-SiGe has a thermal conductivity that is more than 40 times lower than silicon (2.8 W/mK and 120 W/mK for SiGe and Silicon respectively), providing a distinct advantage in this implementation. Since the device is intended for packaged vacuum operation the majority of the thermal losses occur through the supports to the substrate. For the SiGe resonator, the thermal resistance of the T-type support is much larger than for the SOI support, which results in large power savings. The measured power required to heat the resonator to the targeted 95˚C is shown in Figure 7 for SiGe and SOI resonators with the same physical dimensions. The SiGe system needs far 40 times less power (<250µW at -40°C ambient) than an SOI implementation would need (> 10mW at -40°C ambient). In reality the generation of heater currents requires more power than that dissipated in the supports, resulting in a larger effective heater power. These SOI numbers compare to the work in [6]. A drawback for the implementation in SiGe is the lower quality factor of the poly-crystaline structures compared to the crystalline SOI.

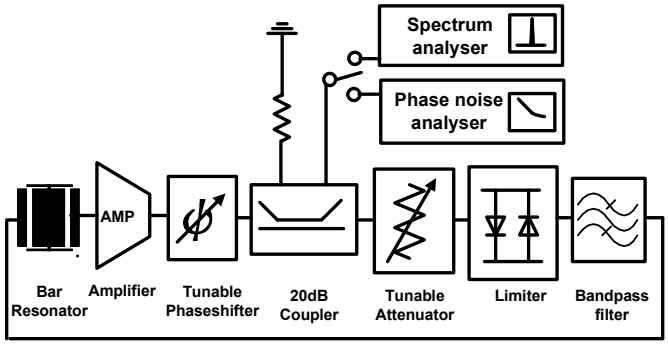


Figure 8, Oscillator setup using COTS sustaining amplifier

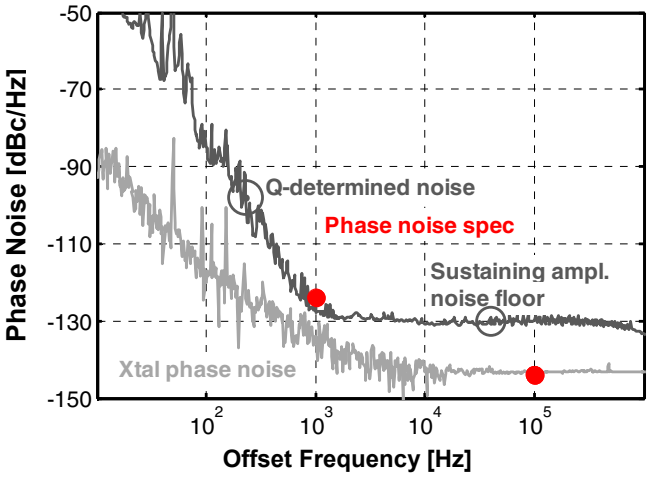


Figure 9, Measured phase noise of the SiGe MEMS oscillator using a COTS sustaining amplifier loop. For reference, the measured phase noise of a 26MHz XTAL oscillator using the same COTS setup. Also depicted is the phase noise specification for GSM.

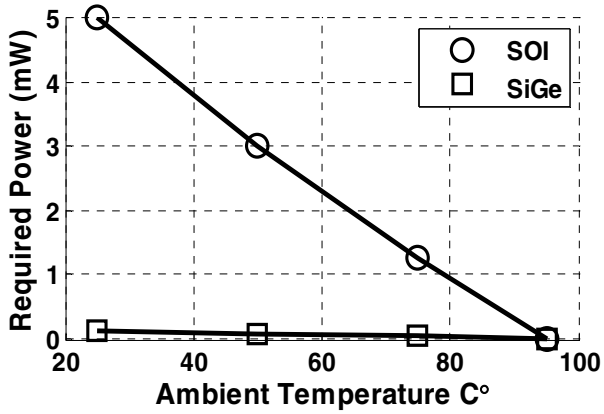


Figure 10, Measurements showing the required heater power applied through support legs of 38µm to maintain resonator temperature at 95˚C with varying ambient temperature for SOI and SiGe resonators with identical configurations.

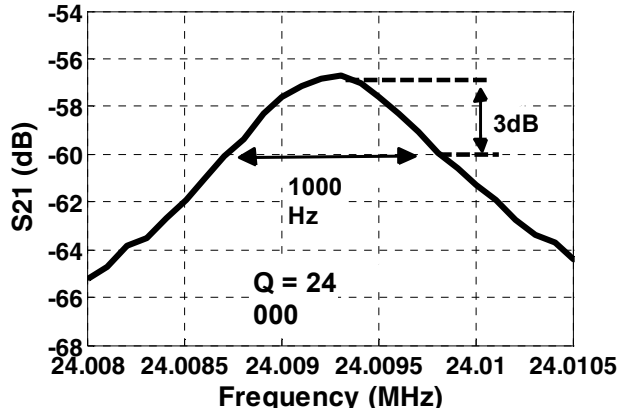


Figure 11, Measured S_{21} of a poly-SiGe resonator with length and width 220 and 110μm respectively. The quality factor is shown to be 24000.

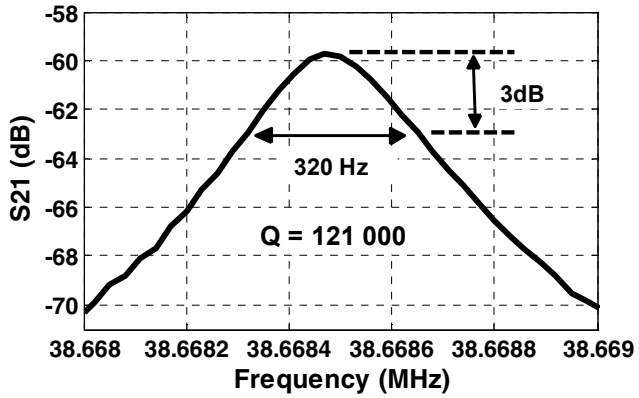


Figure 12, Measured S_{21} of a SOI resonator with length and width 220 and 110μm respectively. The quality factor is shown to be 121000.

Figure 11 and Figure 12 display the measured S_{21} of a SiGe and SOI resonator respectively, showing a quality factor of 24,000 for SiGe (at 24MHz) and 121,000 for SOI device (at 38.7MHz). The higher Q for SOI indicates that SOI resonators have an advantage with respect to SOI for achieving low phase noise in a timing device. However, as we have shown, the system with the SiGe does show potential to meet the GSM specification, indicating that SiGe is *good enough*. On the other hand, the 40 times higher power consumption of a micro-oven in SOI (in excess of 10mW) is prohibitive for a TCXO replacement, making SiGe a more suitable technology in which to implement an OCMO.

V. CONCLUSIONS

We demonstrate an oscillator using a bar-type BAW MEMS resonator that shows promising phase noise performance. The resonator is designed using T-shaped supports, optimizing the Q while providing sufficient thermal isolation to the substrate. The SiGe resonator requires significantly less power to heat than a similar device in SOI. Furthermore, the SiGe platform allows the resonator to be processed on top of CMOS, enabling a fully integrated OCMO system.

REFERENCES

[1] S. Pourkamali, G.K. Ho, F. Ayazi, Vertical Capacitive SiBARs", *MEMS'05 Conf.* 2005, *Miami*, pp. 211-214.

[2] V. Kaajakari, T. Mattila, A. Oja, J. Kiihamäki and H. Seppä, "Square-Extensional Mode Single-Crystal Silicon Micromechanical Resonator or Low-Phase-Noise Oscillator Applications", *IEEE El. Dev. Let.*, vol. 25(4), April 2004, pp. 173-175.

[3] S. Stoffels, "Extensional mode SiGe MEM-Resonators: Design, modeling and fabrication", Ph.D. dissertation, K.U. Leuven, December 2010.

[4] L. Haspeslagha, J. De Coster, O. Varela Pedreira, I. De Wolf, B. Du Bois, A. Verbist, R. Van Hoof, M. Willegems, S. Locorotondo, G. Brycea, J. Vaesa, B. van Drieenhuizen and A. Witvrouw, "A. Highly reliable CMOSintegrated 11MPixel SiGe-based micro-mirror arrays for high-end industrial applications", IEDM, San Francisco, CA, 2008, pp. 655.

[5] M. H. Perrott, S. Pamarti, E.G. Hoffman, F.S. Lee, S.Mukherjee, C. Lee, V. Tsinker, S. Perumal, B. T. Soto, N. Arumugam, and B. W. Garlepp, "A Low Area, Switched-Resistor Based Fractional-N Synthesizer Applied to a MEMS-Based Programmable Oscillator," IEEE J. Solid-State Circuits, vol. 45, Dec 2010, pp. 2566-2581.

[6] J. Salvia, R. Melamud, S. Chandorkar, S. F. Lord, and T. W. Kenny, "Real-Time Temperature Compensation of MEMS Oscillators Using an Integrated Micro-Oven and a Phase Lock Loop," Journal of Microelectromechanical Systems, Vol. 19, No. 1, pp. 192-201, 2010.

[7] R. Jansen, S. Stoffels, X. Rottenberg, Y. Zhang, J. De Coster, S. Donnay, S. Severi, J. Borremans, M. Lofrano, G. van der Plas, P. Verheyen, W. de Raedt and H.A.C. Tilmans, "Optimal T-support anchoring for bar-type BAW resonators", MEMS 2011.

[8] S. Severi, J. Heck, T.-K.A. Chou, N, Belov, J.-S. Park, D, Harrar, A. Jain, R. Van Hoof, B. Du Bois, J. De Coster, O.V. Pedreira, M. Willegems, J. Vaes, G. Jamieson, L. Haspeslagh, D. Adams, V. Rao, S. Decoutere, A. Witvrouw, "CMOS-integrated poly-SiGe cantilevers with read/write system for probe storage device", *TRANSDUCERS*, Denver, CO, 2009, pp. 2409.

[9] B. Guo, S. Severi, G. Bryce, G. Claes, R. Van Hoof, B. Du Bois, L. Haspeslagh, A. Witvrouw, S. Decoutere, "Improvement of PECVD Silicon- Germanium Crystallization for CMOS Compatible MEMS Applications", *Journal of The Electrochemical Society*, 157(2), pp. D103 – D110, 2010.

[10] J. De Coster, M. Lofrano, R. Jansen, X. Rottenberg, S. Severi, J. Borremans, G. Van Der Plas, S. Donnay, I. De Wolf, H.A.C. Tilmans, "Measurement of thermal conductivity, CTE and residual stress of SiGe using a laser doppler vibrometer", submitted to Transducers 2011.

Mechanically-Coupled CMOS-MEMS Free-Free Beam Resonator Arrays with Two-Port Configuration

Ming-Huang Li[1], Wen-Chien Chen[2], and Sheng-Shian Li[1,2]
[1]Institute of NanoEngineering and MicroSystems, [2]Department of Power Mechanical Engineering
National Tsing Hua University
Hsinchu, Taiwan
Email: ssli@mx.nthu.edu.tw

***Abstract*—Integrated CMOS-MEMS free-free beam resonator arrays operated in a standard two-port electrical configuration simultaneously with low motional impedance and high power handling capability centered at 10.5 MHz have been demonstrated by the combination of pull-in gap reduction mechanism and mechanically-coupled array design. The mechanical links (i.e., coupling elements) using short stubs connect each constituent resonator of an array to its adjacent ones at the *high-velocity* vibrating locations to accentuate the desired mode and reject all other spurious modes. A single second-mode free-free beam resonator with quality factor Q > 2,200 and motional impedance R_m < 150 kΩ has been utilized to achieve mechanically-coupled resonator arrays in this work. In array design, a 9-resonator array has been experimentally characterized to have around 10X performance improvement on motional impedance and power handling as compared with that of a single resonator. In addition, two-port electrical configuration is much preferred due to its low feedthrough nature and high design flexibilities for future oscillator and filter implementation rather than its one-port counterpart.**

I. Introduction

Vibrating off-chip mechanical tank components such as quartz crystals and SAW resonators are widely used in wireless transceivers and electronic systems for precision timing and filtering applications. In particular, quartz crystal oscillator, a classical timing reference device, has been used in consumer electronics for more than fifty years because of its exceptional Q and excellent thermal stability. In recent years, electrostatic silicon-based resonators are very promising in terms of their low motional impedance, high Q, small volume, satisfactory power handling capability, and IC compatible processes, paving a way of realizing miniature MEMS oscillators [1]. However, commercialized MEMS oscillators currently available in the timing market are still using two-chip stacking with wire-bonding connections, therefore against the ultimate miniaturization and integration [2][3].

In contrast, CMOS-MEMS based high-Q resonators become a promising candidate for realizing monolithic integration of resonators and circuits [4]. In our previous work, a deep-submicron gap fundamental-mode free-free beam resonator centered at 11.5 MHz was designed and fabricated by utilizing a foundry CMOS-MEMS platform with motional impedance down to 112 kΩ and Q's greater than 1,500 under dc-bias voltage of 60V [5]. However, the previously-designed beam resonator is limited by one-port electrical configuration due to its fundamental vibrating mode shape; therefore, the monolithic CMOS-MEMS oscillator implementation will become very difficult. Moreover, the large dc bias voltage (often greater than 50V) is required to attain acceptable motional impedance for a single resonator due to its insufficient transduction area. In addition, the power handling capability is also restricted by the tiny device size of such a "single" resonator.

To overcome the aforementioned issues without affecting Q, in this work, second-mode free-free beam resonator arrays combined with pull-in mechanism for gap reduction are designed to replace the previous fundamental free-free beam resonator version. The major advantages of resonator array design mainly comes from (i) the large transduction area to effectively reduce the required dc-bias voltage while maintaining reasonable motional impedance and (ii) higher power handling capability compared to that of the single resonator since the maximum handling power is proportional to the effective stiffness of the resonator. Furthermore, the

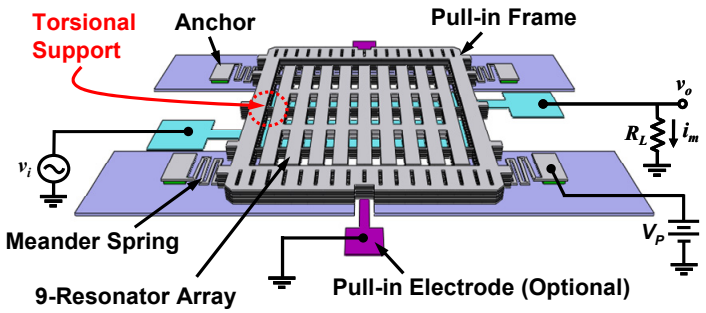


Fig. 1: Perspective view schematic of a 2nd-mode free-free beam resonator array consists of 9 identical resonators in a standard two-port electrical configuration.

This work was supported by National Science Council (NSC) of Taiwan. Travel was supported by IEEE-IFCS. This work was granted the best student paper award at 2011 Joint Conference of the IEEE International Frequency Control Symposium and European Frequency and Time Forum, San Francisco, USA.

two-port electrical configuration proposed in this work not only alleviates the electrical feedthrough between the input and output ports but also separates the signal and dc-bias ports, hence making monolithic CMOS-MEMS oscillator implementation possible. A comprehensive study of CMOS-MEMS resonator arrays is provided in this work to address issues on motional impedance and power handling capability.

II. Array Design and Gap Reduction Mechanism

The perspective view of a deep-submicron gap free-free beam resonator array is shown in Fig. 1 with a typical two-port biasing and excitation configuration. Two parallel electrodes are placed underneath the resonator array with an original gap of around 1μm before pull-in operation for gap reduction. One end of the torsional supporting beams is connected to the outer pull-in frame while the other end is attached to the nodal points of the resonator array to minimize the acoustic loss, hence attaining high Q. As also shown in Fig. 1, the pull-in frame is suspended by four meander springs. Since these meander springs have very low stiffness (i.e., very low pull-in voltage), the optional pull-in electrodes in Fig. 1 can be shorted to ground in a normal operation.

The mechanical resonance frequency of a single free-free beam resonator is determined by its material properties, including equivalent Young's Modulus E and density ρ, and geometric parameters, including beam thickness h_r and length L_r. The vibrating mode shape of a second-mode free-free beam is shown in Fig. 2(a); the corresponding resonance frequency of the beam can be derived using Euler-Bernoulli beam theory, given by

$$f_o = 2.833\sqrt{\frac{E}{\rho}}\frac{h_r}{L_r^2}. \qquad (1)$$

Note that the electrical stiffness is neglected in (1) for simplicity.

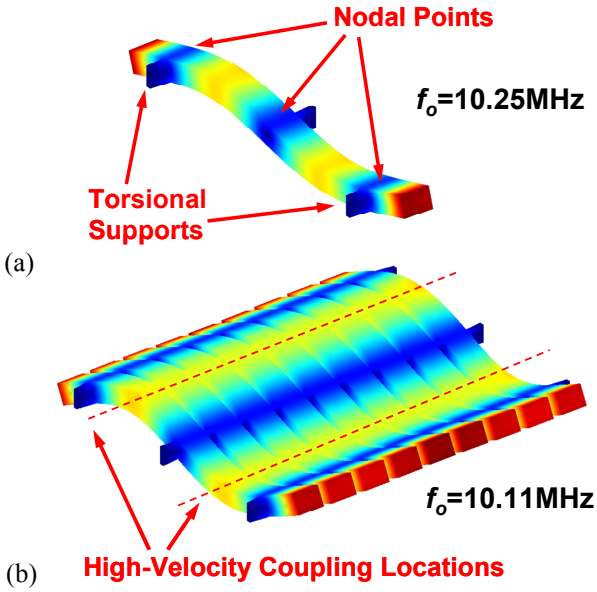


Fig. 2: Finite element simulated mode shapes for (a) a single free-free beam resonator and (b) a 9-free-free beam resonator array. The supporting beams are placed at the nodal locations of both the free-free beam resonator and array to reduce the anchor loss, hence improving the quality factor.

To achieve the mechanically-coupled resonator array, the identical resonators are placed nearby, and each constituent resonator is connected to the adjacent ones by mechanical coupling links, as described in [6]. The resonance frequency of an arrayed resonator is very similar to that of a single resonator if all of the constituent resonators of the array are operated in phase, as shown in Fig. 2(b). However, in general such a multiple degree-of-freedom mechanical system would generate plenty of spurious modes near the desired mode in contrast to single resonance behavior, thereby preventing the implementation of subsequent filter and oscillator applications. To eliminate the spurious modes [7][8], stiff mechanical coupling elements (half-wavelength couplers or short stubs at high-velocity vibrating locations) can be used to effectively push the spurious modes away from the desired frequency; electrical phasing design also helps to select only one vibrating mode and reject all the others. The use of high-velocity coupling and electrical phasing design in this work selects the in-phase vibrating mode of a resonator array as shown in Fig. 2(b) while eliminating all other unwanted modes.

For a capacitively-transduced resonator, the motional impedance R_m and maximum handling power $P_{o\max}$ are governed by the following equations

$$R_m = \frac{k_{re}d_o^4}{\omega_o Q V_P^2 \varepsilon_o^2 A_e^2} \qquad (2)$$

$$P_{o\max} = \frac{\omega_o k_{re}}{Q} X_{\max}^2 \qquad (3)$$

where d_o, k_{re}, ω_o, Q and A_e are electrode-to-resonator gap spacing, equivalent stiffness, radian resonance frequency, quality factor, and electrode-to-resonator overlap area, respectively, and where V_P and ε_o are the dc-polarization voltage and free-space permittivity, respectively. In (3), the parameter $X_{\max}$ denotes the maximum vibration amplitude of the resonator. To reduce the motional impedance R_m and increase the maximum power handling capability $P_{o\max}$ of a resonator, mechanically-coupled array design concept is adopted here to effectively increase both the equivalent stiffness k_{re} and transduction area A_e, hence lowering R_m and

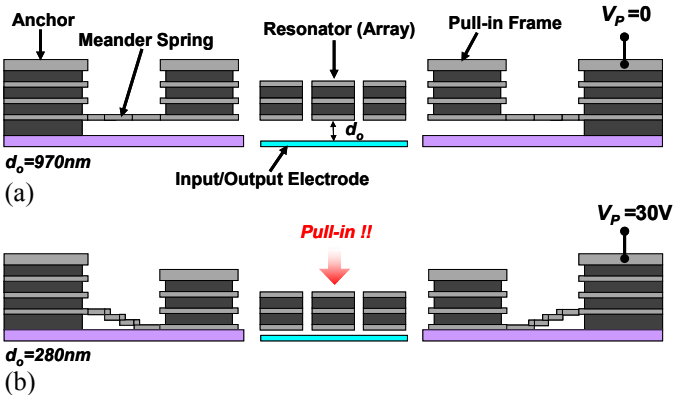


Fig. 3: Cross-section view schematics of the CMOS-MEMS resonator array illustrating resonator status (a) before and (b) after a dc-bias voltage is applied into the resonator structure.

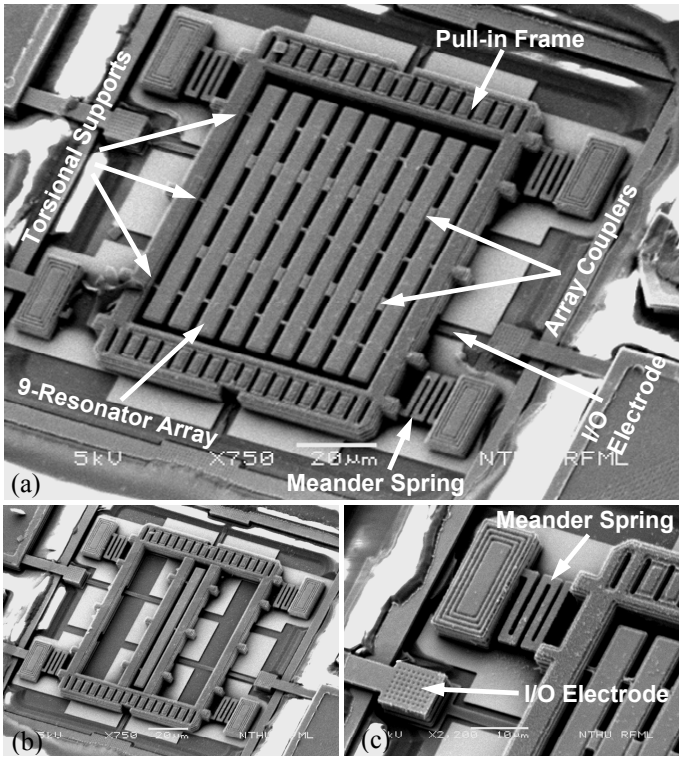


Fig. 4: (a) Global view of a 9-resonator array where the constituent resonator components are located at the center of the pull-in frame. (b) Fabricated single free-free beam resonator using the same pull-in frame. (c) Meander-spring zoomed view of the 9-resonator array.

increasing P_{omax}. To further reduce the motional impedance, the pull-in gap reduction mechanism to enable deep-submicron gaps is an efficient approach to minimize R_m. Fig. 3(a) and (b) present the gap-reduction operation of the resonator array with status (a) before pull-in operation where dc-bias V_P is 0V and (b) after pull-in gap reduction where dc-bias V_P is 30V. The final gap of around 280nm is set by altitude difference of the two polysilicon layers in the 0.35μm CMOS process.

III. Fabrication Process

The free-free beam resonator arrays were fabricated by using a generalized 0.35μm 2P4M CMOS-MEMS platform provided by TSMC followed by a maskless wet-release process [4]. A commercial silicon-dioxide etchant with very high selectivity between metal and oxide was used to perform the maskless oxide release process without the help of critical point dryer. After wet release process, the surrounded sacrificial oxide was removed, leaving the resonator structures together with their pull-in frames suspended above the input/output (I/O) electrodes and pull-in stoppers by the use of the soft meander springs.

Fig. 4(a) presents a global SEM view of a fabricated CMOS-MEMS second-mode free-free beam resonator array containing 9 identical resonators. The array couplers (at high-velocity vibrating locations) and I/O electrodes are also indicated. In this work, a single resonator was also fabricated for comparison with a SEM view shown in Fig. 4(b). The meander-spring zoomed view is also shown in Fig. 4(c). Note

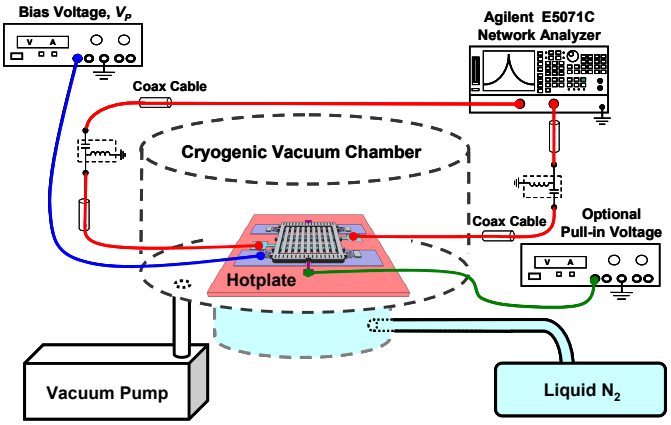


Fig. 5: Two-port measurement setup for a free-free beam resonator array in a cryogenic vacuum probe station.

that this single resonator only has three torsional supports for minimizing the anchor loss while the resonator array in Fig. 4(a) possesses six torsional supports for persevering enough polarization voltage V_P to avoid device breakdown.

IV. Experimental Results

As mentioned above, a typical two-port biasing and excitation configuration is used to characterize the fabricated CMOS-MEMS resonator and array. As shown in Fig. 5, the fabricated resonator was placed in a cryogenic vacuum probe station to measure its capacitive transduction characteristics under a vacuum level of 100μTorr. Since the meander springs of the pull-in frame are of low stiffness, the optional pull-in voltage port shown in Fig. 5 is connected to ground throughout the following measurement results. To operate the device, a dc-bias voltage V_P is applied into the resonator to perform pull-in, as a result producing a deep-submicron electrode-to-resonator gap of around 280nm. Then a small ac signal v_i is applied onto the input electrode. Such a small ac signal v_i together with the dc-bias voltage V_P would generate electrostatic force to drive the resonator into resonance as frequency of v_i matches the resonance frequency of the beam structure, hence sourcing out a motional current i_m to the output electrode. The motional current then goes back to the network analyzer, showing a high-Q frequency response due

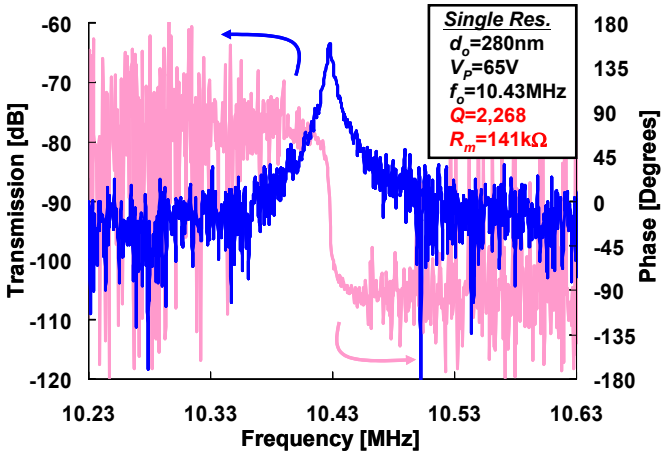


Fig. 6: The measured frequency spectrum of a single resonator under a bias voltage of 65V in a vacuum environment.

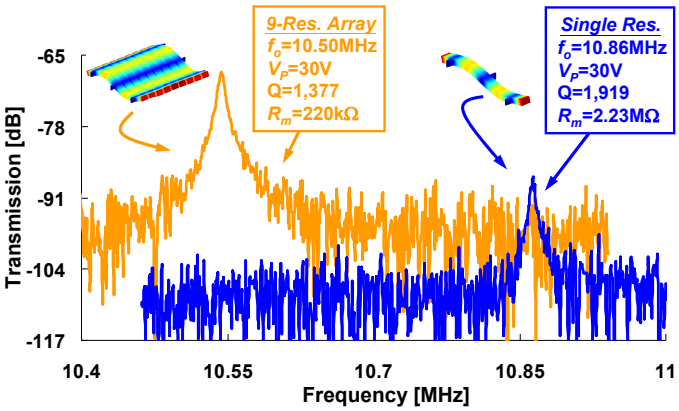


Fig. 7: The measured frequency spectra of single resonator and multi-resonator array under the same bias voltage of 30V in vacuum environment.

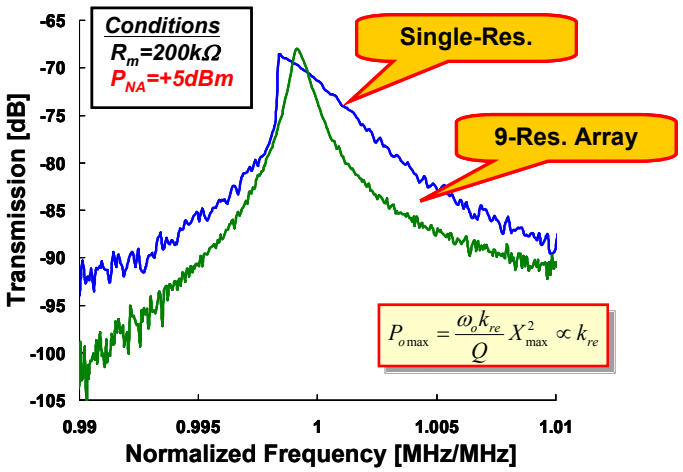


Fig. 8: Comparison of single resonator and 9-resonator array under the same motional impedance and input power level.

to the mechanical vibration.

Fig. 6 presents the measured transmission amplitude and phase plots for a single second-mode free-free beam resonator centered at 10.43 MHz under a dc-bias voltage of 65V in vacuum. The motional impedance of the resonator is less than 150 kΩ with quality factor greater than 2,200, comparable to that of the fundamental mode free-free beam resonators (one-port electrical configuration) fabricated by using the same CMOS-MEMS process [5]. To characterize the performance of the resonator array, Fig. 7 presents measured frequency characteristics of the fabricated single free-free beam resonator and the nine mechanically-coupled resonator array tested under the same bias voltage V_P of 30V and 100μTorr vacuum environment. Obviously, the resonator array exhibits lower motional impedance than that of the single resonator as expected. In theoretical calculation, the motional impedance reduction ratio is proportional to the number (9 in this work) of constituent resonators used in an array. In Fig. 7, motional impedances of 2.23 MΩ and 220 kΩ for single resonator and 9-resonator array, respectively, were measured, showing the reduction ratio is around 10.1 which is in good agreement with the ideal ratio of 9. The discrepancy of R_m reduction ratio mainly comes from the contribution of additional transduction areas between short stubs and I/O electrodes in an array.

To characterize the relative enhancement of the power handling capability for the mechanically-coupled resonator array, single resonator and 9-resonator array were tested under the same transmission conditions, i.e., the same motional impedance and input power level. As shown in Fig. 8, these two resonators with similar motional impedance R_m and quality factor Q were operated under the same driving condition of P_{NA} = 5dBm, showing different measured transmission spectra. P_{NA} denotes the power directly from the network analyzer in Fig. 5. In such a power level, the single resonator was driven into the Duffing nonlinear regime while the 9-resonator array was still far away from its bifurcation point, showing significant power handling improvement using mechanically-coupled array approach.

V. Conclusion

In this work, two-port CMOS-MEMS free-free beam resonator arrays combined with pull-in gap reduction mechanism to enable deep-submicron electrode-to-resonator gap spacing have been successfully demonstrated with preferred electrical configuration (i.e., two-port operation) and much lower dc-bias voltage (only 30V versus more than 60V used in previous single beam resonator) for attaining satisfactory motional impedance. In addition, the power handling capability of the mechanically-coupled resonator array is significant improved, hence paving a way for realizing low phase noise monolithic CMOS-MEMS oscillators in the future.

Acknowledgment

This work was supported by the National Science Council (NSC) of Taiwan under grant of NSC-99-2220-E-007-012. The authors also wish to appreciate the TSMC and National Chip Implementation Center (CIC), Taiwan, for supporting the IC Manufacturing.

References

[1] C. T.-C. Nguyen, "MEMS technology for timing and frequency control," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 54, no. 2, pp. 251-270, Feb. 2007.

[2] Discera, Inc., http://www.discera.com/

[3] SiTime, Inc., http://www.sitime.com/

[4] W.-C. Chen, W. Fang, and S.-S. Li, "A generalized CMOS-MEMS platform for micromechanical resonators monolithically integrated with circuits," *J. Micromech. Microeng.*, 21 (2011) 065012.

[5] W.-C. Chen, W. Fang, and S.-S. Li, "Quasi-linear frequency tuning for CMOS-MEMS resonators," *Proceedings,* 24th Int. IEEE Micro Electro Mechanical Systems Conf. *(MEMS 2011)*, Cancun, Mexico, pp. 784-787, Jan. 2011.

[6] S. Lee and C. T.-C. Nguyen, "Mechanically-coupled micromechanical arrays for improved phase noise," *Proceedings*, IEEE Int. Ultrasonics, Ferroelectrics, and Frequency Control 50th Anniv. Joint Conf., Montreal, Canada, pp. 280-286, Aug. 2004.

[7] M. Demirci and C. T.-C. Nguyen, "Mechanically corner-coupled square microresonator array for reduced series motional resistance," *IEEE/ASME J. Microelectromech. Syst.*, vol. 15, no. 6, pp. 1419-1436, Dec. 2006.

[8] Y.-W. Lin, "Low phase noise micromechanical reference oscillators for wireless communications," *Ph.D. Dissertation*, University of Michigan, Ann Arbor, May 2007.

Bulk Modes in Silicon Crystal Silicon

Gavin K. Ho
ParibX, Inc.
Mountain View, CA USA
gavinho@ieee.org

Siavash Pourkamali
Dept. of Electrical and Computer Eng.
University of Denver
Denver, CO USA

Farrokh Ayazi
School of Electrical and Computer Eng.
Georgia Institute of Technology
Atlanta, GA USA

***Abstract*—Bulk dilation modes in rectangular plates of particular orientations in single crystal silicon are found to have distinct advantages over to those in isotropic materials. Finite element modeling using an anisotropic model reveals plates aligned to the ‹110› direction in (100) silicon have modes which offer excellent capacitive coupling and good isolation to the supports. Low impedance and high *Q* are predicted for select geometries. Experimental results are included, showing a 100-MHz resonator having an unloaded *Q* of 90,000, which demonstrate an *f*-*Q* product of ~1×10^{13}.**

I. Introduction

Micromechanical resonators [1] are a promising alternative to traditional macro-scale frequency control devices. Silicon is an attractive material since it is well characterized, it has deterministic material properties, and it offers low acoustic loss.

Electrostatic micromechanical resonators, also known as capacitive resonators, are attractive for a number of reasons. Since metallization is not necessary on the resonating structure, the resonator can be monolithic in the strictest sense. Very high quality factors, no built-in stress, and excellent long-term stability are possible in silicon capacitor resonators.

The characteristic motional resistance R_1 in early demonstrations of capacitive resonators was 10 kΩ to 1 MΩ. It is desirable for resonators, especially those at higher frequencies, to have lower R_1. Reducing R_1 (*i.e.*, increasing output current from an input voltage) involves increasing electrode area, improving electromechanical coupling, and optimizing the mode shape of the resonator.

High *Q* reduces the close-to-carrier phase noise. Ensuring high *Q* requires minimizing the magnitude of the loss mechanisms. It is necessary to minimize both the intrinsic and extrinsic loss mechanisms. The design-dependent extrinsic losses are reduced through mode isolation (from other modes) and designing for mode shapes which limit energy leakage into the surrounding medium.

Good linearity is also desired in the resonator for a low noise floor. The linearity limit of a resonator is determined by material properties and physical dimensions. It is evident that a large resonator is desirable.

A silicon plate resonator with a large lateral dimension is introduced to satisfy the requirements for low R_1, high *Q*, and good linearity. See Figure 1. As the lateral dimension is increased, it is known that wavelike properties commonly appear in the observed modes. The modes for some isotropic plates are shown in Figure 2.

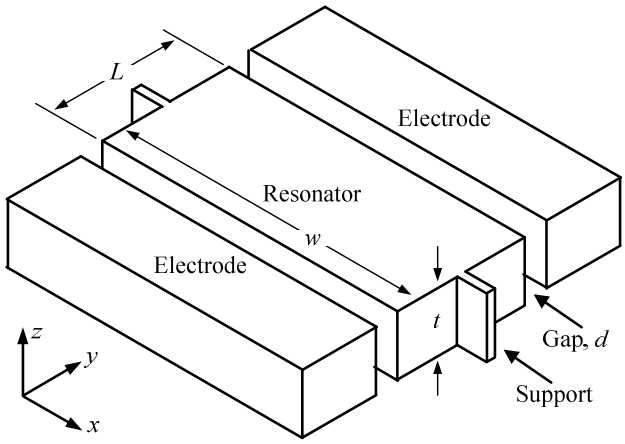


Figure 1. Perspective view of a capacitive plate resonator with two supports and two fixed electrodes.

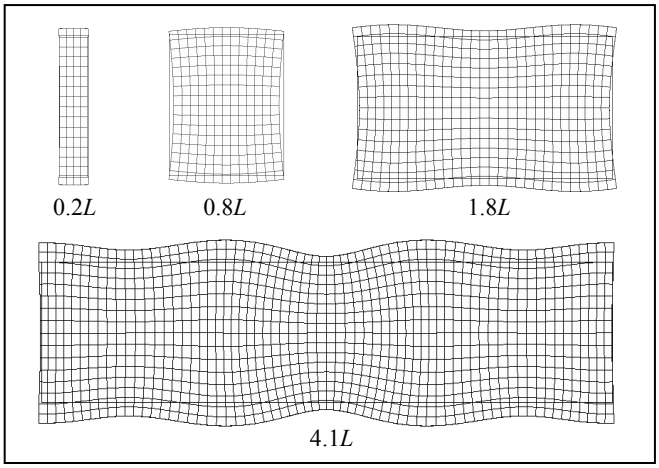


Figure 2. Typical mode shapes in narrow, wide, and very wide rectangular plates in isotropic materials (ν=0.28).

The objective of this work includes: 1) characterizing the dilation modes of rectangular plates in isotropic and anisotropic materials and 2) understanding whether plate modes can be optimized for the metrics of interest. Finite element analysis is performed in ANSYS using SOLID186 elements and isotropic and anisotropic material models.

II. Bulk Dilation Modes

A brief discussion of bulk dilation modes in capacitive resonators follows. The natural frequencies are

$$f_n = \frac{n v_a}{2L}, \tag{1}$$

where n is the mode number, v_a is the acoustic velocity and L is the length of the resonator. The motional resistance of a capacitive bulk dilation mode resonator may be simplified to

$$R_1 = \frac{k_r \pi \sqrt{E_i \rho_m} d^4}{2Q\varepsilon^2 V_P{}^2 A}, \tag{2}$$

where k_r is the relative dynamic stiffness, E_i is the elastic modulus in the i direction, ρ_m is the mass density, d is the gap size, ε is the permittivity of the gap, V_P is the polarization voltage, and A is the overlap area of the electrodes. It is evident that increasing A reduces R_1 – more current is available when the transduction area is increased. Reducing k_r also reduces R_1. An interesting observation is that R_1 is not directly dependent on frequency. Contrary to the trend in prior art, higher frequency resonators do not necessarily have higher R_1.

k_r is a comparison between the dynamic stiffness k_n and the stiffness of a slender ideal 1-dimensional resonator.

$$k_r = \frac{k_n}{k_{n,1D}} \tag{3}$$

It is strongly dependant on the mode shape. One goal of this work is to understand the effect of increasing A on k_r.

The dynamic stiffness k_n for each mode is found from the energy in the mode W and the average of the displacement of the resonator electrode surface normal to the fixed electrode u_y. For a low k_n (and consequently low k_r), all the energy in the mode is preferably attributed to u_y.

$$k_n = \frac{2W}{\overline{u_y}^2} \tag{4}$$

III. Isotropic Plates

The analysis begins by extending the findings on bulk dilation modes of isotropic plates summarized in [2]. In a plate with a small width w (*e.g.*, 0.2L), the mode is entirely uniform along the ends. See Figure 2. As w is increased to 0.8L, some non-uniformity is observed. Lateral motion at the support location exists. For a plate with a width of 1.8L, the most prominent mode is a third mode along the x direction. For a width of 4.1L, the most prominent mode is a seventh-order mode along x. Although the frequencies of these modes are very similar, the latter three modes are in three different modal families.

Focusing on the seventh-order modal family of an isotropic material with a Poisson ratio v of 0.28, observe the attributes of the mode as the width is increased. See Figure 3. For w=3.5L, there are out-of-phase regions which cause charge cancellation. The mode in this geometry has poor coupling and a large k_r. For w=4.1L, the displacement is somewhat uniform. With this geometry, the u_x motion at the support is minimal. For w=4.3L, the displacement is most uniform. This is the optimal plate geometry for coupling and the k_r is 1.2. Further increasing w to 4.8L leads to a more wavelike mode with substantial u_x motion at the support. In these thin isotropic plates, the optimal coupling and optimal support isolation occur for different values of w.

The solutions for isotropic plates having widths of [0,10]L, thickness t of L/5, and v = 0.28 are summarized in Figure 4. k_r and normalized u_x (to the average u_y) are plotted. The minima for k_r, for w>2L, are approximately 1.2. There are geometries in which the support u_x is small. However, u_x is quite sensitive to the width.

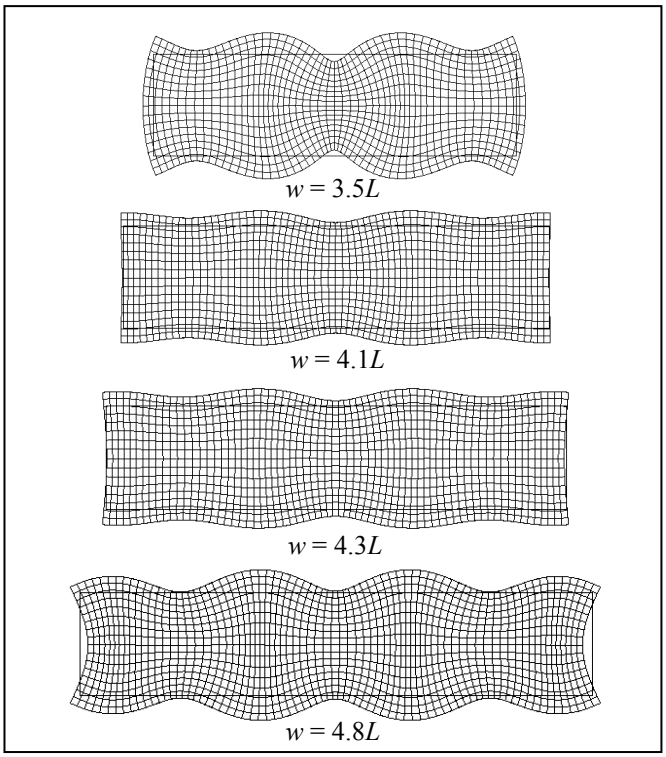


Figure 3. Variation of mode shape within a modal family for width of 3.5L, 4.1L, 4.3L, and 4.8L (isotropic, v=0.28).

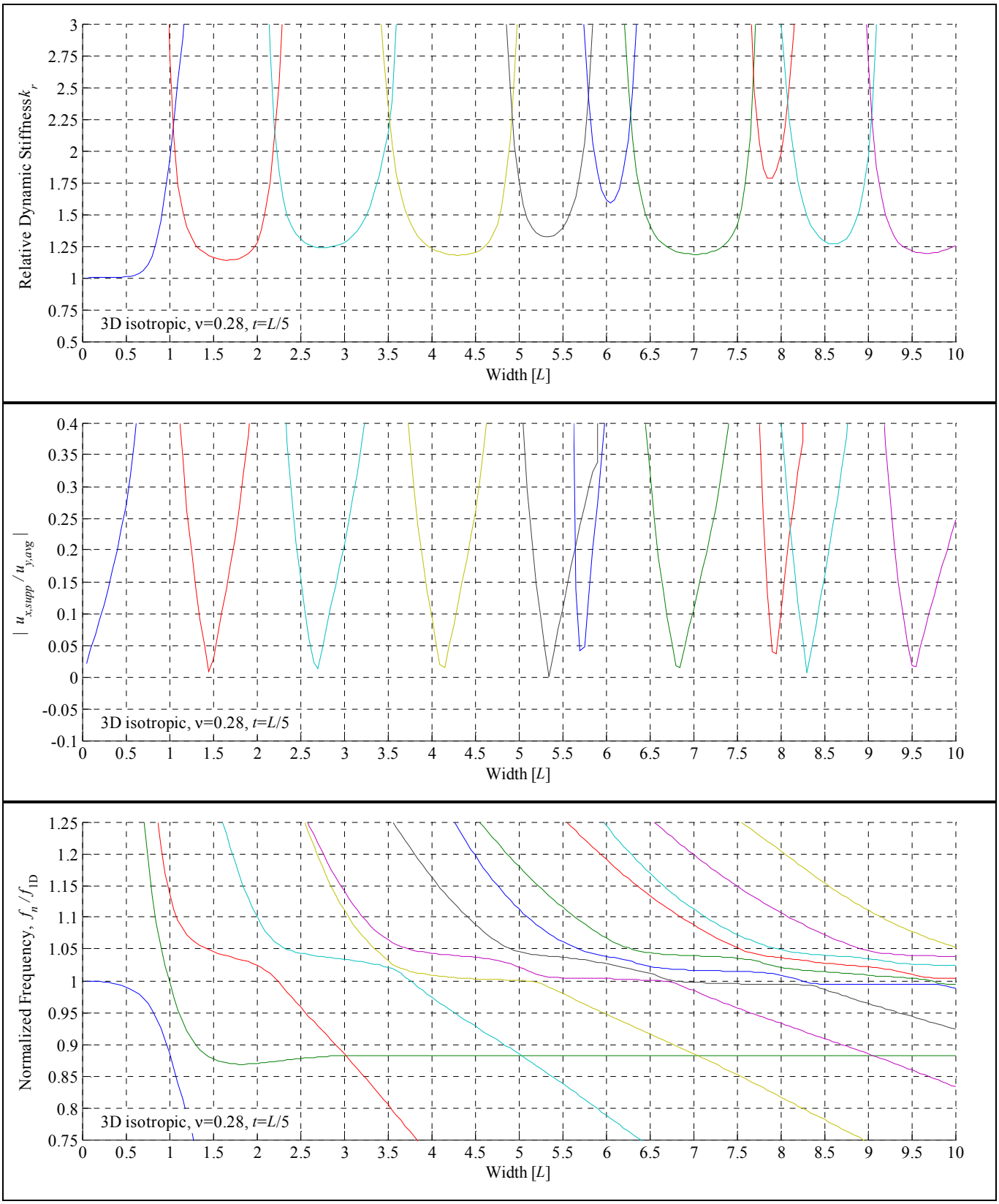


Figure 4. Relative dynamic stiffness, normalized support motion, and normalized frequency in isotropic plates with ν=0.28 and a thickness of L/5.

IV. Anisotropic Plates

The analysis for anisotropic plates is performed similarly. A very interesting observation is that uniform modes are observed in plates aligned to the ‹110› direction in (100) silicon. A thorough analysis was completed for plates with t equal to $L/5$ and $L/2$. The mode shape for $L/5$-thick plates with w=[3.4,4.1]L are very uniform (Figure 5). For w=3.7L, the mode has the smallest u_x response at the support. The k_r and normalized support u_x are shown in Figure 6.

There are several advantages for these anisotropic plates. First, k_r is ~1 for a large variety of widths. Next, u_x is smaller compared to isotropic plate modes. The u_x minima coincide with the k_r minima. Also, u_x has low sensitivity to w.

Increasing the thickness to $L/2$ reveals further advantages. See Figure 7. Compared to $L/5$ plates, the range of widths for which $k_r \approx 1$ is broader. The u_x also have lower sensitivity to w. The most beneficial attribute is that the transduction area is increased without increasing die size.

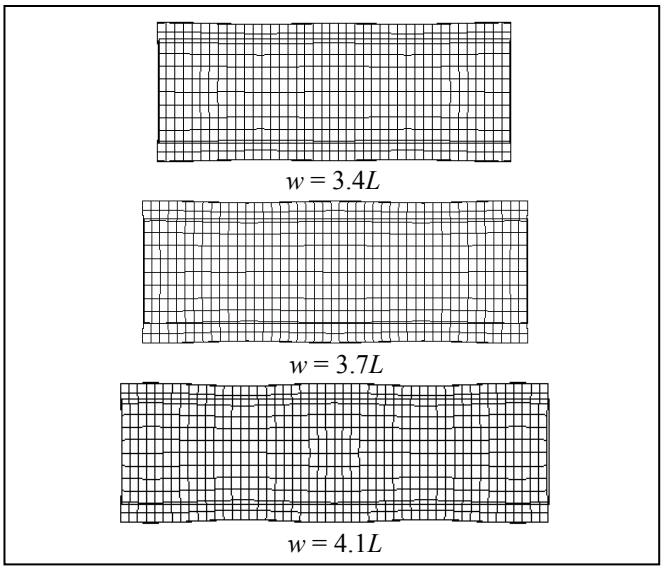


Figure 5. Variation of mode shape within a modal family for width of 3.4L, 3.7L, and 4.1L (anisotropic, t=L/5).

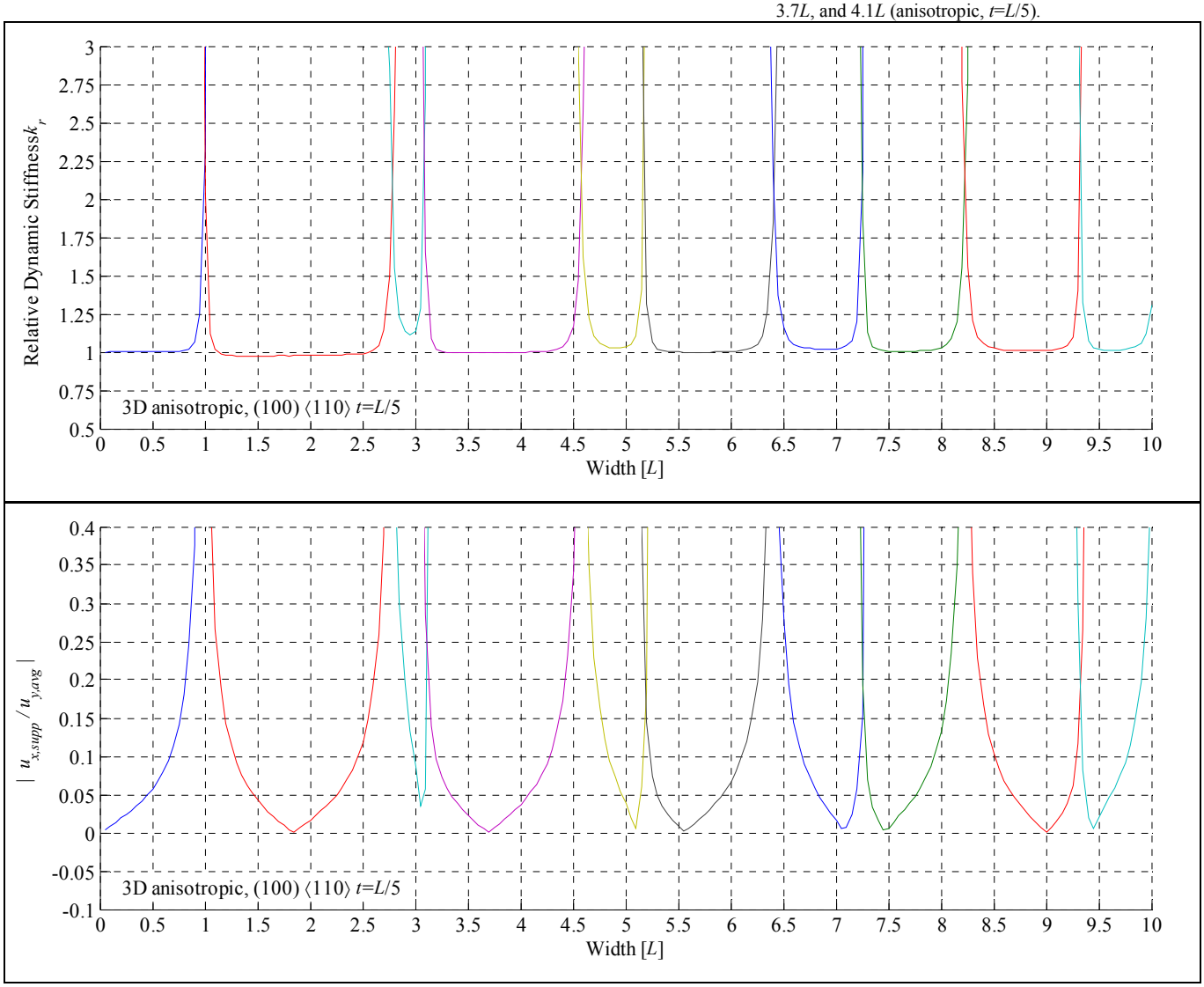


Figure 6. Relative dynamic stiffness and normalized support motion in anisotropic plates aligned to ‹110› in (100) silicon having a thickness of $L/5$.

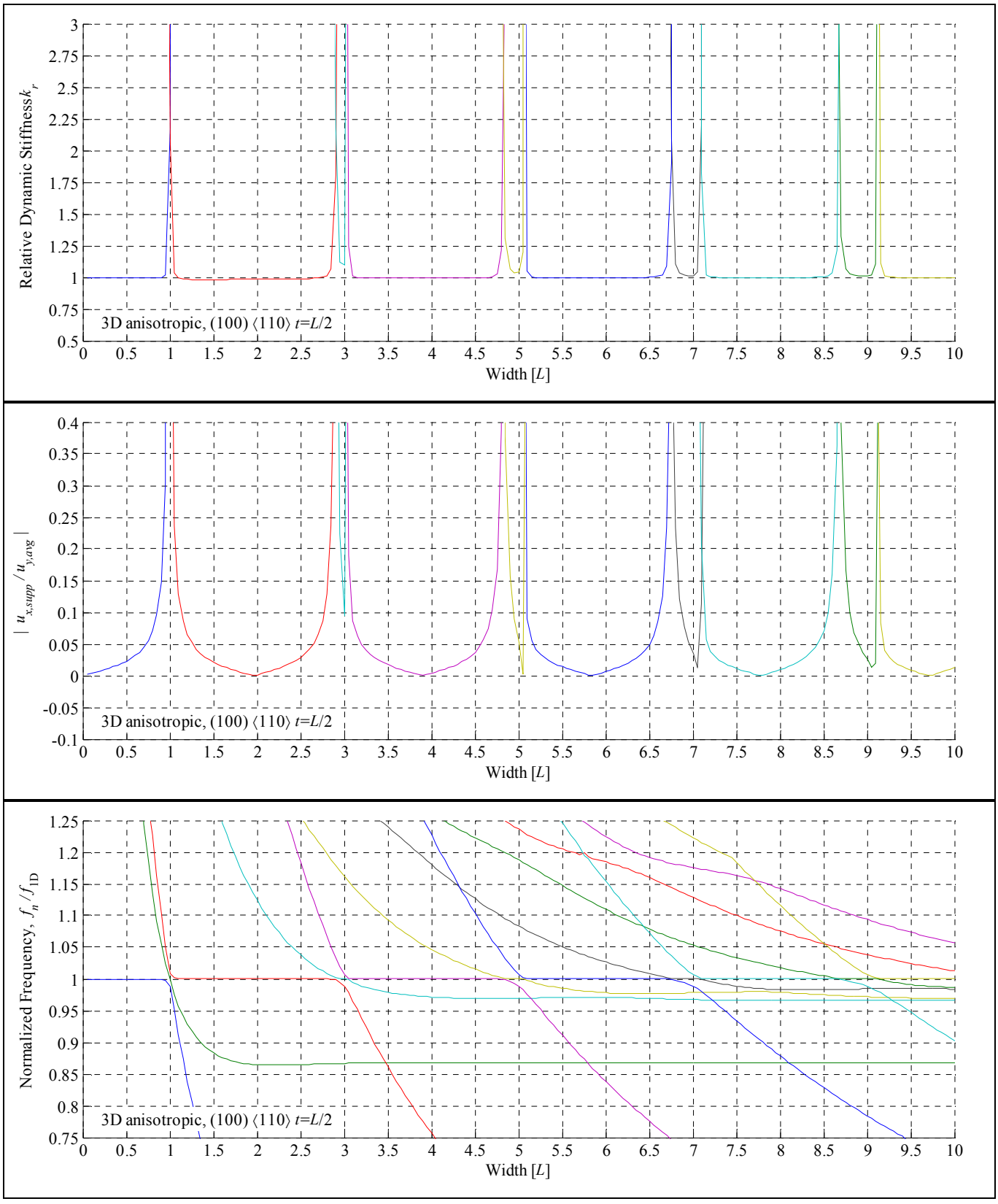


Figure 7. Relative dynamic stiffness, normalized support motion, and normalized frequency in anisotropic plates aligned to ‹110› in (100) silicon having a thickness of L/2.

Periodicity is evident from the results. Widths of even multiples of L provide small motion at the supports and good coupling. The optimal widths for minimizing support loss are 2.0L, 3.9L, 5.8L, 7.7L, and 9.7L.

V. Experimentation

Resonators with L=40μm and t=L/2 were fabricated using the HARPSS-on-SOI process [3]. A resonator with a width of 9.7L and 225-nm gaps is shown in Figure 8. The resonance frequency of this resonator was measured to be 103 MHz along with a quality factor of 90000. See Figure 9. The optimized design enabled the f-Q product of ~1×10^{13}.

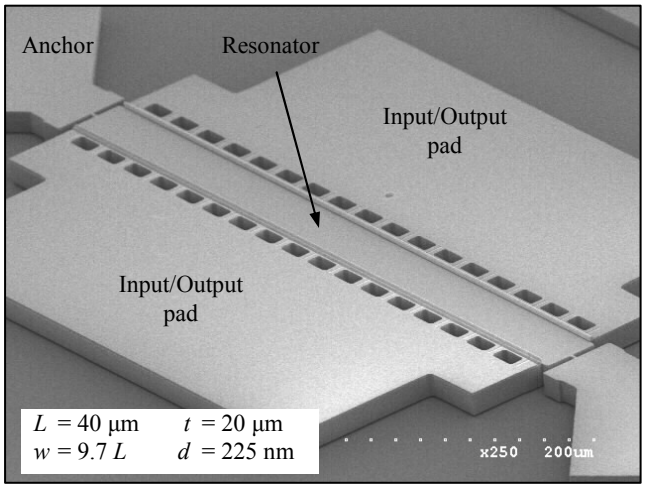


Figure 8. Scanning electron micrograph of a resonator fabricated using the HARPSS-on-SOI process.

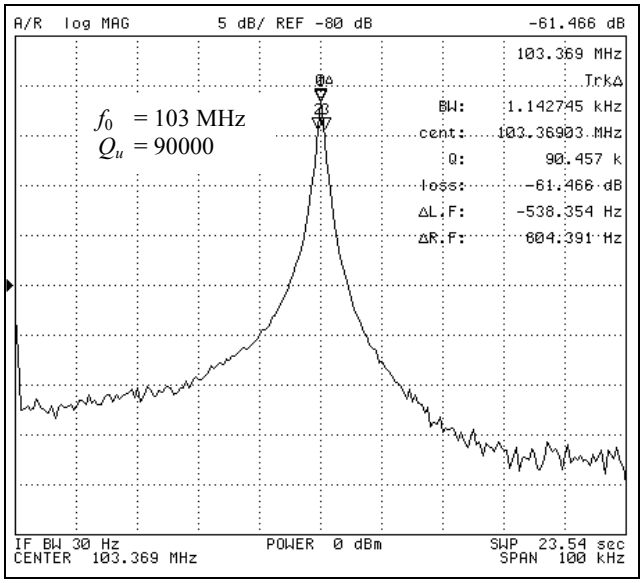


Figure 9. Transmission frequency response of the resonator showing an unloaded Q of 90000 at a frequency of 103 MHz.

Conclusions

Bulk dilation modes of plates aligned to the ‹110› direction in (100) silicon were found to have improved attributes over plates in isotropic materials. In anisotropic plates, the relative dynamic stiffness is unity for many geometries. For a broad range of widths, a 1-dimensional assumption applies. The thorough exercise also revealed specific widths at which the motion of the support location is minimal. This enabled a 100-MHz resonator to demonstrate an f-Q product of 1×10^{13}. Future work includes characterizing modes in plates with larger thickness.

References

[1] C. T.-C. Nguyen, "MEMS technology for timing and frequency control," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, pp. 251-270, Feb. 2007.

[2] R. Holland, "Contour extensional resonant properties of rectangular piezoelectric plates", *IEEE Trans. Son. Ultrason.*, SU-15, n. 2, pp97-105, Apr 1968

[3] S. Pourkamali, G. K. Ho, and F. Ayazi, "Low-Impedance VHF and UHF Capacitive Silicon Bulk Acoustic Wave Resonators – Part I: Concept and Fabrication", *IEEE Trans. Elec. Dev.*, vol. 54, n. 8, pp. 2017-2023, Aug. 2007.

ACES MWL Status and Test Results

M.P. Hess, J. Kehrer,
Astrium Space Transportation, Friedrichshafen, Germany
Email: Marc-Peter.Hess@astrium.eads.net

L. Cacciapuoti, R. Much, R. Nasca
European Space Agency, ESTEC, Noordwijk, The Netherlands

M. Kufner, S. Durand, G. Hejc, H. Frühauf,
TimeTech, Stuttgart, Germany

***Abstract*—Atomic Clock Ensemble in Space (ACES) is a mission using high-performance clocks and links to test fundamental laws of physics in space. The ACES microwave link (MWL) will make the ACES clock signal available to ground laboratories equipped with atomic clocks. The ACES MWL will allow space-to-ground and ground-to-ground comparisons of atomic frequency standards. The MWL comprises the Flight Segment (FS) as part of the ACES payload as well as a distributed set of Ground Terminals (GT), collocated with the ground atomic clocks. MWL is a two-way, two-frequency link transmitting a carrier signal modulated by a PN code, both phase coherent with the local clocks. MWL is designed to reach a time resolution (TDEV) of 0.3 ps in less than 300 s of integration time. In addition to the MWL, operating in microwave domain, ELT is an optical link based on the exchange of laser pulses detected and time stamped in MWL in the local time scale in space and on ground.**

The engineering model of the MWL FS and the first prototypes of the MWL GT will be used for end-to-end testing of the MWL performance. The end-to-end test is using a set-up, where the MWL FS and GT are connected by cables through a test equipment that allows the full simulation of the signal dynamics during an overflight of the ISS over a MWL GT. The MWL design and test results of the individual elements will be briefly reported. The results of the end-to-end test will be presented and perspectives for clock comparisons with ACES discussed.

I. INTRODUCTION

ACES is a distributed system designed to disseminate a high stability and accuracy clock signal [1,2]. It consists of a space payload generating the ACES atomic frequency reference and a network of ground terminals connected to high-performance atomic clocks on ground.

Transported on the International Space Station (ISS) by the Japanese transfer vehicle HTV in the 2014-2015 timeframe, the ACES payload will be installed at the external payload facility of the Columbus module, using the space station robotic arm.

The ACES payload (Fig. 1) accommodates two atomic clocks: PHARAO (acronym of "Projet d'Horloge Atomique par Refroidissement d'Atomes en Orbit"), a primary frequency standard based on samples of laser cooled caesium atoms, and SHM (acronym of "Space Hydrogen Maser"), an active hydrogen maser for space applications. The performances of the two clocks are combined to generate an on-board timescale with the short-term stability of SHM and the long-term stability and accuracy of the PHARAO clock. The on-board comparison of PHARAO and SHM and the distribution of the ACES clock signal are ensured by the Frequency Comparison and Distribution Package (FCDP), while all data handling processes are controlled by the eXternal PayLoad Computer (XPLC). A GNSS receiver installed on the ACES payload and connected to the on-board time scale will provide orbit determination of the ACES clocks. One of the main objectives of the ACES mission consists in maintaining a stable and accurate on-board timescale that will be used to perform space-to-ground as well as ground-to-ground comparisons of atomic frequency standards. The ACES clock signal will be compared to ground clocks using two separate links: a time and frequency transfer link in the microwave domain (MWL) and an optical link (ELT). These comparisons will enable fundamental physics tests and applications in different areas of research.

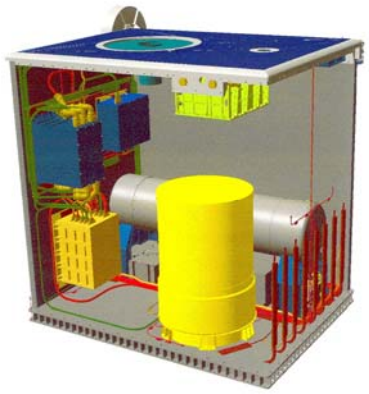


Figure 1. ACES Payload Design. Front: SHM; back: PHARAO, side panels removed for better visibility.

The planned mission duration is 18 months. During the first two weeks, the functionality of the clocks and of MWL will be tested. Then, a period of 6 months will be devoted to

The work presented is performed under ESA contract 16242/02/NL/JS.

the characterization and performance evaluation of the clocks. During this phase, a clock signal with frequency inaccuracy in the 10^{-15} range will be available to ground users. Under microgravity conditions, it will be possible to tune the linewidth of the atomic resonance of PHARAO by two orders of magnitude, down to sub-Hz values (from 11 Hz to 110 mHz). After the clocks optimization, performances in the 10^{-16} range both for frequency instability and inaccuracy are expected. In the second part of the mission (12 months, possibly extended up to 30 months), the on-board clocks will be compared to a number of atomic clocks on ground operating both in the microwave and optical domain.

ACES will perform worldwide comparisons of advanced clocks operating on different atoms or molecules reaching a frequency resolution in the 10^{-17} regime. These measurements will test general relativity and seek for new interactions beyond the Standard Model.

The ACES mission is developed by ESA with ASTRIUM as prime contractor. PHARAO is funded and developed by CNES as contribution to the ACES mission.

The ACES Mission is presently in phase C/D. All instruments and subsystems are in an advanced state of development with engineering models delivered and flight models manufacturing started. The ACES Payload and Ground Segment Critical Design Reviews (CDR) have been successfully passed in 2009 and 2010, respectively.

Development is now proceeding towards a delivery of the ACES payload flight model in 2013. ACES is designed for a mission duration of 1.5 years with the possibility for extension to a minimum of 3 years.

II. ACES Scientific Objectives

ACES will conduct the first experiments with cold atoms under microgravity conditions equivalent to a freely falling laboratory; it will perform fundamental physics tests to high resolution, and develop applications in different areas of research.

A. New Generation of Microwave Clocks for Space

A new generation of space clocks reaching frequency instability and inaccuracy of few parts in 10^{16} will be validated by ACES. PHARAO will combine laser cooling techniques and microgravity conditions to significantly increase the interaction time and consequently reduce the linewidth of the clock transition. Improved stability and better control of systematic effects will be demonstrated in the space environment. PHARAO will have a fractional frequency instability of $1 \cdot 10^{-13} \cdot \tau^{-1/2}$, where τ is the integration time expressed in seconds, and an inaccuracy of few parts in 10^{16}. The reliability offered by active H-masers will be made available for space applications by SHM. SHM will demonstrate a fractional frequency instability of $1.5 \cdot 10^{-15}$ after only 10000 seconds of integration time. Two servo-loops will lock together the clock signals of PHARAO and SHM generating an on-board time scale, combining the short-term stability of the H-maser with the long-term stability and accuracy of the caesium clock (Fig. 2).

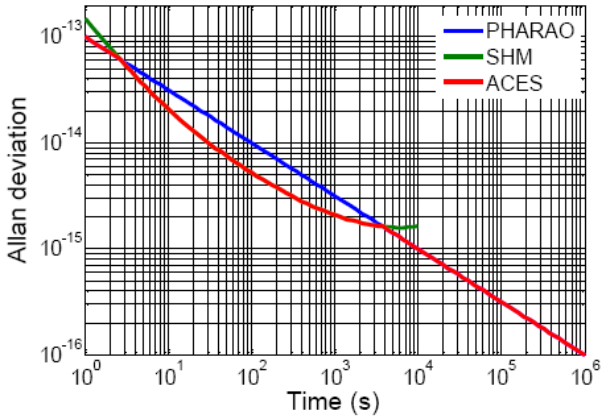


Figure 2. Fractional frequency instability specified for the PHARAO, SHM, and ACES clock signals in space

The ACES clock signal will be distributed via a dedicated Microwave Link (MWL). Frequency transfer with time deviation better than 0.3 ps at 300 s, 7 ps at 1 day, and 23 ps at 10 days of integration time will be demonstrated (Fig. 3). These performances, surpassing existing techniques (TWSTFT and GPS) by one to two orders of magnitude, will enable common view and non-common view comparisons of ground clocks with 10^{-17} frequency resolution after a few days of integration time. Thanks to the recent development of optical frequency combs [3,4], which significantly simplify the link between optical and microwave frequencies, ACES will be able to take full advantage of the progress of optical clocks [5,6], today reaching instability and inaccuracy levels of a few parts in 10^{17}.

ACES will also deliver a global atomic time scale with 10^{-16} accuracy, it will allow clock synchronization at an uncertainty level of a few hundreds of ps, and contribute to international atomic time scales (TAI, UTC…).

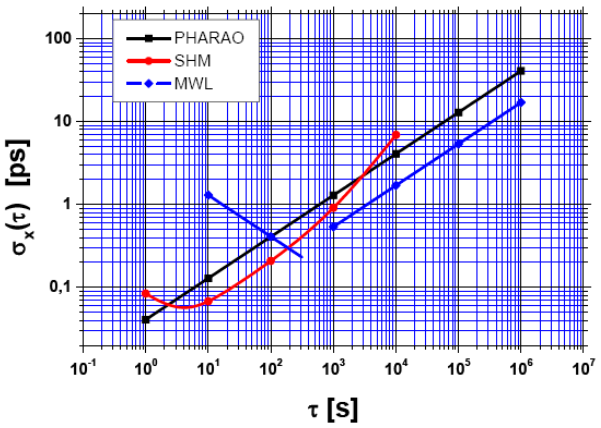


Figure 3. Time instability specified for PHARAO, SHM, and MWL.

B. Fundamental Physics Tests with ACES

According to Einstein's theory of general relativity, identical clocks placed in different gravitational fields experience a frequency shift that, in the frame of the PPN

approximation, depends on the difference between the Newtonian potentials at the clocks positions. The comparison between the ACES on-board clocks and ground-based atomic clocks will measure the frequency variation due to the gravitational red-shift with a 35-fold improvement on previous experiments [7], testing Einstein's prediction at the 2 ppm uncertainty level.

Time variations of fundamental constants can be measured by comparing clocks based on different transitions or different atomic species [8]. Indeed, any transition energy can be expressed in terms of the fine structure constant α and the two dimensionless constants m_q/Λ_{QCD} and m_e/Λ_{QCD}, depending on the quark mass m_q, the electron mass m_e, and the QCD mass scale Λ_{QCD} [9,10]. ACES will perform crossed comparisons of ground clocks both in the microwave and in the optical domain with a frequency resolution of $1\cdot10^{-17}$ in a few days of integration time. These comparisons will impose strong and unambiguous constraints on time variations of fundamental constants reaching an uncertainty of $1\cdot10^{-17}$/year in case of a 1-year mission duration, down to $3\cdot10^{-18}$/year after three years.

The foundations of special relativity lie on the hypothesis of Local Lorentz Invariance (LLI). According to this principle, the outcome of any local test experiment is independent of the velocity of the freely falling apparatus. In 1997, LLI tests based on the measurement of the round-trip speed of light have been performed by comparing clocks on-board GPS satellites to ground hydrogen masers [11]. In such experiments, LLI violations would appear as variations of the speed of light c with the direction and the relative velocity of the clocks. ACES will perform a similar experiment by measuring relative variations of the speed of light at the 10^{-10} uncertainty level.

C. Applications

ACES will also demonstrate a new technique, called "relativistic geodesy", to map the Earth gravitational potential. This technique uses a precision measurement of the Einstein's gravitational red-shift between two clocks to determine the corresponding difference in the local gravitational potentials. The possibility of performing comparisons of ground clocks at the 10^{-17} frequency uncertainty level will allow ACES to resolve geopotential differences at 10 cm.

A dedicated GNSS receiver on-board the ACES payload will ensure orbit determination, important for comparing clocks and performing fundamental physics tests. In addition, the GNSS subsystem will be connected to the ACES clock signal, opening the possibility to use the GNSS network for clock comparisons or remote sensing applications (GNSS radio-occultation and reflectometry).

In addition to the ACES MWL, the European Laser Timing (ELT) link will allow clock comparisons, time transfer, and ranging experiments in the optical domain. The combination of ELT and MWL will provide a bench to test two different time transfer and ranging techniques, also opening the door to studies of atmospheric propagation delays.

III. MWL Status

The ACES clock signal distributed by FCDP is finally transmitted to ground stations by the ACES microwave link. The proposed MWL concept is an upgraded version of the Vessot two-way technique used for the GP-A experiment in 1976 [7] and the PRARE geodesy instrument. The system operates continuously with a carrier frequency in the Ku-band. The high carrier frequencies of the up and down links (13.5 GHz and 14.7 GHz respectively) allow for a noticeable reduction of the ionospheric delay. A third frequency in the S-band (2.2 GHz) is used to determine the Total Electron Content (TEC) and correct for the ionospere time delay. A PN-code modulation (100 Mchip/s) of the carrier removes the phase ambiguity between successive comparison sessions separated by large dead times. The system is designed for multiple access capability, allowing up to 4 simultaneous ground users distinguished by the different PN-codes and Doppler shifts.

A. MWL Flight Segment

The engineering model of the flight segment electronic unit (Fig. 4) has been completed and tested.

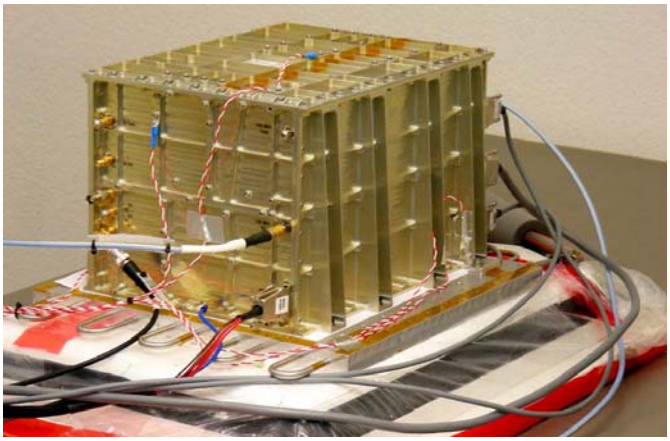


Figure 4. MWL Flight Segment engineering model under test.

MWL long-term stability is ensured by the continuous calibration of the receiver channels provided by a built-in test-loop translator. For shorter durations (< 300 s), time stability is driven by the noise performance of the Ku transmitter and receiver and of the DLL (Delay-Locked Loop) boards. The 100 MHz chip rate has already enabled a time stability better than 2 ps to be achieved with code measurements. Carrier phase stability is shown in Fig. 5, where time deviations down to 80 fs at about 100 s are reported. For longer durations, time deviation remains well below the 1 ps level even in the worst conditions of signal to noise density ratio (C/N), corresponding to very low elevation angles of the ISS over a ground terminal. The thermal sensitivity of the system has been measured and used to calibrate MWL phase comparison data against temperature variations. The sensitivity to a series of key parameters such as clock input power, received signal C/N, supply voltage, Doppler, Doppler rate, etc. has been measured. The susceptibility of the system to narrowband and broadband interference, as well as to multipath effects has been characterized.

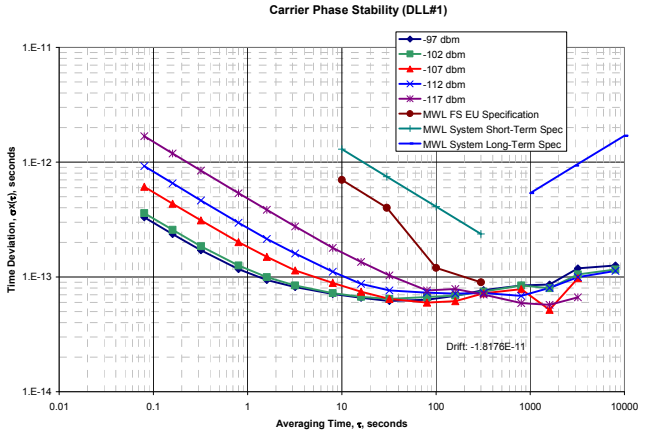


Figure 5. MWL Flight Segment stability of the carrier phase, expressed in time deviation, for different carrier-to-noise power density ratios. Measurement are compared to MWL system requirements and derived MWL FS contributions.

The engineering model of the flight segment S-band and Ku-band antennas (fig. 6) have been completed and tested. The antennas exhibit a extremely flat phase pattern over the complete field of view covering 140°.

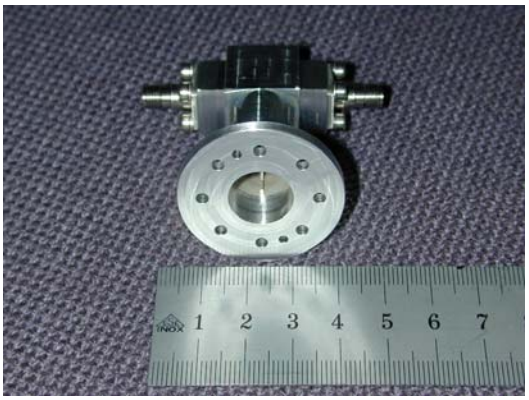


Figure 6. Engineering Model of the MWL Flight Segment Ku-Band antenna.

B. MWL Ground Terminal

The MWL ground terminal electronics is similar to the MWL flight hardware, symmetry being important in a two-way system to reduce instrumental errors. The ACES MWL Ground Terminal (MWL GT) is a microwave station interfacing the local clock on ground to the ACES payload allowing space-to-ground clock comparisons. To reduce phase instabilities due to the tracking motion, the electronic unit of the MWL GT has been rigidly attached to the antenna unit. The Ku-band signal is delivered to the antenna feeder via a waveguide, a high stability RF cable is used for the S-band. The antenna is a 60 cm offset reflector with a dual-band feed system automatically pointed in azimuth and elevation by a steering mechanism. A computer controls the steering unit based on ISS orbit prediction files, collects telemetry and science data both from the local clock and the MWL GT electronics, and interfaces directly with the ACES Users Support and Operation Center (USOC). The system is housed below a protective radome cupola (fig. 8), which also allows to stabilize the temperature of the enclosed volume by an air conditioning system, part of a separate service pallet. The MWL ground terminal is presently being assembled (fig. 7). After dedicated tests on the ground terminal, a MWL system level test will be conducted. The test will be performed using dedicated test equipment that will allow reproducing the signal dynamics both in terms of amplitude and frequency (Doppler) variations.

Figure 7. MWL Ground Terminal under assembly (top: steering unit with antenna and electronics unit, middle, bottom: electronics unit front panel, backplane).

Figure 8. MWL Ground Terminal protective radome

The MWL ground terminals will be collocated with the ground atomic clocks operated by metrology institutes and research laboratories. The hosting site will provide the clock signal to the MWL ground terminal (table 1) ensuring ground clock synchronisation to UTC to 0.5 μs (100 ps after 10 days for at least two of the ground clocks).

The MWL ground terminal comprises the radome pallet including antenna, electronics unit, steering unit and radome and the service pallet including control computer, steering unit controller, power supply (and UPS), and water chiller to thermally stabilise the MWL electronics unit. All these elements are accommodated in an air-conditioned outside cabinet. A dedicated air-conditioning system will maintain and stabilise the temperature inside the radome.

The total weight of system is 270 kg for the radome pallet and 380 kg for the service pallet and air-conditioning system. Power supply demand include about 6 kW for all electronics, thermal chiller and cabinet cooling with an additional 3-5 kW for the radome thermal control, capable to operate the MWL GT in a large range of ambient environments from -30 °C to +45 °C.

TABLE I. CHARACTERISTICS OF THE MWL GROUND TERMINAL TO LOCAL GROUND CLOCK INTERFACE.

Signal	**Characteristics**
Receiver characteristics	Single ended, AC coupled, no galvanic isolation w.r.t. chassis ground
Nominal frequency	100 MHz ± $1 \cdot 10^{-7}$
Waveform	Sine
Input level	6 dBm ± 2 dB
Spurious of input signals	Harmonic: ≤ 69 dBc No harmonic: ≤ 69 dBc in a bandwidth of 100 kHz around the carrier
Nominal input impedance	50 Ω
VSWR	≤ 1.20
S11 stability	≤ $1.5 \cdot 10^{-3}$ pp/day
Connector	3.5 mm female

Signal	PPS
Slope	Rising slope
Trigger level	0-4 V (programmable)
Impedance	50 Ω
Connector	3.5 mm female

C. MWL Data Analysis

MWL provides time-tagged code and carrier phase measurements on the two Ku-band signals (up and down-link) and on the S-band down-link signal. Housekeeping and telemetry data, collected both at the ACES payload and on ground, provide additional calibration parameters (temperature, RF power, etc.) to be applied to MWL raw measurements. ACES scientific products are generated by analyzing MWL raw data, together with additional information, in particular ISS orbitography, parameters for the troposphere model, etc. Data analysis algorithms will be able to deal with the intrinsic noise of the measurements, possible events such as temporary signal loss, etc. and will provide results at the expected performance levels even in adverse conditions. Algorithms include a DRVID (Di erenced Range Versus Integrated Doppler)-based multipath detection. Ionospheric delay calculation will be performed zusing the Ku- and S-band downlink signals, possibly using again DRVID in addition.

For this purpose, a complete model describing the space-to-ground comparison of clocks via MWL has been developed. The model includes relativistic corrections up to third order in *v/c*, which are still non negligible for reaching a frequency uncertainty level well below $1 \cdot 10^{-16}$. Algorithms to measure and control the instabilities introduced by atmospheric propagation delays (ionosphere and troposphere) are included in the code, as well as algorithms for the correct identification of phase ambiguities and cycle slips. The code has been tested on raw MWL data having the same noise behaviour measured during of MWL EM tests and processed by an independent program which also models atmospheric propagation delays and link-related signal attenuation. The results presented in [12,13] show a good agreement with the ACES system requirements.

Based on this work, algorithms are presently developed and implemented in the ACES ground segment infrastructure to operationally exploit the MWL measurement and to generate the ACES data products available to the scientific community.

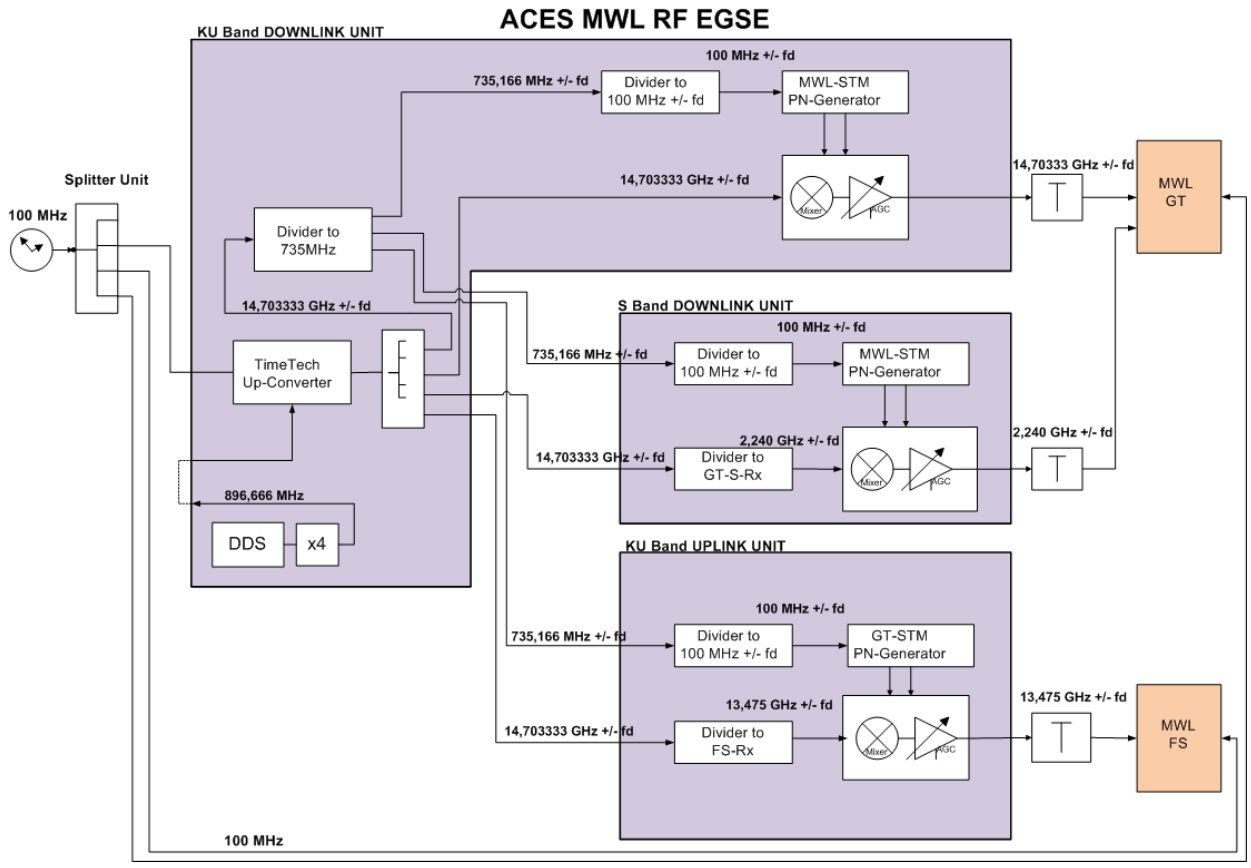


Figure 9. MWL end-to-end test RF EGSE providing full Doppler simulation.

IV. MWL Test Results

A. MWL Test Program

The MWL system is thoroughly tested in incremental steps. The MWL FS and MWL GT electronic units are measured against a dedicated EGSE as well as using their internal test loop translators to determine intrinsic stability. The sensitivity against variations in the environment (temperature, supply voltage, clock signal level, etc.) is independently measured and sensitivity coefficients are validated using realistic variation patterns. The antennas are characterised for gain as well as for their individual phase pattern, allowing to apply corrections in post-processing depending on line-of-sight angles.

Following stand-alone tests, the system is validated in an end-to-end test comprising MWL FS and MWL GT electronic units connected by cables through a signal simulator RF EGSE (fig. 10).

Within the end-to-end test the system stability with Doppler, Doppler rate and variations in signal amplitude is verified. Delay calibration between uplink and downlink is demonstrated in the end-to-end test.

Figure 10. MWL end-to-end test set-up (MWL FS on the right, MWL GT electronics on the left, RF EGSE in the two center racks)

B. MWL End-To-End Test Configuration

The MWL end-to-end test is presently conducted using the MWL FS electronics unit Engineering Model and the MWL GT electronics unit of the first item being build. Both items under test are connected by cable including appropriate and variable attenuation to offset the free-space range. An RF

EGSE simulates Doppler, generated in a DDS and applied coherently on the PN code and carrier signals (fig 9).

The end-to-end test configuration supports both, common clock and dual clock measurements. The common clock configuration will be used to test the MWL intrinsic stability, where the dual clock measurement resembles the operations set-up during the mission; results obtained in this configuration will be compared to measurements performed externally between the two reference clock using a separate phase comparator.

The end-to-end test includes static tests without Doppler and signal amplitude variations followed by Doppler tests. It concludes with a test including the full signal dynamics for a typical ISS passage over a MWL GT.

In the end-to-end test configuration, the differential delay calibration of uplink versus downlink is demonstrated. Delay calibration is performed by comparing the ground clock PPS to a PPS generated internally to the MWL GT.

The end-to-end test data evaluation makes use of extensive data post-processing developed specifically for the test and being a part of the algorithms under development for the ACES mission data exploitation itself.

C. MWL End-to-End Test Results

First results have been obtained from the end-to-end test (fig 11) with constant Doppler.

The figure shows the MWL end-to-end carrier phase stability, expressed in time deviation. The upper curve represents the measurement in the MWL FS of the MWL GT received carrier phase against the local reference signal; measured with DLL2. The lower curve represents the measurement in the MWL GT of the MWL FS received carrier phase against the local reference; measured with DLL1. Results are based on raw data, i.e. before establishing the lambda configuration and applying the two-way formula.

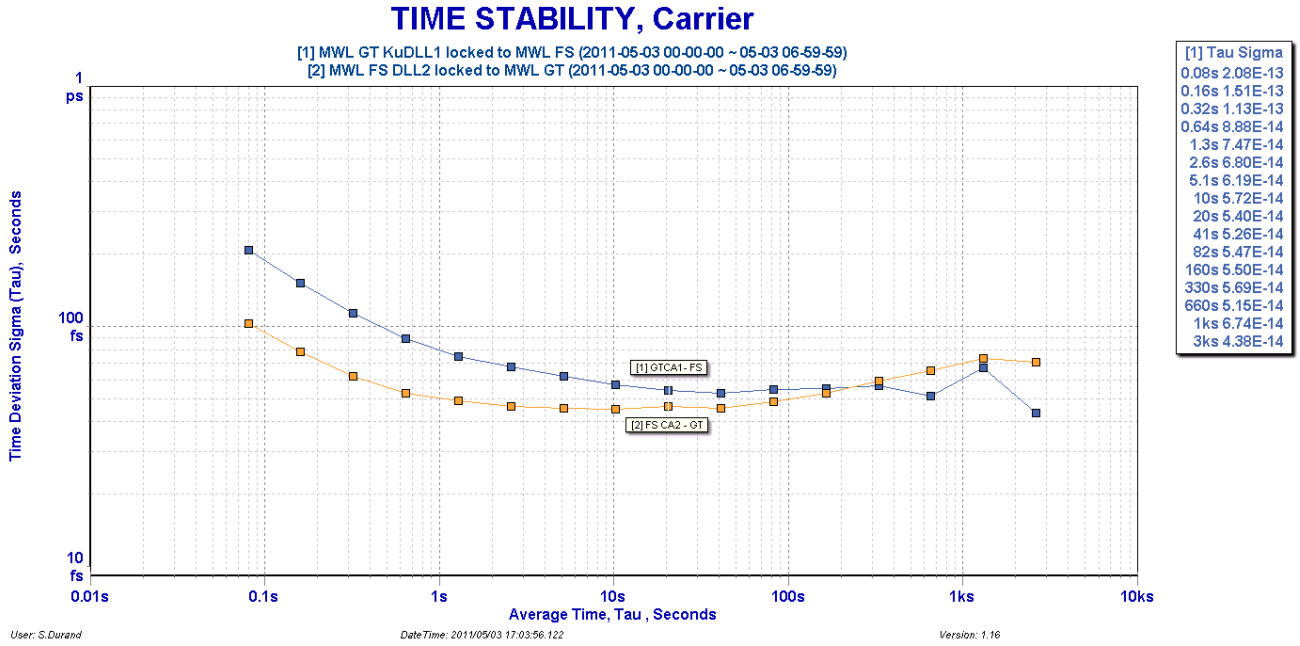


Figure 11. MWL end-to-end test results (carrier phase stability with constant Doppler measured in the MWL FS (blue curve) and in the MWL GT (orange curve)

The system performance reaches stability in the order of 50-60 fs, which is well within the specification requesting 240 fs. The MWL system remains stable up to several hundred seconds of integration time, performance at integration time larger than 1 thousand seconds is dominated by thermal noise in the EGSE.

Remarkably, the system performance is already reached with the code measurements at 300 seconds of integration time (fig 12).

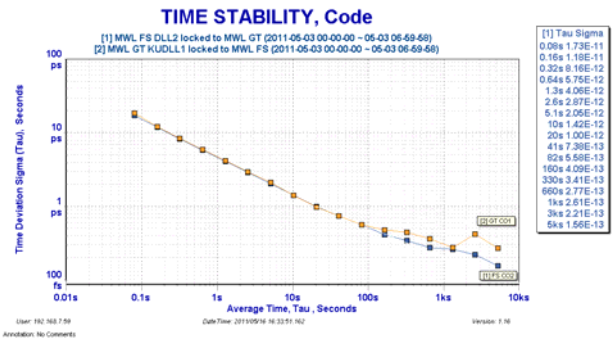


Figure 12. MWL end-to-end test results (code phase stability with constant Doppler measured in the MWL FS (upper curve) and in the MWL GT (lower curve)

The code phase is primarily used to resolve carrier phase ambiguities, the performance is well suited to obtain this goal.

V. OUTLOOK

The MWL end-to-end test is planned to be completed by June 2011. Based on a successful demonstration of performance, the MWL FS flight model will be build and manufacturing of the full set of MWL Ground Terminals will be released.

REFERENCES

[1] C. Salomon et al. (2001). Cold Atoms in Space and Atomic Clocks: ACES. *C. R. Acad. Sci. Paris* **t.2 Séries IV**, 1313.

[2] L. Cacciapuoti and C. Salomon (2009). Space clocks and Fundamental Tests: The ACES Experiment. *Eur. Phys. J. Special Topics* **172**, 57.

[3] R. Holzwarth, et al.(2000). Optical Frequency Synthesizer for Precision Spectroscopy. *Phys. Rev. Lett.* **85**, 2264.

[4] S.A. Diddams, et al.(2000). Direct Link Between Microwave and Optical Frequencies with a 300 THz Femtosecond Laser Comb. *Phys. Rev. Lett.* **84**, 5102.

[5] T. Rosenband, et al. (2008). Frequency Ratio of Al+ and Hg+ Single-Ion Optical Clocks; Metrology at the 17th Decimal Place: *Science* **319**, 1808.

[6] A.D. Ludlow, et al. (2008). Sr Lattice Clock at 1·10–16 Fractional Uncertainty by Remote Optical Evaluation with a Ca Clock: *Science* **319**, 1805.

[7] R.F.C. Vessot, et al. (1980). Test of Relativistic Gravitation with a Space-Borne Hydrogen Maser: *Phys. Rev. Lett.* **45**, 2081.

[8] T.M. Fortier, et al. (2007). Precision Atomic Spectroscopy for Improved Limits on Variation of the Fine Structure Constant and Local Position Invariance: *Phys. Rev. Lett.* **98**, 070801.

[9] V.V. Flambaum, et al. (2004). Limits on Variations of the Quark Masses, QCD Scale, and Fine Structure Constant. *Phys. Rev. D* **69**, 115006.

[10] V.V. Flambaum and A.F. Tedesco (2006). Dependence of Nuclear Magnetic Moments on Quark Masses and Limits on Temporal Variation of Fundamental Constants from Atomic Clock Experiments. *Phys. Rev. C* **73**, 055501.

[11] P. Wolf and G. Petit. (1997). Satellite Test of Special Relativity Using the Global Positioning System. *Phys. Rev. A* **56**, 4405.

[12] L. Duchayne et al. (2009). Orbit Determination for Next Generation Space Clocks. *A&A 504, 653-661 (2009)DOI: 10.1051/0004-6361/200809613*

[13] L. Duchayne et al. (2008). Data Analysis and Phase Ambiguity Removal in the ACES Microwave Link. *Proc. of 2008 IEEE Frequency Control Symposium*, DOI 10.1109/FREQ.2008.4623052.

[14] M.P. Heß et al. (2010). The ACES Mission: System Development and Test Status. *61st International Astronautical Congress*

Single-frequency time and frequency transfer with Galileo E5

Mari Carmen Martínez-Belda
Applied Mathematics Department, University of Alicante
San Vicente del Raspeig, Alicante, Spain
carmen.martinez@ua.es

Pascale Defraigne[1], Quentin Baire and Wim Aerts
Royal Observatory of Belgium (ROB)
Brussels, Belgium
[1]p.defraigne@oma.be

***Abstract*— The GNSS main technique currently used for accurate time and frequency transfer is the so-called "geodetic time transfer", based on a consistent modeling of dual-frequency code and carrier-phase measurements. This technique enables frequency transfer with an uncertainty of 0.1 nanoseconds thanks to the very low noise level of the carrier phase measurements, but the accuracy of time transfer can only reach the level of a couple of nanoseconds in the best cases, due to the noise and multipath of the GPS and GLONASS codes. However, in the near future Galileo will offer one promising signal, namely the broadband signal E5, with an ultimate low range noise in the cm range and with the lowest multipath error impact observed before. This paper investigates the use of this precise code for accurate time transfer.**

I. INTRODUCTION

Measurements from Global Navigation Satellite Systems (GNSS) are used since the eighties to perform precise and accurate Time and Frequency Transfer (TFT) [1-2]. Nowadays, the main technique used for TFT is the so-called "geodetic time transfer", based on the joint analysis of GPS (and GLONASS) dual-frequency code and carrier-phase measurements. It is widely recognized for its high frequency stability (see for instance [3-5]), thanks to the very low noise level of the carrier phases, enabling time transfer with an uncertainty of 0.1 ns (nanoseconds), when ignoring the uncertainty of the instrumental hardware delays calibration. This calibration value can be determined with an uncertainty at the level of about 1 ns for GPS links, as it has been shown in recent experiments [6-7]. However, the time transfer accuracy is also limited by the colored signature of the code noise, affecting the medium-term stability of the solution and inducing possible discontinuities at the day boundaries [8-10]. Therefore, the accuracy of time transfer given by this technique can only reach the level of a couple of nanoseconds in the best cases.

On the other hand, the European Galileo system is currently under development, and will be interoperable with GPS and GLONASS. Galileo signal is on the air since December 2005, when the first experimental Galileo satellite GIOVE-A was launched. Presently, there are two experimental Galileo satellites in operation, GIOVE-A and GIOVE-B, and the first four operational Galileo satellites (building the Galileo In-Orbit Validation (IOV) constellation) are due for launch in 2011. GIOVE satellites transmit the ranging signals using all the code modulations currently foreseen for the future Galileo, and provide therefore a foretaste of their performance in real-life applications. Due to the use of a higher PRN chip rate of 10.23 MHz and two subcarriers spaced at 30.69 MHz resulting in a 51 MHz wide signal bandwidth, the ranging signals of Galileo provide significant improvement of the (long) multipath performance as compared to current GPS/GLONASS. In 2008, some studies with GIOVE-A data evidenced that the E5AltBOC code demonstrates the highest multipath suppression as compared to other signals and very low magnitude of average multipath errors, down to the values about 20 cm [11]. Previous research done by the authors of this paper [12] showed that this very precise GNSS pseudorange is very promising for improving time transfer applications.

The purpose of this article is to develop accurate methods for time transfer with Galileo E5AltBOC (E5 hereafter). In section 2 we revise the current GNSS main time transfer technique; section 3 presents the noise level of the E5 code using the experimental GIOVE satellites observations and investigates the possible approaches to use it for time transfer; the fourth section presents some results from tests with Galileo simulated data, and the last section is dedicated to the conclusions of this work and perspectives for the future.

II. STATE OF THE ART

The current geodetic time transfer method is based on a consistent analysis of ionosphere-free combinations for the code and carrier-phase measurements on L1 and L2 bands, that is,

$$P_3 = \alpha_f \cdot P_1 - \beta_f \cdot P_2,$$
$$L_3 = \alpha_f \cdot L_1 - \beta_f \cdot L_2,$$

978-1-61284-110-6/11/$26.00 ©2011 IEEE

with

$$\alpha_f = \frac{f_1^2}{f_1^2 - f_2^2} \approx 2.545, \quad \beta_f = \frac{f2}{f_1^2 - f_2^2} \approx 1.545,$$

which are finally given by

$$\begin{aligned} P_3 &= \| x_{rec} - x^{sat} \| - c\Delta t_{rec} + c\Delta t^{sat} + Trop + \delta_3^P + \varepsilon_3^P, \\ L_3 &= \| x_{rec} - x^{sat} \| - c\Delta t_{rec} + c\Delta t^{sat} + Trop + \lambda_3 N_3 + \delta_3^\phi + \varepsilon_3^\phi, \end{aligned} \quad (1)$$

where x_{rec} is the receiver position, x^{sat} is the satellite position, c is the vacuum speed of light, Δt_{rec} is the receiver synchronization clock error, Δt^{sat} is the satellite synchronization clock error, *Trop* is the tropospheric delay, λ_3 is the wavelength of the carrier-phase combination, N_3 is the integer phase ambiguity of L3, δ_3^P contains the receiver hardware delays for the P3 code, δ_3^ϕ is the initial phase delay of the phase combination, ε_3^P is the pseudorange noise and multipath, and ε_3^ϕ is the ionosphere-free carrier-phase noise.

The carrier phases themselves allow to give a precise signal evolution, at the subnanosecond level, as already mentioned, but they need to be combined with code data, due to the fact that they contain an unknown initial ambiguity (integer number of cycles, see (1)) which must be solved. The P3 code is used for this task, determining then the absolute receiver clock offset. The noise of the ionosphere-free P3 combination is given by the expression

$$\sigma_{P_3} = \sqrt{2.545^2 \cdot \sigma_{P_1}^2 + 1.545^2 \cdot \sigma_{P_2}^2} \approx 3 \cdot \sigma_{P_1}.$$

The ionosphere-free combination therefore amplifies the pseudorange noise and multipath with respect to the individual code measurements. As a consequence, the accuracy of the time transfer solution is at the level of up to a couple of nanoseconds.

The most important parameter for time transfer is Δt_{rec}, i.e. the receiver clock synchronization error with respect to the reference of the satellite clocks. When the receiver frequency is driven by or synchronized with an external atomic clock frequency, Δt_{rec} gives indeed access to this external clock. The solution for Δt_{rec} is obtained from a least-squares fitting of the code P3 and carrier phase L3 observations. Using a Precise Point Positioning (PPP) approach [14-15], it is possible to determine at the same time the Δt_{rec}, the station position, the tropospheric zenith delay *Trop*, and the float carrier phase ambiguities (λ_3 N_3+ δ_3^ϕ); the satellite positions and clocks are provided by external products, e.g. the IGS products. Note that δ_3^P will be absorbed in the clock solution Δt_{rec}, and should be corrected for using external calibration parameters (receiver, antenna and cable delays).

III. USING THE GALILEO E5ALTBOC PSEUDORANGE FOR TIME TRANSFER

A. E5 observation equation

The observation equation for E5 is quite similar to the observation equation for P3, with an additional term containing the ionospheric delay at frequency E5:

$$E_5 = \| x_{rec} - x^{sat} \| - c\Delta t_{rec} + c\Delta t^{sat} + Iono_5 + Trop + \delta_5^P + \varepsilon_5^P, \quad (2)$$

where δ_5^P contains the receiver hardware delays for this frequency and ε_5^P is the pseudorange noise and multipath.

In order to obtain the receiver clock error, Δt_{rec}, all the other contributions to the E5 code must be corrected. As explained before, the satellite positions and clocks can be provided by external products. The troposphere zenith delay can be either obtained from external products (if such products exist for the station where the clock is located), or estimated as explained before for PPP. The ionospheric delay is the main limitation and requires being determined at the highest possible precision level. Different ways can be used to correct for the ionospheric delay. Combining the E5 code with another code in order to perform an ionosphere-free solution, as done with the GPS/GLONASS P3, will cause the loss of the properties of this very precise code, due to the noise of the second pseudorange used for the combination. This will thus offer no advantage with respect to the present geodetic time transfer approach. Alternative approaches exist and are presented in the next section.

B. Ionosphere correction of the E5 pseudorange measurement

1) Ionosphere correction for single-frequency receivers

Having only single-frequency measurements at our disposal requires external products for the ionosphere correction. It is therefore necessary to use maps providing the Total Electron Content (TEC) as a function of the latitude and longitude, in order to get the ionospheric delay for each signal path. Using global TEC maps proposed in the IONEX format, as for example those provided by the IGS analysis center CODE [16], however only corrects for the long wavelength and long term variations (above 2 hours) of the TEC. It was demonstrated with GPS that this is not sufficient for time and frequency transfer: the ionosphere-free combination P3 improves time transfer by a factor of 2 with respect to using P1 code corrected with the CODE TEC maps, for intercontinental baselines [17].

On the other hand, regional ionosphere maps, currently under development (as for instance the ones produced by the US National Oceanic and Atmospheric Administration (NOAA)), offer a better spatial and temporal resolution with respect to the global maps, as they are built with more dense networks. They could then provide a more precise correction for the ionospheric delay on pseudoranges; this will be tested in the future.

Another approach to correct the ionosphere could be to estimate this effect by means of an appropriate model from the Code-minus-Carrier (CMC) observable. The latter is obtained from the difference between the E5 code and carrier-phase measurements, and retains only the ionosphere correction and the carrier-phase inherent ambiguity term. Indeed, the L5 carrier-phase observation equation reads:

$$L_5 = \| x_{rec} - x^{sat} \| - c\Delta t_{rec} + c\Delta t^{sat} - Iono_5 + Trop + \\ + \lambda_5 N_5 + \delta_5^{\phi} + \varepsilon_5^{\phi}. \quad (3)$$

Note that the ionosphere contribution $Iono_5$ contains here only the first order (i.e. in $1/f^2$ where f is the frequency), which affects the code and carrier phase with the same magnitude but opposite signs. By differencing equations (2) and (3), we obtain the CMC combination:

$$CMC = \frac{E_5 - L_5}{2} = Iono_5 + \frac{\delta_5^P}{2} - \frac{(\lambda_5 N_5 + \delta_5^{\Phi})}{2} + \varepsilon_{CMC}. \quad (4)$$

CMC therefore contains only the ionospheric delay, plus an ambiguity term (constant over each continuous satellite track) and a constant term. The ionospheric delay, $Iono_5$, can be estimated by using a single-layer spherical ionospheric shell model [18], i.e.,

$$CMC = MF \cdot \left(ZID_0 + \nabla_{\varphi} \cdot (\varphi_{IP} - \varphi_0) + \nabla_{\lambda} \cdot (\lambda_{IP} - \lambda_0) \right) \\ + \frac{\delta_5^P}{2} - \frac{(\lambda_5 N_5 + \delta_5^{\phi})}{2} + \varepsilon_{CMC},$$

where MF is the elevation-dependent mapping function, ZID_0 denotes the zenith ionospheric delay at the station position, ∇_{φ} and ∇_{λ} are latitudinal and longitudinal gradients, respectively, φ_{IP} and λ_{IP} are the geodetic latitude and longitude of the ionosphere pierce point (IPP); φ_0 and λ_0 are the latitude and longitude of the station. Using a least- squares fitting of all observations above a given elevation, it is possible to determine the TEC at the zenith plus latitudinal and longitudinal gradients. We tested this approach using simulated Galileo data (see later), and using the mapping function $MF=1/\cos z$, where z denotes the zenith angle of the satellite. The differences between the VTEC retrieved with this model and the simulated VTEC can reach the level of some TEC units (see Fig. 1), corresponding to more than one nanosecond for the GNSS frequencies. The technique can therefore not be used for our purpose as not enough precise.

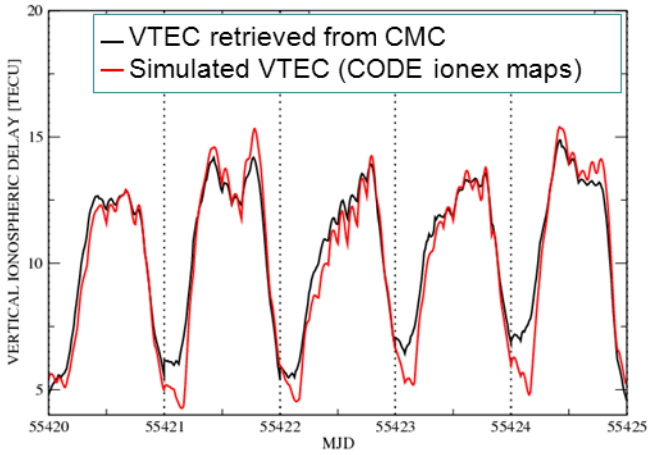


Figure 1. Comparison between the vertical ionospheric delay in TEC units for a 5-day period estimated with surface fitting based on a ionospheric shell model from CMC , and the simulated VTEC (taken from CODE IONEX maps).

2) Ionosphere correction for dual-frequency receivers

If dual-frequency data are available, a combination of carrier phases measured on two frequencies can be used to get the first order ionospheric delay for each signal path. First of all, as the ionosphere is a dispersive medium, the ionospheric delays for different frequencies f_i and f_j are related by the expression

$$\frac{Iono_i}{f_j^2} = \frac{Iono_j}{f_i^2}.$$

Then, the geometry-free combination of L1 and L5 carrier-phase observations can be expressed as

$$L_{GF} = L_1 - L_5 = Iono_5 \left(1 - \frac{f_5^2}{f_1^2} \right) + N'_{GF} + \varepsilon_{GF}^{\phi}, \quad (5)$$

where N'_{GF} is an ambiguity parameter that includes the difference between the ambiguities in the two frequencies, plus some biases. In order to estimate the ionosphere, this ambiguity must be first solved for. The latter can be done by means of the geometry-free combination of E1 and E5 codes, which retains the ionospheric delay plus a collection of biases of the receiver and satellite hardware (the so-called differential code biases, or DCBs), that is,

$$E_{GF} = E_5 - E_1 = Iono_5 \left(1 - \frac{f_5^2}{f_1^2} \right) - c(DCB_{rec} + DCB^{sat}) + \varepsilon_{GF}^{P}. \quad (6)$$

Then, by making the difference between (5) and (6) we can estimate the ambiguity parameter as

$$N'_{GF} = \left\langle (L_{GF} - E_{GF}) \right\rangle_{\substack{arc\ without \\ cycle\ slips}} - c(DCB_{rec} + DCB^{sat}).$$

And finally, we can obtain the ionospheric delay of each E5 measurement as

$$Iono_5 = \frac{1}{\alpha_{GF}}\left\{ L_{GF} - \left[\left\langle (L_{GF} - E_{GF}) \right\rangle_{\substack{arc\ without \\ cycle\ slips}} - \right.\right. \quad (7)$$
$$\left.\left. -c(DCB_{rec} + DCB^{sat}) \right] \right\},$$

where $\alpha_{GF} = 1 - f_5^2 / f_1^2$.

C. The E5 Code-Plus-Carrier combination

Instead of using the E5 observation only, we propose here to take advantage of the Code-plus-Carrier observable (CPC), which is at first-order ionosphere-free. By adding equation (3) to (2), we get CPC, which reads

$$CPC = \frac{E_5 + L_5}{2} = \| x_{rec} - x^{sat} \| - c\Delta t_{rec} + c\Delta t^{sat} + Trop + \quad (8)$$
$$+ \frac{\delta_5^P}{2} + \frac{\left(\lambda_5 N_5 + \delta_5^\phi\right)}{2} + \varepsilon_{CPC}.$$

It is important to note that the noise and multipath ε_{CPC} of this combination are mainly the noise and multipath of the pseudoranges, as the carrier-phase noise is significantly lower. Then,

$$\sigma_{CPC} = \frac{1}{2}\sqrt{\sigma_{E_5}^2 + \sigma_{L_5}^2} \approx \frac{\sigma_{E_5}}{2},$$

that is, CPC has half the noise and multipath of the E5 code. Therefore, the time transfer solution obtained with CPC should have a noise reduction by a factor of 2 with respect to the solution obtained with E5 only. Moreover, the CPC combination also mitigates higher-order ionospheric effects, with a reduction of 25% of the second-order delay (*~1/f³*) and 33% of the third-order delay (*~1/f⁴*) with respect to the code observable [19]. However, the CPC combination is ambiguous, so that the time transfer method based on this combination will have to deal with the ambiguity resolution in the first step.

In [12] we presented some results of the time transfer solution obtained with CPC once the ambiguities were solved, with current data available from GIOVE satellites. These are reported in Fig. 2, which presents for the time link GNOR-GUSN, i.e. Noordwijk (Netherlands)-USNO (US), the clock solutions obtained with either the CPC, or the ionosphere-free codes using the combination of E1 and E5AltBOC (labelled as E5(a+b)) for GIOVE, and P1 and P2 for GPS. The results are shown for one satellite track of GIOVE-B and of GPS PRN 19, these two satellites appearing visible at the same time for both stations. The clock solutions have been computed using pseudorange observations at a 30 second sampling rate, and correcting them for the geometric distance satellite-receiver, for the satellite clock and for the relativistic effect due to the satellite orbit eccentricity, using the IGS orbits for GPS and the CONGO orbits [20] for GIOVE-B. The tropospheric delays were taken from the IGS products [19] for USNO and for DLFT (25 km far from GNOR). The station positions were fixed to *a priori* known coordinates. The clock solutions are presented in Common View in order to get rid of the reference time scale which is different for GPS and for GIOVE satellites clock products.

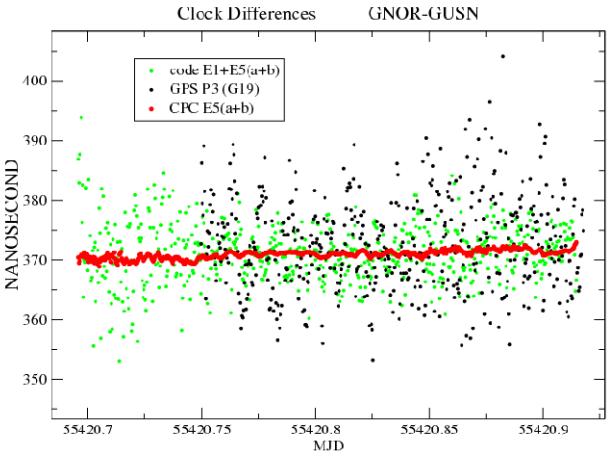


Figure 2. From [12]: Comparison between the clock solutions obtained with either dual-frequency ionosphere-free combination of pseudoranges, or the CPC combination.

The improvement in terms of noise level obtained with the CPC with respect to the classical dual-frequency ionosphere-free codes appears clearly. The rms of the GPS P3 and of the GIOVE E1+E5 combinations are about 7.8 and 4.8 ns, respectively, whereas the rms of the CPC E5 falls down to 0.4 ns. That means that the latter has a noise level about 10 times lower than the ionosphere-free E1+E5 and more than 15 times lower than GPS P3. The small variations of the CPC curve correspond to E5 multipath.

If we do not consider the instrumental calibration issues, the clock solutions so-obtained correspond to the quantity which determines the accuracy of the time transfer solutions. The advantage of the CPC will be therefore to improve the accuracy of time transfer with respect to the dual-frequency ionosphere-free combination, i.e. with CGGTTS (Common GPS GLONASS Time Transfer Standard) or PPP, provided that the ambiguities of the CPC are precisely determined. In order to do that, we can for instance use the Code-minus-Carrier combination (4), as it has the same ambiguity that appears in CPC (see (8)). However, this approach requires again the correction of the ionosphere effect, as with the E5 code only approach. To that purpose, the double-frequency geometry-free combination of carrier-phase and code measurements on E1 and E5 bands can be used.

Let's recall now that the time transfer solution with CPC E5 should be 2 times less noisy than that using only E5. This will be verified in the next section, where we investigate the two approaches that have been presented, with data simulated with the ASiF software [13]. To summarise, these two approaches are:

<u>E5-only approach:</u>

1. - Estimate the ionospheric delays from the geometry-free code and carrier-phase combinations of E1 and E5.

2. - Correct the E5 observable with the previous value and then, solve for the clock solution with a least-squares fitting: either determine the receiver clock, the station position and the wet troposphere path delays; or fix the position and troposphere, and determine only the receiver clock.

CPC E5 approach:

1. - Estimate the ionospheric delays from the geometry-free code and carrier-phase combinations of E1 and E5.

2. - Correct the ionosphere in the CMC equation to determine the ambiguities.

3. – Enter these ambiguities in the CPC observable and then, solve it as in the E5-only approach.

IV. Clock Results From Simulated Data

The ASiF software [13] simulates measurements in RINEX 3.0 format and satellite products in clk formats for the GPS and Galileo constellations. The satellite orbits are modelled with a keplerian motion; receiver and satellite clocks are simulated using the procedure proposed in [21]; tropospheric delays are taken from the IGS products [22]; ionospheric delays are generated from CODE IONEX maps [16]; a white noise is then added to the measurements. In the present study, we considered pseudorange white noise with a standard deviation of 6 cm for E5 and 13 cm for E1, with no multipath effect, and white noise with standard deviation of 13 mm for the carrier phases.

Following the two procedures previously described, we estimated the receiver clock using a least-squares inversion as in the Atomium PPP software [23], i.e. with an analysis of one-day data batches. The clock solution was computed either fixing the position and troposphere to *a priori* known values, or determining them in the same analysis. Figs. 3 and 4 show the difference between the simulated clock and the clock retrieved by the least-squares inversion with the two approaches (E5-only or CPC E5), for 5 consecutive days. The first conclusion that arises from these results is that, as expected, the noise of the clock solution obtained from the CPC combination is half the noise level of the solution obtained directly from the E5 pseudoranges. Secondly, we see that estimating also the troposphere and the position, as it is done in classical PPP, induces errors on the clock solution of the order of about 50 ps in the present case. Finally, we can consider that the accuracy of our clock solution is given by the average difference between the computed and simulated clocks over one day. This difference is smaller than 20 picoseconds in both cases. However, it is important to note that only white noise has been considered for E5 and E1 pseudoranges. Although E5 is not much influenced by multipath, this effect should be also taken into account in the future, as it will directly impact the clock solution accuracy. In the case of E1, or any other different from E5, the noise and multipath are more significant, and will then affect the precision on the ionosphere correction (see (7)). Moreover, the simulation did not consider different hardware delays for each frequency, which play also a very important role in the ionosphere determination. The uncertainty on these parameters will therefore be a parameter influencing the accuracy of the clock solution and will require further investigations.

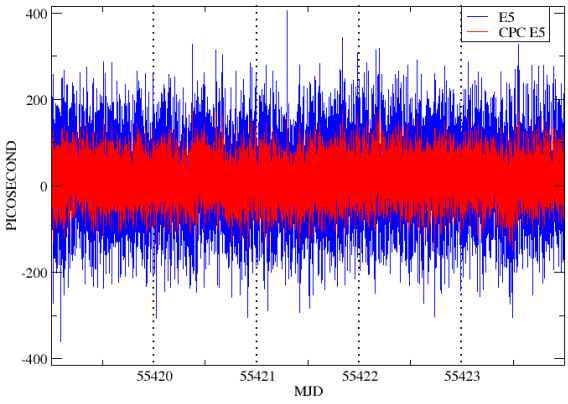


Figure 3. Differences between the computed and simulated clock; the computed clock was obtained using either the E5 code directly or the E5 CPC combination. The position and tropospheric delays are fixed to *a priori* known values.

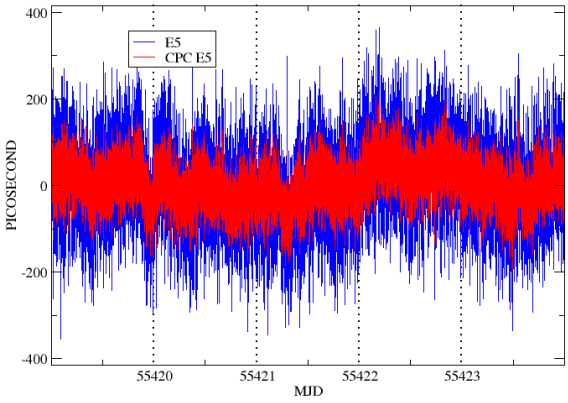


Figure 4. Same as figure 3, but here the position and tropospheric delays are also estimated in the least-squares analysis.

V. Conclusions

This paper investigated the applications of the future Galileo code E5AltBOC for time transfer. Thanks to its high precision and low multipath, this code can offer large improvement to the time transfer accuracy. The sum of code and carrier phase measurements (called CPC for Code plus Carrier) on E5AltBOC has been proved to be two times less noisy than the E5AltBOC code; its use for time transfer shows therefore a big interest. Some comparisons with true GPS and GIOVE data demonstrated that the CPC combination is more than 15 times less noisy than GPS P3. As the CPC is an ambiguous observable, the combination CMC (Code Minus Phase) has been proposed to solve for the ambiguities. This

combination needs however to be first corrected for the ionospheric effect, which can be precisely determined only from dual-frequency code and carrier-phase measurements. Some tests with simulated data have shown that the accuracy of the E5-based clock solution, measured as the mean offset between the computed and simulated clocks, is lower than 20 ps. However, the impact on the clock solution of multipath and of the precision of the differential code biases which were not considered here, need to be carefully investigated in future work.

ACKNOWLEDGMENT

This work has been partially supported by Spanish Science and Technology Ministry (Project AYA2007-67546), and by the Solar - Terrestrial Centre of Excellence (STCE, 2011). The authors also acknowledge the IGS for their data and products used in this study

REFERENCES

[1] D.W. Allan, and M. Weiss, "Accurate time and frequency transfer during common-view of a GPS satellite", Proc. 1980 IEEE Freq. Contr. Symp., Philadelphia, PA, pp. 334-356, 1980.

[2] P. Defraigne, P. Banerjee, and W. Lewandowski, "Time transfer through GPS", Indian Journal of Radio & Space Physics, vol. 36, August 2007, pp. 303-312.

[3] Th. Schildknecht, G. Beutler and M. Rotacher, "Towards sub-nanosecond GPS time transfer using geodetic processing technique", Proc. of the 4th EF TF, 1990, pp. 335-346.

[4] K.M. Larson, J. Levine, L.M. Nelson, and T. Parker, "Assessment of GPS carrier-phase stability for time-transfer applications". IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 47(2), pp. 484-494, 2000.

[5] C. Bruyninx, and P. Defraigne, "Frequency Transfer Using GPS Codes and Phases: Short and Long Term Stability", Proc. of the 31th PTTI meeting, Dana Point, California, December 1999, Ed. L.A. Breakiron, 2000, pp. 471-478.

[6] H. Esteban, J. Palacio, F.J. Galindo, T. Feldman, A. Bauch, and D. Piester, "A GPS calibration trip experience between ROA and PTB", Proc. IEEE FCS-EFTF, 2009.

[7] A. Niessner, W. Mache, B. Blanzano, O. Koudelka, J. Becker, D. Piester, Z. Jiang, and F. Arias, "Calibration of the BEV GPS receiver by using TWSTFT", Proc. 40th PTTI Meeting, 2008.

[8] P. Defraigne, and C. Bruyninx "On the link between GPS pseudorange noise and day-boundary discontinuities in geodetic time transfer solutions", GPS solutions, 2007, vol. 11(4), pp. 239-249.

[9] K. Senior, and J. Ray, "Accuracy and precision of carrier-phase clock estimates", Proc. 33rd PTTI meeting, Long Beach, CA, November 2001, pp. 199-217.

[10] J. Ray, and K. Senior, "IGS/BIPM Pilot Project: GPS carrier phase for time/frequency transfer and time scale formation", Metrologia**,** vol. 40**,** S270-88.

[11] A. Simsky, J.-M. Sleewaegen, M. Hollreiser, and M. Crisci, "Experimental results for the multipath performance of Galileo signals transmitted by GIOVE-A satellite", Int. J. of Nav. and Obs., vol. 2008, article ID 416380, 13 pages, 2008.

[12] P. Defraigne, and M.C. Martínez-Belda, "Time Galileo Common View: format, processing and tests with GIOVE",Proc. 42nd PTTI meeting, Reston, Virginia, November 2010.

[13] Q. Baire, P. Defraigne, and E. Pottiaux, "Influence of Troposphere in PPP Time Transfer", Proc. of EFTF 2009.

[14] J.F. Zumberge, M.B. Heflin, D.C. Jefferson, M.M. Watkinsand F.H. Webb, "Precise point positioning for the efficient and robust analysis of GPS data from large networks", Journal of Geophysical Research, vol. 102(B3), pp. 5005-5017, 1997.

[15] J. Kouba, and P. Heroux, "GPS Precise Point Positionning using GPS orbit products", GPS solutions, vol. 5, pp. 12-28, 2001.

[16] http://aiuws.unibe.ch/ionosphere

[17] P. Defraigne, and C. Bruyninx, "GPS time and frequency transfer: state of the art", Proc. "Journées 2005 Systèmes de Référence Spatio-Temporels", Eds. A. Brzezinski, N. Capitaine and B. Kolaczek , Space Research Centre PAS, Warsaw, Poland, 2006.

[18] R.F. Leandro, M.C. Santos, and R.B. Langley, "Analyzing GNSS data in precise point positioning software", GPS Solutions, DOI 10.1007/s10291-010-0173-9, 2010.

[19] T. Schüler, H. Diessongo, and Y. Pku-Gyamfi, "Precise ionosphere-free single-frequency GNSS positioning", GPS Solut., DOI 10.1007/s10291-010-0177-5, 2010.

[20] O. Montenbruck, A. Hauschild, U. Hessels, P. Steigenberger, and U. Hugentobler, "CONGO - First GPS/GIOVE Tracking Network for Science, Research", GPS World, September 2009, pp. 36-41, 2009.

[21] G. Panfilo, and P. Tavella, "Algorithms for the Atomic Clock Prediction Within the GALILEO System", Proc. of 1st Colloquium Scientific and Fundamental Aspects of the Galileo Program, 1-4 October, Toulouse, France, 2007.

[22] ftp://cddis.nasa.gov/gps/products/troposphere/zpd

[23] P. Defraigne, N. Guyennon, and C. Bruyninx, "GPS Time and Frequency Transfer: PPP and Phase-Only Analysis", Int. J. of Nav. and Obs., vol. 2008, Article ID 175468, 7 pages, 2008.

New realization strategy for UTC(PTB)

A. Bauch, E. Staliuniene, D. Piester, W. Yang*
Physikalisch-Technische Bundesanstalt (PTB)
Bundesallee 100, 38116 Braunschweig, Germany
andreas.bauch@ptb.de

***Abstract*—The Physikalisch-Technische Bundesanstalt (PTB) realizes its atomic time scale UTC(PTB) as approximation to the international time reference UTC. It serves as the basis for PTB's time services, for the local clock comparisons and for international time comparisons. Since February 2010 UTC(PTB) has been realized using an active hydrogen maser steered in frequency via a phase micro stepper according to an algorithm described in this paper. Thereby the long-term stability and accuracy of PTB's primary clocks is combined with the typical short-term frequency stability of a hydrogen maser. During the last 12 months, the caesium fountain clock CSF1 was used as the steering reference on more than 98% of the days, which resulted in an excellent stability of UTC(PTB). During the last six months, including March 2011, its deviation from UTC was less than 4 ns.**

I. INTRODUCTION

The German Units and Time Act entrusts the Physikalisch-Technische Bundesanstalt PTB to realize UTC(PTB), a representation of Coordinated Universal Time UTC, as the basis for legal time in the country. It serves thus as the reference for PTB's time dissemination services, for the local clock comparisons and for international time comparisons. According to recommendations of the Consultative Committee for Time and Frequency the deviation of UTC(k) from UTC, where k stands for any laboratory active in this field, should be as small as possible, preferably below 100 ns. No unique prescription for the realization of UTC(k) follows thereof. UTC(PTB) used to have its physical representation as standard frequency and one pulse per second (1 PPS) signals based on its primary clock CS2 for the last 20 years [1]. Since February 2010 UTC(PTB) has instead been realized using an active hydrogen maser (AHM) steered in frequency via a phase micro stepper (PMS) according to an algorithm described below. Thereby the long-term stability and accuracy of PTB's primary clocks is combined with the short-term frequency stability of a hydrogen maser.

Subsequently we lay down the algorithm in current use (Section II) and briefly describe the practical realization (Section III). In Section IV results are presented. A discussion and outlook on further work closes the paper.

* on leave from NUDT, ChangSha. PR China

II. CONSTRAINTS ON THE REALIZATION OF UTC(PTB)

A. Available equipment

PTB currently operates two cold-atom based caesium fountain clocks, CSF1 [2] and CSF2 [3], two older thermal beam primary clocks, CS1 and CS2 [1], three commercial caesium clocks of type Symmetricom 5071 (high-performance option), and three AHM produced by Vremya-Ch. One motivation of going to a new scheme of realization of UTC(k) was to make better use of the available resources in the laboratory. As long as UTC(PTB) was based on CS2 alone, it already had a good reputation as one of the most long-term stable and predictable time scales. A deviation from UTC by more than 50 ns occurred only for a few ten days during the last 10 years. The day-to-day stability was, however, worse than that of time scales of other institutes realized with hydrogen masers as the signal source. This fact became even more obvious since two-way satellite time transfer and GPS carrier phase-based time transfer became standard methods in the field, and was considered no longer state-of-the-art. A final motivation was a rather technical one: The CS2 ovens were almost depleted of caesium, and clock operation could not be assured beyond some point in time in 2010.

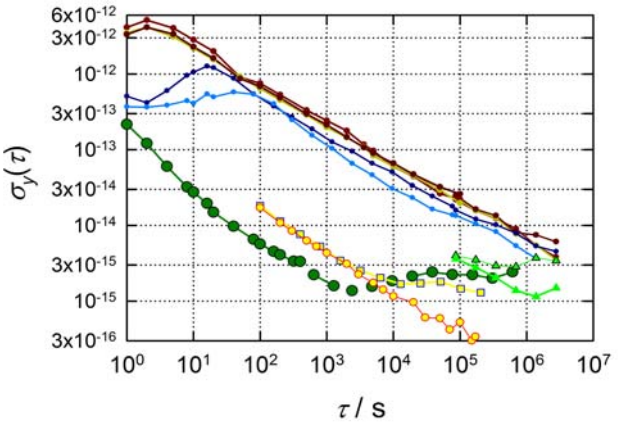


Figure 1. Relative frequency instability of PTB's atomic clocks, detailed explanations in the text

In Figure 1, the frequency instability of the available frequency standards in PTB is depicted. It illustrates that the performance of the various devices is quite different. Results of AHM comparisons (green symbols) are shown for local comparisons among two of them with averaging times between 1 s and 10 days, and for comparisons of the same two masers with the time scale UTC(NIST) of the US National Institute of Standards and Technology for averaging times between 1 day and 30 days. The performance of the fountain CSF1 is shown in yellow symbols, with respect to an AHM (squares) and to an optical frequency standard of PTB [4] (circles) with an instability superior to CSF1. The data span analyzed here is longer than that reported in Figure 3 of [4]. These data advised us to steer the maser frequency through the PMS towards CSF1 at least once per day based on daily average values of the frequency difference. The frequency instability of all other available standards (CS1: dark blue, CS2: light blue, three commercial clocks: earth tones) equals that of the masers only after far more than 10 days of averaging. Thus daily steering the maser frequency towards those references has been based on linear fits to a number of daily average values, extrapolated to the current day.

B. *Basic algorithm design*

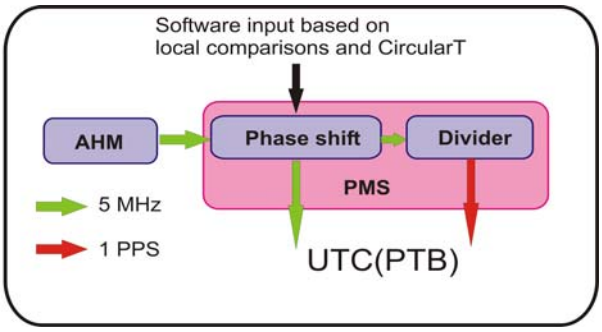


Figure 2. UTC(PTB) realization from a hydrogen maser

In the past, a frequency offset between UTC(PTB) and CS2 used to be introduced through a phase micro stepper (PMS) so that agreement of UTC(PTB) and UTC was assured in the long term. This offset was introduced manually once every few months based on the results published in the BIPM Time Department Circular T. Now this "steering of the AHM" is under software control as illustrated in Figure 2. The PMS is commanded daily to shift the 5 MHz input frequency by the relative amount δf_{steer} which has three additive components, δf_1, δf_2 and δf_3. The first component, δf_1, is intended to adjust the frequency of the maser to the rate of the reference clocks as just described. The second term, δf_2, represents the individual reference clock rates with respect to TAI. Three steering values $\delta f_1 + \delta f_2$ are calculated in parallel each day and one is selected based on a pre-selected priority and the availability of data:

Option 1: The AHM frequency is compared to PTB's fountain clock CSF1 hourly. If during day N-1 6 hours of data are available, then the mean value, corrected for the daily frequency drift of the maser is used as δf_1 during day N. When no valid data are available for day N-1, the steering for day N is calculated from a linear fit to the last three daily values. If less than two such data are available (e.g. two extrapolated, but not measured data), the steering is calculated based on another option. For δf_2 we use the estimated d, representing the fractional deviation of the scale interval of TAI from the SI second on the geoid, as published in the last line of the Circular T Section 4.

Option 2: The AHM frequency is compared to all caesium beam clocks. δf_1 is updated daily, based on extrapolation of past data for an interval adapted to the frequency instability of the reference clocks, mostly 25 days. To estimate δf_2 the monthly published rates rTAI of the clocks with respect to TAI from the ALGOS data base are used. Individual clock results are combined using their most recent statistical weights wTAI as published in the ALGOS data base.

Option 3: As a fall-back option, steering is made based either on CS2 or CS1, similar to Option 2.

The third term δf_3 is common to all three options. It assures the long term steering of UTC(PTB) to UTC and is currently calculated as $\{\text{UTC-UTC(PTB)}\}_{\text{LRD}}$ / 60 d, where the time difference is taken at the last reported day in the most recent issue of the Circular T, LRD. The values of δf_3 (and also δf_2) are thus constant for about one month after a new issue of the Circular T was received.

C. *How to ensure reliability and availability?*

The first steps of realizing a time scale based on CSF1 were made in 2002 [5]. Since then we collected a lot of experiences of equipment failures and weaknesses of the chosen strategies. Figure 2 should make clear that a maser output failure is fatal. Therefore the realization of the time scale is made in duplicate, named UTC(PTB) and UTH(PTB), based on two different AHM, and both time scales should be in close agreement. Switching of all signals (5 MHz and 1 PPS) requires manual intervention, but is simple to do. The PMS has proven very reliable, so it appears a more likely event that no software input is available. In this case the frequency steering remains unchanged, and alarms are distributed. The data required for the calculation of δf_1 are collected using well developed measurement infrastructure.

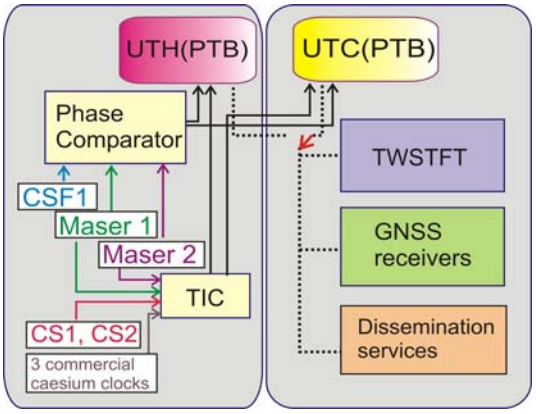


Figure 3. Implementation of duplicate realization of PTB's reference time scale UTC(PTB)

The generation of three optional inputs has proven as sufficient to provide a steering value each day and also allows cross-check among the results. The software can be re-started manually at any point in time so that during working days corrected input data could be used to produce a refined steering value if necessary.

III. Practical Implementation

Figure 3 gives a simplified view of the implementation. In order to implement Option1, a phase comparator is used to get hourly mean frequency differences between CSF1 and two AHM which provide the physical input to the generation of UTC(PTB) and UTH(PTB), respectively. The same measurement infrastructure is available for CSF2, but the interfaces to the time scale generation have not been established yet. These interfaces proved as one of the key issues to be solved. Other than most fountain clocks operated world-wide, both fountains output 5 MHz signals from a quartz oscillator (XO). Such a signal is available even if the XO is deliberately not locked to the fountain or if the regular fountain operation failed, e. g., because of a loss-of-lock of one of the lasers used for atom cooling and detection. Using data in such a condition would result in an erroneous steering command. The interface thus contains a two-level check of data quality: use or do not use.

Figure 3 shows also part of the measurement infrastructure linking the CS1 and CS2 in the clock hall and the commercial caesium clocks with the two AHM, in order to implement Options 2 and 3. In duplicate (not shown), the 1 PPS signals are routed to signal switches and compared by two time interval counters (TIC). Routinely the software accesses the data files produced daily after midnight to get the required clock differences. These are combined with previously acquired data to get the frequency prediction of each clock with respect to both AHM. These predictions are then combined with rTAI-values of each clock, and in Option 2 a weighted average is calculated based on the recent wTAI-values.

The software selects the daily steering command according to pre-selected priorities: Priority 1 has always been given to the use of CSF1 data (Option 1). A newly calculated value δf_{steer} should agree with the value calculated one day before within 2×10^{-14}. If this condition is not met, the software checks all available options. If a large step is found in all options, then this step is applied and an alarm is generated. Such a large step would point to a malfunction of the maser involved.

After a steering command was sent to the PMS, the instrument is queried, and proper reception of the command is verified. For each option and for each implementation, daily log-files and archive files are produced which allow a check of all calculations after the fact. Last but not least the software supports two events which cannot be excluded to happen. Based on auxiliary data, a change between the AHM in use to the third one available in the lab is quickly made in case that service is needed to the AHM in operation. Even an AHM with cavity auto-tuning exhibits a (linear) frequency drift, and the synthesizer which is part of the AHM electronics needs to be stepped once in a couple of months as for practical reasons the relative frequency difference between the AHM frequency and the fountains should be limited to about 10^{-13}.

IV. Results

The new realization scheme was implemented after 27 February 2010, MJD 55254, and was accompanied by a replacement of some older cables (RG 58) by a more appropriate double-shielded variant. All cable delay changes were measured and later documented as required such that the transition could be made without a step in the time scale. On the other hand, the UTC(PTB) stability improved immediately, as illustrated by a comparison to UTC(NIST) using a GPS carrier phase comparison (NRCan PPP software [6]) whose result is shown in Figure 4.

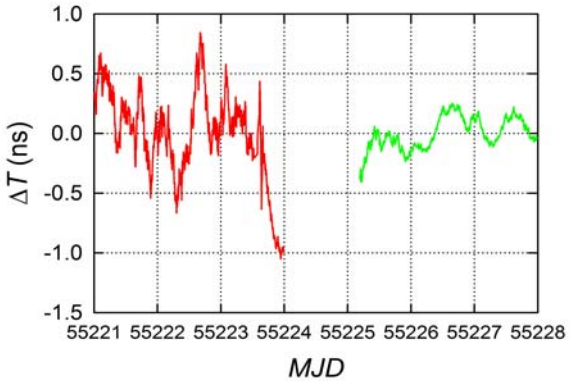


Figure 4. Phase comparison UTC(PTB) – UTC(NIST) based on a GPS carrier phases analysis using the NRCan PPP software [6] for a few days before and after the change of the UTC(PTB) realization; independent phase offsets removed in both data sets.

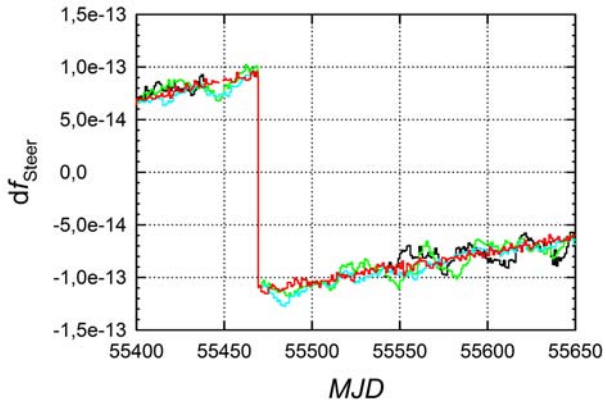


Figure 5. Steering corrections calculated since October 2010, red: Option 1, cyan: Option 2, green Option 3 (CS1), black based on CS2 with showing the interruption of its operation in fall 2010

Since then, including March 2011, the frequency steering could be based on Option 1 (CSF1 data) on about 98% of the days. Of course, during many days CSF1 was useful for less than 24 hours, but during the 13 months CSF1 was also used in 7 measurements of the TAI scale unit (see Figure 9) during which its operation has been ensured with a duty cycle

approaching 100%. The combination software (Option 2) became operational only in October 2010. Before that CS2 alone was used in Option 2 and CS1 in Option 3. In Figure 5 we show the steering values calculated since October 2010. They include one intentional maser frequency adjustment. As a matter of fact, the AHM selected as the physical source of UTC(PTB) behaved quite predictable during the last months. The alternate channel (UTH(PTB)) faced some abnormal events. In Figure 6 we compare six months of *UTC-UTC(PTB)*, in 2009 - 2010 (CS2 based) and 2010 – 2011 (hydrogen maser based) , respectively. The improvement in stability and thus predictability is appealing. During the last six months, including March 2011 UTC(PTB) deviated from UTC by less than 4 ns. In comparison with other timescales of renown quality, we see that UTC(PTB) is now quite competitive, despite of the fact, that a rather limited number of clocks has been involved in its realization.

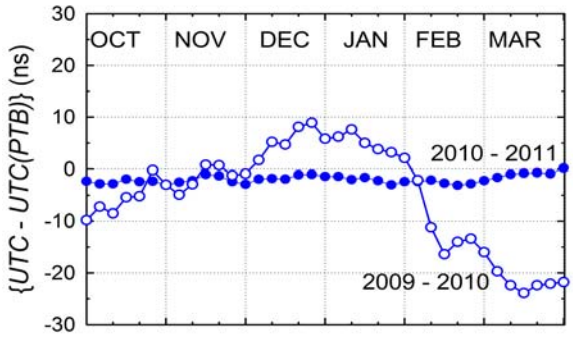


Figure 6. Time scale comparison UTC-UTC(PTB) in 2009 - 2010 and 2010 – 2011, respectively

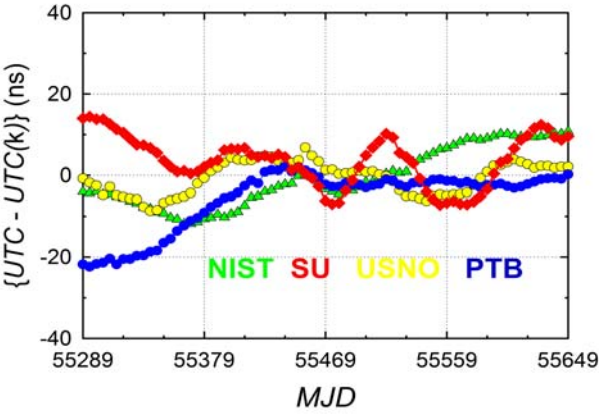


Figure 7. Time scale comparison UTC-UTC(k) during one year, including March 2011; k as indicated in the figure

As mentioned before, the realization of UTC(PTB) is made in duplicate, and in order to facilitate switching between the two signals they should be in close agreement. One option is that one realization is given the master status and the other slaved to follow it. This is not what we do. In Figure 8 we illustrate the relation between the two independently realized time scales, UTC(PTB) and UTH(PTB) (see Figure 3), during 100 days. The cause of the oscillatory behavior is under study, nevertheless the two independently realized time scales are typically within ±1 ns.

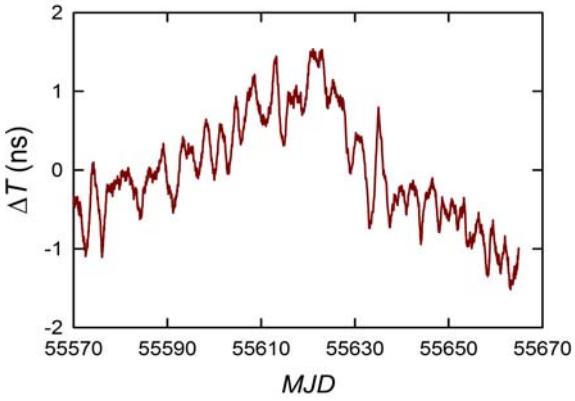


Figure 8. Comparison among the two realizations of PTB's time scale during 100 days

V. Discussion and Outlook

From the very beginning of the project, the use of fountain data as the reference for steering the AHM frequency was given the highest priority. One issue in debate was the optimum way to determine at all times the relation between CSF1 and the rate of TAI. We opted initially to use the estimated *d* from Circular T, Section 4, although *d* depends in a complicated way on the past contributions of several primary frequency standards, including CSF1 from time to time [7]. Another choice would have been to rely always on the last published value *d(CSF1)* or a mean value over recent data of that kind.

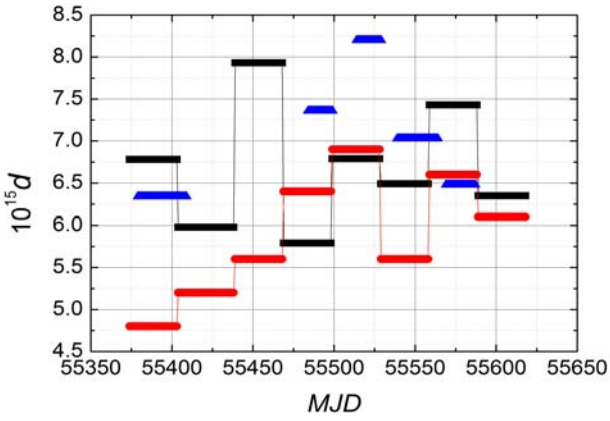


Figure 9. Three options to make the relation between CSF1 and the rate of TAI: blue: reported vales *d(CSF1)* for periods when CSF1 is used to measure the TAI scale unit; black: monthly values r(TAI-CSF1) as described in the text; red: monthly reported value *d* from Circular T, Section 4, which was actually used.

We also combined monthly average frequency values $\langle y(\text{CSF1-AHM})\rangle_M$ built from the daily values, collected as described above, with the monthly values rTAI for the AHM

involved to yield r(CSF1-TAI) for a full month. There are two subtle difficulties involved. As described before, a value $\langle y(\text{CSF1-AHM})\rangle_M$ practically never covers 100 % of a month, but the rTAI value for the AHM does. These rTAI-values, on the other hand, do not represent true frequency averages but are calculated by the BIPM from the slope of a linear fit to the seven time difference values reported monthly according to the ALGOS schedule. In Figure 9 we illustrate that the three options lead to steering values differing by a few parts in 10^{15} in some cases. In consequence, the time scale built thereon would behave differently. The result of one simulation made is shown in Figure 10 where the fictive time difference $\{UTC\text{-}UTC(PTB)\}_F$, based on the monthly r(CSF1-TAI), shows significantly larger deviations from zero as what was achieved in reality. Fortunately, the use of *d* is the simplest to implement, and as it provided good results we will continue to follow this practice for a while.

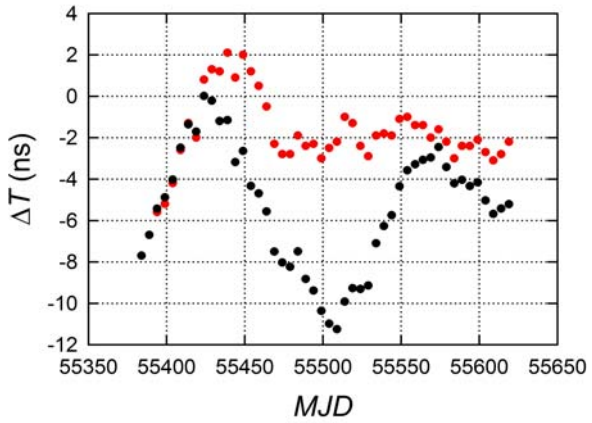


Figure 10. Deviation of UTC(PTB) from UTC: fictively calculated (black) based on the monthly values r(TAI-CSF1) (black in Figure 9) and realized in practice (red).

Simulations also covered the situation that one of the other options had to be used all times instead of Option 1. They included deviations from the practice of calculations described above, namely the use of the most recent or an average value for rTAI, and also the averaging interval for the calculation of δf_1 was questioned. We prefer to skip all details here, as in all cases a realization of UTC(PTB) based thereon would by far not have had the quality that was actually achieved using Option 1.

In view of this, the work program ahead of us is obvious. At first we implement the interfaces between the fountain CSF2 and the time-scale generation software and later use that input either as a separate option or we combine the data from the two fountains. Having achieved this, the steering of the hydrogen maser can be made twice per day. We also noted a few minor inconveniences in the operation of the software which shall be avoided in a future update.

One has to keep in mind that the UTC-UTC(PTB) differences obtained during the last 6 months are no longer large compared to the uncertainty with which they have been determined [8]. Delay changes in the time transfer equipment have to be avoided or at least have to be corrected for as otherwise they spoil the quality of the timescale. Periodic calibrations of the signal delays in the time transfer equipment are a way to support this. Such activities are, however, on our agenda anyway [9, 10] since PTB serves as the pivot laboratory for all time comparisons made in context of the realization of TAI.

ACKNOWLEDGMENT

The availability of fountain data was instrumental for the success. We appreciate the devotion of Stefan Weyers and his team at PTB to operate CSF1 in an extremely reliable way. Testing and documentation of the software was supported by Jürgen Becker and Thomas Polewka. Yang Wenke thanks China Scholarship Council for funding her one year scholarship in PTB.

REFERENCES

[1] A. Bauch, "The PTB primary clocks CS1 and CS2," Metrologia, vol. 42, pp. S43–S54, 2005.

[2] S. Weyers, U. Hübner, R. Schröder, Chr. Tamm, and A. Bauch, "Uncertainty evaluation of the atomic caesium fountain CSF1 of PTB", Metrologia, vol. 38(4), pp. 343–352, 2001, and S. Weyers, A. Bauch, R. Schröder, and Chr. Tamm, "The atomic caesium fountain CSF1 of PTB", Proceedings of the 6th Symposium on Frequency Standards and Metrology 2001, University of St Andrews, Fife, Scotland, S. 64–71, ISBN 981-02-4911-X (World Scientific).

[3] V. Gerginov, N. Nemitz, S. Weyers, R. Schröder, D. Griebsch, and R. Wynands, "Uncertainty evaluation of the caesium fountain clock PTB-CSF2", Metrologia, vol. 47, pp. 67-79, 2010.

[4] Chr. Tamm, S. Weyers, B. Lipphardt, and E. Peik, "Stray-field-induced quadrupole shift and absolute frequency of the 688-THz $^{171}Yb^+$ single-ion optical frequency standard", Phys. Rev A. vol. 80, 043403, 2009.

[5] A. Bauch, D. Piester, and E.Staliuniene, "A new realization strategy for the time scale UTC(PTB)", Proceedings of the 2004 IEEE International Frequency Control Symposium and Exposition, Montreal, 2004, pp. 518 – 523, 2004.

[6] J. Kouba and P. Héroux, "Precise Point Positioning using IGS orbit and clock products", GPS Solutions, vol. 5(2), pp. 12-28, 2001.

[7] G. Petit, "Use of primary frequency standards for estimating the duration of the scale unit of TAI", Proc. 31st Precise Time and Time Interval (PTTI) systems and Applications Meeting, Dana Point, 1999, pp. 297-304, 2000.

[8] W. Lewandowski, D. Matsakis, G. Panfilo, and P. Tavella, "The evaluation of uncertainties in [UTC-UTC(k)]", Metrologia, vol. 43, pp. 278-286, 2006.

[9] T. Feldmann, A. Bauch, D. Piester, M. Rost, E. Goldberg, S. Mitchell, and B. Fonville, "Advanced GPS based time link calibration with PTB´s new GPS calibration set-up", Proceedings of the 39th Annual Precise Time and Time Interval (PTTI) Conference, Reston, VA, October 2010, in press.

[10] K. Liang, T. Feldmann, A. Bauch, D. Piester, A. Zhang, and X. Gao, "Summary of the link calibration between NIM and PTB using a traveling GPS receiver", in this Proceedings.

Contribution to TAI frequency by a travelling Primary Frequency Standard

G. D. Rovera, M. Abgrall, Ph. Laurent
LNE-SYRTE, Observatoire de Paris - CNRS - UPMC, France
e-mail: daniele.rovera@obspm.fr

***Abstract*— The Mobile Atomic Fountain has been used to steer the TAI frequency when operating in sites without regular link to BIPM. A brief description of the method used to link this standard to a fly-wheel clock included in BIPM database and results are presented. There is no evidence of different behavior when the FOM is sitting in LNE-SYRTE or in remote laboratories.**

I. Introduction

The steering of the frequency of TAI by the Primary atomic Frequency Standards (PFS) has been introduced after that was found a systematic bias in the majority of the commercial cesium beam frequency standards of first generation [1].

With a few exceptions PFS do not run continuously to allow the necessary tests necessary to the accuracy evaluation. For this reason they are only used to steer the TAI frequency. In normal practice a PFS measures the frequency of a high stability flywheel oscillator. This oscillator participates to the computation of TAI and a monthly report of its the frequency measured by the PFS is sent to BIPM. The link uncertainty in this case is very low because usually the PFS and the flywheel clock are located in the same building, or at least in the same campus as is the case for FO1 and FO2 at Paris Observatory.

The Mobile atomic FOuntain FOM developed at SYRTE has contributed to TAI in this way for several years. In June 2008 the FOM has been moved to CNES in Toulouse for a long campaign of test of the space clock PHARAO[2]. Two Hydrogen Maser were available as flywheel at CNES but unfortunately both of them were not included in the BIPM database and therefore useless to report to BIPM. The main goal of the campaign was to check the PHARAO Performances, and a complementary task was to continue the synchronous comparison between all the PFS of LNE-SYRTE, and possibly continue to report to TAI.

First result of comparison of distant fountain have been achieved in October 2008 via a GPS carrier phase time transfer link. After a period of validation of the method a first report has been sent to BIPM with data of April 2009, followed by 3 others from the same location. In all report from CNES the carrier phase GPS data processing has been performed by CNES staff using a proprietary software[3].

In the following FOM has been moved to Max Plank Institute in Garching for a measurement of the frequency of 1S-2S Hydrogen transition [4]. During summer 2010 FOM has been moved to Observatoire de la Cote d'Azur OCA, near Grasse, to serve as a reference clock for the T2L2 experiment

Fig. 1. FOM Hardware

and from here 3 reports have been sent to BIPM. For the last two campaigns the GPS data processing has been done with the PPP software kindly licensed to SYRTE from NRCan. The method of frequency transfer described here is a demonstration of the possibility of operate a PFS outside of Metrology Institute regularly contributing to TAI.

II. The Mobile Atomic Fountain

Motivations:

- The mobile fountain FOM has been built when there was no link allowing frequency comparisons at 10^{-15} in a few days of operation.
- It was necessary to build a reference clock for the ground test of PHARAO clock at CNES integration facility.
- In almost 10 years of operation FOM has been used as primary frequency standard for several measurements of

978-1-61284-110-6/11/$26.00 ©2011 IEEE

optical/microwave frequency standards.

- The mobile fountain FOM can be used to test the performances of newly developed time/frequency links.
- Since 2006 the mobile fountain FOM has contributed 25 times to the steering of TAI.
- Since 2008 the mobile fountain FOM has contributed 8 times to the steering of TAI from 3 different sites not usually linked to UTC.

FOM can be operated in three different configurations to match the availability of a local frequency reference:

- Autonomous mode, by using a low noise BVA quartz-oscillator. This configuration delivers signals at 10 and 100 MHz with a short term instability of $1.4 \cdot 10^{-13} \tau^{-1/2}$, limited by the noise of local oscillator.
- Maser mode, by filtering the 100 MHz signal from an H maser with a low noise BVA quartz-oscillator. This configuration measure the maser frequency with a short term instability of $1.2 \cdot 10^{-13} \tau^{-1/2}$.
- Cryo mode, by using directly a 100 MHz or 1 GHz signal from a cryogenic oscillator[5]. This configuration measure the cryo frequency with a short term instability of $8 \cdot 10^{-14} \tau^{-1/2}$. In this case the stability is limited by the quantum projection noise

A. FOM measurement campaigns

- 1999 MPQ Garching: Hydrogen 1s-2s measurement [6].
- 2003 PTB Braunschweig: Fountain comparison
- 2003 MPQ Garching: Hydrogen 1s-2s measurement [7].
- 2007 Innsbruck University: Measurement of Ca+ optical clock [8].
- 2008-2010 CNES Toulouse: Test of PHARAO space clock.
- 2010 MPQ Garching: Hydrogen 1s-2s measurement[4].
- 2010 OCA Grasse: T2L2 [9].
- When in Paris: Comparison with FO1 and FO2[10], measurement of Rb, Sr[11], Hg [12], ...

B. FOM uncertainty budget

TABLE I
FOM UNCERTAINTY BUDGET

	Correction 10^{-16}	Uncertainty 10^{-16}
Quadratic Zeeman effect	-295.9	1.2
Black body radiation	165.8	0.6
Cold collisions & cavity pulling	25.1	5.0
Microwave power dependence	0	6
Ramsey & Rabi pulling	0	<0.1
Microwave recoil	0	<1.4
Second order Doppler	0	<0.1
Background gas collisions	0	<1.0
Total	-105.0	8.1
Red shift OCA	-1384.1	3.0
Red shift MPQ	-520.5	3.0
Red shift CNES	-160.1	1.0
Red shift LNE-SYRTE	-68.7	1.0

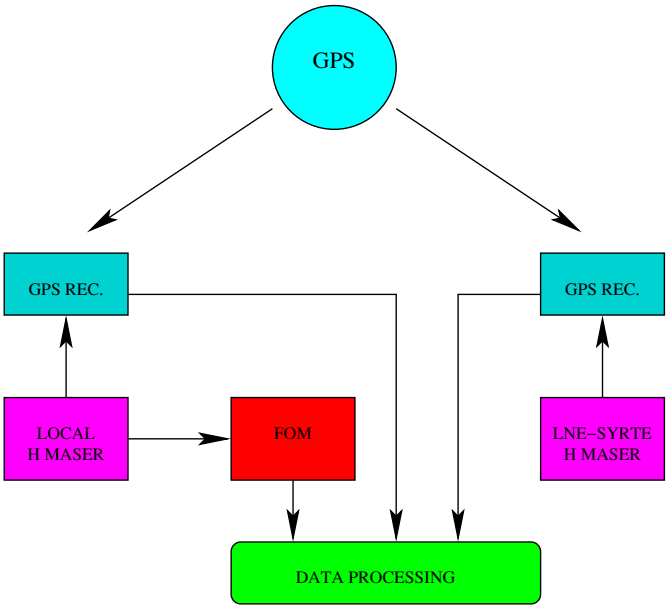


Fig. 2. Link Diagram

A typical uncertainty budget of FOM is presented in table I. The largest value of the red shift is obviously obtained at OCA (h=1268.72 m) A recent a survey[13] report the height of a reference point located nearby the fountain with a uncertainty of 1 cm. The same source also reports the local gravity g with 0.1 mGal. As the value of the red shift depends on the local potential we take a conservative uncertainty of 3^{-16} We take the same uncertainty in Garching because we have a poorer determination of height. The agreement of the frequency of the FOM with the SI at different heights is a confirmation of the good estimation of the red shift.

III. LINK OF FOM TO TAI

The basic idea is to set up a link allowing the measurement of the frequency of a flywheel oscillator that is inserted in BIPM database and contributes to the generation of TAI. Considering the experience gathered in years of fountain comparison inside LNE-SYRTE we have decided to continue on the same basis, that is the measurement of the maser frequency averaged over a period that is slightly shorter then the optimum time where the fountain and maser noises crosses.

Figure 2 shows the block diagram of the link between the FOM and one of the Maser of LNE-SYRTE used as flywheel oscillator for FO1 and FO2 fountains. In all the visited laboratories the local maser is not inserted in BIPM database and in some case not stable as the one in Paris.

When operating in Maser or Cryo mode FOM produces files with 1 line of data per cycle. The most relevant information is the central frequency of the microwave synthesizer, that is the measured uncorrected frequency of the maser.

- The 1 cycle raw files are processed line by line by applying all the estimated correction, included gravitational red

shift. The result is a 2 column file with the cycle time-stamp and the relative frequency of the local H maser.
- To avoid manipulation of huge amount of data dominated by white frequency noise the 1 cycle files are packed in files with one point ever 0.1 day.
- GPS rinex files are processed with the NRCan PPP software and used to calculate the time difference between the local H maser and the H maser 140 0890 located at LNE-SYRTE.
- The maser's frequency difference, averaged over period of 0.1 day, is obtained by smoothing and differentiation of time differences.
- The frequency of H maser 140 0890 is obtained by difference of local and GPS frequency files.
- Usual techniques are applied to estimate the frequency and drift of the LNE-SYRTE H maser 140 0890 in a standard BIPM interval, as well the Allan standard deviation and the link uncertainty.

TAI Contributions

List of FOM contributions from different sites

- 54919-54944 CNES GPS carrier phase CNES software
- 55044-55074 CNES GPS carrier phase CNES software
- 55074-55104 CNES GPS carrier phase CNES software
- 55104-55119 CNES GPS carrier phase CNES software
- 55344-55359 MPQ GPS PPP NRCan software
- 55364-55399 OCA GPS PPP NRCan software
- 55404-55434 OCA GPS PPP NRCan software
- 55439-55469 OCA GPS PPP NRCan software

Figure III shows the frequency difference between FOM and TAI. For comparison on the same graph the red line shows the difference between SI, that is the best estimation of the realization of the second, and TAI. From this plot it is evident that the frequency of FOM is in good agreement with SI not only when it is sitting in Paris Observatory but also when it is travelling around Europe.

IV. Conclusion

This is the first time that a Primary Frequency Standard contributes to steering of TAI from different sites with different gravitational red shift.

The perfect agreement of the frequency of the standard with SI from locations differing in height more than a thousand meter is confirmation of the correctness of the evaluation of gravitational red shift.

Acknowledgments

Thanks to colleagues of LNE-SYRTE for supporting this activity. CNES has supported FOM development, its utilization in Toulouse and Grasse and has provided GPS data analysis for the Toulouse campaign. Thanks to MPQ and OCA for hosting the FOM. Thanks to NRCan for the PPP software, with a special mention to François Lahaye for the support. One of the authors (G. D. Rovera) will acknowledge INRiM for the intensive training in GPS data analysis.

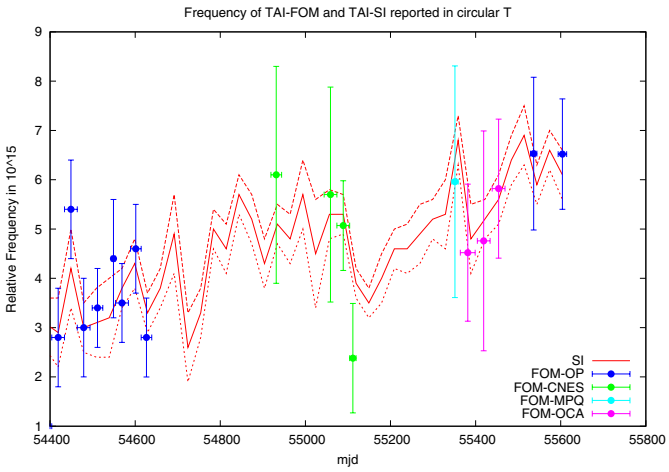


Fig. 3. Frequency of FOM - TAI from BIPM Circular T. The red line shows SI-TAI, with dotted lines as error bars

References

[1] A. De Marchi, G. D. Rovera, and A. Premoli, "Pulling by neighbouring transitions and its effects on the performance of Caesium beam frequency standards," *Metrologia*, vol. 20, pp. 37–47, 1984.

[2] P. Laurent et al., "Design of the cold atom PHARAO space clock and initial test results," *Appl. Phys. B*, vol. 84, no. 4, pp. 683–690, Sept. 2006.

[3] J. Delporte, F. Mercier, D. Laurichesse, O. Galy, GPS carrier phase time transfer using single-difference integer ambiguities, International Journal of Navigation and Observation, Hindawi Publishing Corporation, Volume 2008, Article ID 273785

[4] C. G. Parthey et al., "Improved measurement of the H ydrogen 1S-2S transition frequency," 2004, submitted.

[5] D. Chambon, S. Bize, M. Lours, F. Narbonneau, H. Marion, A. Clairon, G. Santarelli, A. Luiten, and M. Tobar, "Design and realization of a flywheel oscillator for advanced time and frequency metrology," *Rev. Sci. Instrum.*, vol. 76, p. 094704, 2005.

[6] M. Niering et al. "Measurement of the hydrogen $1s$- $2s$ transition frequency by phase coherent comparison with a microwave cesium fountain clock," *Phys. Rev. Lett.*, vol. 84, no. 24, pp. 5496–5499, Jun 2000.

[7] M. Fischer, N. Kolachevsky, M. Zimmermann, R. Holzwarth, T. Udem, T. W. Hänsch, M. Abgrall, J. Grünert, I. Maksimovic, S. Bize, H. Marion, F. P. D. Santos, P. Lemonde, G. Santarelli, P. Laurent, A. Clairon, C. Salomon, M. Haas, U. D. Jentschura, and C. H. Keitel, "New limits on the drift of fundamental constants from laboratory measurements," *Phys. Rev. Lett.*, vol. 92, no. 23, p. 230802, Jun 2004.

[8] M. Chwalla et al., "Absolute frequency measurement of the $^{40}\mathrm{Ca}^{+}$ $4\mathrm{s}^2\mathrm{S}_{1/2} - 3\mathrm{d}^2\mathrm{D}_{5/2}$ clock transition," *Phys. Rev. Lett.*, vol. 102, p. 023002, 2009. [Online]. Available: http://link.aps.org/abstract/PRL/v102/e023002

[9] E. Samain et al., "Time transfer by laser link - T2L2: Current status and future experiments (invited)," in *IEEE International Frequency Control Symposium, 2011 Joint with the 24nd European Frequency and Time forum.*, 2011.

[10] J. Guena, P. Rosenbusch, P. Laurent, M. Abgrall, D. Rovera, G. Santarelli, M. Tobar, S. Bize, and A. Clairon, "Demonstration of a dual alkali Rb/Cs fountain clock," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 57, no. 3, pp. 647–653, 2010.

[11] X. Baillard et al. "An optical lattice clock with spin-polarized 87sr atoms," *Eur. Phys. J. D*, vol. 48, p. 11, 2008.

[12] M. Petersen, et al. "Doppler-free spectroscopy of the $^1S_0 - ^3P_0$ optical clock transition in laser-cooled fermionic isotopes of neutral mercury," *Phys. Rev. Lett.*, vol. 101, no. 18, p. 183004, 2008.

[13] http://geodesie.ign.fr/fiches

Comparing Room Temperature and Cryogenic Cesium Fountains

T. P. Heavner, T. E. Parker, J. H. Shirley, L. Donley, S.R. Jefferts
NIST – Time and Frequency Division
325 Broadway
Boulder, CO 80305 USA

F. Levi, D. Calonico, C. Calosso
INRIM – Str. delle Cacce 91
10135 Torino Italy

G. Costanzo, B. Mongino
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino Italy

***Abstract*—We have compared the frequency of a room-temperature cesium-fountain primary frequency standard with that of a cryogenic (~80K) cesium fountain. This comparison yields a measurement of the blackbody frequency shift of the room-temperature fountain.**

I. INTRODUCTION

NIST-F1 is a well characterized cesium-fountain primary fractional frequency standard with frequency inaccuracy at the $\delta f/f \sim 4\times10^{-16}$ level. This inaccuracy is dominated by the uncertainty in the blackbody correction of about $\delta f/f \sim 3\times10^{-16}$. NIST-F2 and IT-CsF2 are cryogenic (80K) cesium-fountain frequency standards coming into operation at the National Institute of Standards and Technology (NIST) in Boulder, CO and at Istituto Nazionale di Ricerca Metrologica (INRIM) in Torino, Italy. Frequency comparisons among these fountains yield a measurement of the blackbody shift

II. ROOM TEMPERATURE FOUNTAINS

A. NIST-F1

NIST-F1 is the US primary frequency standard and has been previously described in great detail [1],[2],[3]. Thus we will give only a cursory overview of the salient points regarding its performance and uncertainties.

B. Error Budget

An abbreviated error budget for NIST-F1 is shown in Table 1. It can be seen that the overall uncertainty of the standard is dominated by the $\delta f/f \sim 2.8\times10^{-16}$ fractional uncertainty in the blackbody shift. This uncertainty corresponds to a 1 K uncertainty in the effective radiation temperature of the standard, and is the uncertainty that the cryogenic fountains are designed to effectively eliminate. A comparison of the room temperature fountains with the cryogenic fountains combines the systematic errors in the fountains in ways that require individual treatment of these systematic uncertainties: an obvious example is the gravitational shift that largely cancels for both clocks with the uncertainty in the shift dropping from $\delta f/f \sim 3\times10^{-17}$ to $\delta f/f \sim 1\times10^{-18}$.

III. CRYOGENIC FOUNTAINS

A. NIST-F2 and INRIM CsF-2

The NIST and INRIM cryogenic fountains are similar in that the design of the Physics package is essentially identical. Details of the optical systems and control electronics are however idiosyncratic to each particular fountain. The physics package has been described previously and only a short review is given here [4],[5].

The fountains operate on a pure optical-molasses atom-loading scheme (no MOT), with the beam geometry being (1,1,1) (crystallographic notation that defines the beam geometry). The atoms are launched upwards at about 4.5 m/s and post-cooled in the moving frame to temperatures around 500 nK. The

atoms fly upward through the room-temperature vacuum system before entering the magnetically shielded cryogenic region. In the cryogenic region the atoms are state-selected and Ramsey interrogation takes place. The microwave structure is essentially identical to that of NIST-F1 [1],[2],[3] with the exception that the microwave cavities are tuned to resonance at 80 K instead of 318 K (NIST-F1) or 340 K (IT CsF-1). Detection takes place at room temperature after the atoms have left the magnetically shielded cryogenic region. The detection region is very similar to that used in NIST-F1, and was specifically designed to minimize vignetting, with the modeled detection efficiency being uniform over the entire cloud and vignetting being less than 10 %.

TABLE I. NIST-F1 ERROR BUDGET – BIASES AND UNCERTAINTIES ARE GIVEN IN UNITS OF $\delta F/F = 10^{-15}$. THIS IS CURRENT AS OF MARCH 2011.

Physical Effect	Bias	Type B Uncertainty
Gravitational Red shift	+179.95	0.03
Second-Order Zeeman	+180.25	0.01
Blackbody	-22.98	0.28
Microwave Effects	-0.026	0.12
Spin Exchange (density =8)	0.0 (-0.56)	0.06 (0.16)
AC Zeeman (heaters)	0.05	0.05
Cavity Pulling	0.02	0.02
Rabi Pulling	10^{-4}	10^{-4}
Ramsey Pulling	10^{-4}	10^{-4}
Majorana Transitions	0.02	0.02
Fluorescence Light Shift	10^{-5}	10^{-5}
Second-Order Doppler	0.02	0.02
DC Stark Effect	0.02	0.02
Background Gas Collisions	10^{-3}	10^{-3}
Bloch-Siegert	10^{-4}	10^{-4}
RF Spectral purity	$3x10^{-3}$	$3x10^{-3}$
Integrator offset	0	0.01
Total Type B Standard Uncertainty (including Spin Exchange)		0.30 (0.34)

B. Error Budget – NIST-F2 & IT CSF-2

An abbreviated error budget for NIST-F2 is shown in Table 2. This is a preliminary error budget and several systematic shifts, notably systematic frequency shifts associated with microwave effects (eg. microwave leakage, distributed cavity phase, microwave spurious signals etc) are not yet completely evaluated. The entries in the error budgets generally reflect our most pessimistic measurements and estimates and are usually limited by statistical uncertainties in those measurements. At this time the error budget for INRIM IT-CsF-2 is similar to that for NIST-F2.

TABLE II. NIST-F2 ERROR BUDGET – BIASES AND UNCERTAINTIES ARE GIVEN IN UNITS OF $\delta F/F = 10^{-15}$. THIS IS CURRENT AS OF MARCH 2011.

Physical Effect	Bias	Type B Uncertainty
Gravitational Red shift	+179.15	0.03
Second-Order Zeeman	+287.178	0.03
Blackbody	-0.096	0.005
Microwave Effects	-0.0025	0.27
Spin Exchange (density =10)	0.0 (0.07)	0.01 (0 .24)
Cavity Pulling	0.02	0.02
Rabi Pulling	10^{-4}	10^{-4}
Ramsey Pulling	10^{-4}	10^{-4}
Majorana Transitions	0.02	0.02
Fluorescence Light Shift	10^{-5}	10^{-5}
Second-Order Doppler	0.00	0.01
DC Stark Effect	0.02	0.02
Background Gas Collisions	10^{-3}	10^{-3}
Bloch-Siegert	10^{-4}	10^{-4}
RF Spectral purity	$3x10^{-3}$	$3x10^{-3}$
Integrator offset	0	0.01
Total Type B Standard Uncertainty (Including Spin Exchange)		0.28 0.36

IV. BLACKBODY SHIFT MEASUREMENT

Three measurement campaigns using NIST-F1 and NIST-F2 have been completed. These campaigns (in Sept 2010, Dec 2010 and March 2011) after appropriate data processing yield a measurement of the blackbody shift in NIST-F1. In the course of a measurement, NIST-F1 and NIST-F2 are run concurrently. Each fountain is corrected for systematic frequency shifts associated with the Zeeman effect, spin-exchange collisions, and differential gravitational effects. The frequency difference between the two fountains is then compared to the frequency difference calculated for the differing blackbody shifts in NIST-F1 and NIST-

F2. That comparison yields a measurement of the blackbody shift in NIST-F1.

Preliminary results from comparison of the NIST fountains suggests that the difference between the blackbody shift as predicted by theoretical calculations [6] and the measured blackbody shift is well within the experimental uncertainties ($\delta f/f$ ~5-7*10^{-16}) of the comparison.

Further measurements are underway and full results will be published in a forthcoming paper.

ACKNOWLEDGMENT

The authors are indebted to many people who contributed to this effort over the past several years. In particular, Doug Gallagher of the NIST Instrument Shop for his fine workmanship in the construction of the cryogenic fountains, and Mike Lombardi and David Smith for their always useful and pertinent suggestions on manuscripts.

REFERENCES

[1] S.R.Jefferts, et al., "Accuracy Evaluation of NIST-F1," Metrologia, vol. 39, pp. 321–336, 2002.

[2] T.P. Heavner, S.R. Jefferts, E.A.Donley, J. Shirley, T.E. Parker, "NIST-F1:recent improvments and accuracy evaluations," Metrologia, vol. 42, pp. 411–422, 2005.

[3] T.E. Parker, S.R. Jefferts, T.P. Heavner, E.A. Donley, "Operation of the NIST-F1 cesium fountain primary frequency standard with a maser ensemble, including the impact of frequency transfer noise" Metrologia, vol. 42, pp. 423–430, 2005.

[4] T.P. Heavner, T.E. Parker, J.H. Shirley and S.R. Jefferts "NIST-F1 and F2," Proc. 2008 Symp. on Freq. Stds. Metrology, pp.299–307, 2008.

[5] F. Levi et al, "The Cryogenic Fountain ITCs-F2," Proc. 2009 EFTF, pp 769-773, 2009.

[6] K. Beloy, U.I. Safronova, A. Derevianko, "High-Accuracy Calculation of the Blackbody Radiation Shift in the 133 Cs Primary Frequency Standard," Phys. Rev. Lett. **97** 040801 (2006).

Uncertainty evaluation of the continuous cesium fountain frequency standard FOCS-2

Gianni Di Domenico*, Laurent Devenoges*, Alain Joyet*, André Stefanov† and Pierre Thomann*
*Laboratoire Temps-Fréquence, Université de Neuchâtel, 2000 Neuchâtel, Switzerland
Email: gianni.didomenico@unine.ch
†Federal Office of Metrology, 3003 Bern-Wabern, Switzerland

***Abstract*— The continuous atomic fountain primary frequency standard FOCS-2 is quite unique since all the atomic fountain clocks presently contributing to the international atomic time (TAI) are operating in pulsed mode [1]. Our alternative approach offers the possibility to operate with a high atomic flux without the collisional shift limiting the accuracy or the Dick effect limiting the stability when a quartz is used as local oscillator [2], [3]. Moreover, it contributes to the metrological diversity since the relative importance of the error budget contributors is different in the two types of fountains, notably for microwave cavity and density related effects. In this contribution we present the status of the metrological evaluation of the continuous fountain FOCS-2. Thanks to an innovative atomic beam source followed by an atomic state preparation stage, we reach a signal-to-noise ratio of $600\sqrt{\text{Hz}}$ and thus a relative frequency instability of $6 \times 10^{-14}\,\tau^{-1/2}$. A detailed investigation of the frequency shifts which are specific to the continuous operation (light-shift from the source and cavity end-to-end phase shift) shows that the continuous fountain is compatible with an inaccuracy at the 10^{-15} level or below.**

I. Introduction

In this work, we report on the evaluation of the frequency shifts in our continuous atomic fountain clock FOCS-2. We start in Section II by describing the atomic resonator with an accent on the elements which are unique to the continuous mode of operation: the light-trap and the microwave cavity with two distinct interaction regions. In Section III we describe the operation of the fountain, from the acquisition of Ramsey fringes to frequency stability measurements and state preparation. In Section IV, we discuss the evaluation of the frequency shifts which are specific to the continuous fountain and therefore required the development of new experimental methods. The light-shift produced by the source of the atomic beam is reduced to a negligible level with a rotating light-trap as described in Section IV-A. The microwave cavity is substantially different in a continuous fountain since the upward and downward interactions zones must be separated. As a consequence, a cavity reversal technique was developed for the evaluation of the end-to-end phase shift and the results are presented in section IV-B. Evaluation of the second-order Zeeman shift requires a precise knowledge of the magnetic field in the atomic free evolution zone. The methods developed to measure this magnetic field are presented in Section IV-C. Finally, we present the status of the inaccuracy budget in Section V and conclude in Section VI.

II. Description of the continuous atomic fountain FOCS-2

A. Atomic resonator

A scheme of the continuous atomic fountain FOCS-2 is shown in Fig. 1. The atomic beam is produced with a two-dimensional magneto-optical trap (2D-MOT) [4]. Then, the atoms are captured, cooled, and launched at a speed of 4 m/s with a moving molasses [5]. The atomic beam temperature at the exit of the moving molasses is approximately 70 μK. Before entering the microwave cavity, the atomic beam is further collimated by transverse laser cooling. Moreover, the atoms are pumped into $|F = 3, m = 0\rangle$ with a state preparation scheme combining both optical pumping and laser cooling [6]. After these two steps, the transverse temperature is reduced to ≈ 3 μK.

Then, the atomic beam passes through the light-trap which stops the light coming from the source. A photodiode has been installed which measures the light intensity after the light-trap and is used for the evaluation of the light-shift and long term monitoring, see Sec. IV-A.

During the first passage into the microwave cavity, we apply a $\pi/2$-pulse (duration of 10 ms) and thereby produce a superposition state which evolves freely for 0.52 s before the second $\pi/2$-pulse is applied during the second passage. Finally, the transition probability between $|F = 3, m = 0\rangle$ and $|F = 4, m = 0\rangle$ is measured by fluorescence detection of the atoms in $F = 4$.

B. Light-trap

The atomic beam source described in the previous section operates in a continuous mode, which means that the lasers producing the cold atomic beam are always switched on, and as a consequence their light may produce a shift of the clock transition frequency. Therefore, a rotating light-trap has been installed to prevent the source light to reach the free evolution zone situated above. It is composed of 18 glass blades oriented at 45^o as shown in Fig. 2. The rotation speed is adjusted such that the horizontal velocity v_t of the blades matches the longitudinal velocity v_l of the atomic beam, see Fig. 2(a). When this condition is satisfied, the atoms can pass with a minimal attenuation $< 15\%$. On the contrary, the photons are either absorbed or deviated into a beam block after two reflexions on the blades as shown in Fig. 2(b). The light

978-1-61284-110-6/11/$26.00 ©2011 IEEE

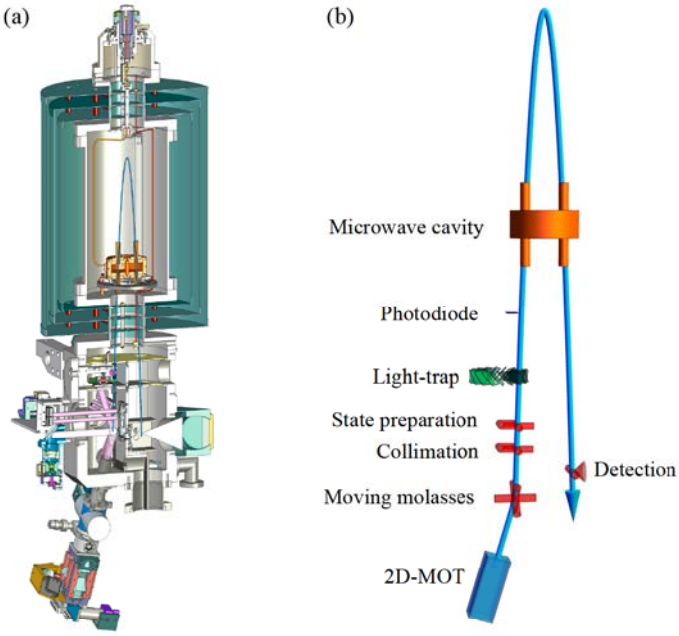


Fig. 1. Scheme of the continuous atomic fountain FOCS-2. (a) Atomic resonator. (b) Scheme of principle. An intense atomic beam of pre-cooled cesium atoms is produced in the two-dimensional magneto-optical trap (2D-MOT). The atoms are then captured by the 3D moving molasses which further cools and launches the atoms at a speed of 4 m/s. Then the atomic beam is collimated with Sisyphus cooling in the transverse directions and, before entering the microwave cavity, the atoms are pumped in $|F = 3, m = 0\rangle$ by state preparation. Finally, after the second passage in the microwave cavity, the transition probability is measured by fluorescence detection of the atoms in $F = 4$. The free evolution time between the two $\pi/2$-pulses is $T = 0.52$ s. The light-trap prevents the stray light from the atomic beam source from reaching the free evolution zone situated above.

attenuation factor was measured to be 10^4. The trap is driven by a home-made electrostatic motor which is compatible with ultra-high vacuum and entirely non-magnetic [7].

C. Microwave cavity

We use a coaxial cylindrical cavity with two interaction regions as shown in Fig. 3. The cavity is mounted on a rotation stage which allows us to exchange the interaction zones for the evaluation of the end-to-end phase shift. There are two feeding ports which are placed at 90 degrees with respect to the plane containing the interaction zones and we excite the mode TE_{021}. The quality factor is 15000 loaded and 22000 unloaded. The coupling factors are approximately 0.2.

III. Operation of the Continuous Atomic Fountain FOCS-2

A. Ramsey fringes

In our continuous fountain, Ramsey fringes are measured by continuously recording the transition probability while scanning the microwave frequency. Figure 4 shows a measurement of the Ramsey fringes observed on the clock transition, acquired in a single scan, at a speed of 2 Hz per second. The

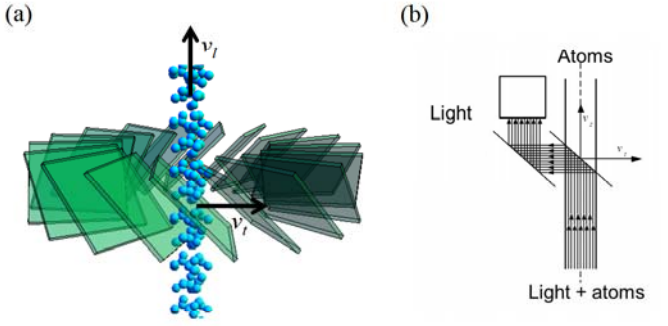


Fig. 2. Scheme of the light-trap. (a) The atomic beam can pass through the trap when the blades horizontal velocity v_t matches the atomic beam vertical velocity v_l. (b) The light is stopped after two reflexions on the blades.

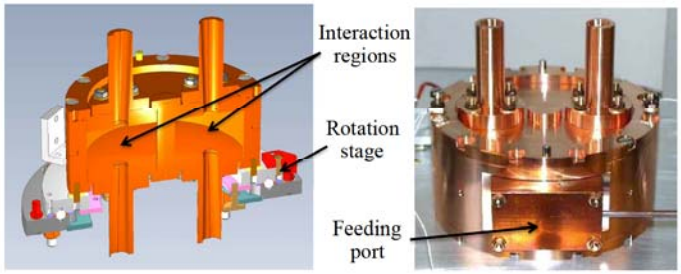


Fig. 3. The microwave cavity has two distinct interaction regions, and two feeding ports. It is mounted on a rotation stage which enables to exchange the interaction zones for the evaluation of the end-to-end phase shift.

contrast is $> 98\%$ and the signal-to-noise ratio is approximately $600\sqrt{\text{Hz}}$ [1].

B. Frequency stability measurement

We compare the frequency of the fountain FOCS-2 to an active hydrogen maser using the scheme presented in Fig. 5. A commercial microwave synthesizer (from SpectraDynamics Inc.) referenced to the hydrogen maser is used to interrogate the atomic transition. We use square wave phase modulation

[1] Here, the signal is the peak-valley amplitude of the central fringe, and the noise is measured on the top of the central fringe.

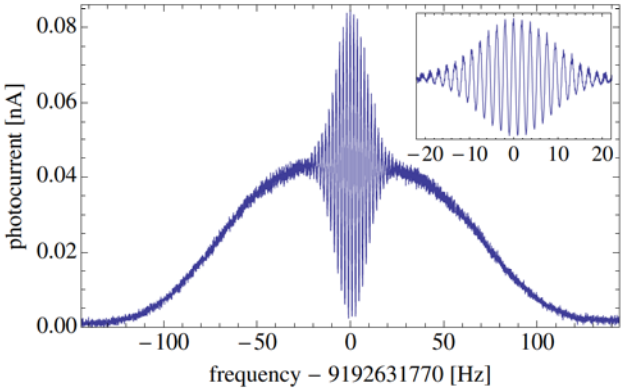


Fig. 4. Measurement of the Ramsey fringes observed on the clock transition, acquired in a single scan, at a speed of 2 Hz per second.

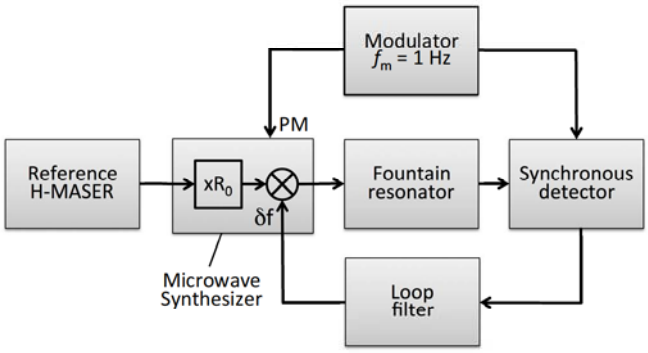


Fig. 5. Scheme of the frequency control loop. A commercial synthesizer, referenced to an active hydrogen maser, generates the microwave which is used to interrogate the atoms. By modulating the phase of the microwave at 1 Hz and demodulating the resulting transition probability, one obtains the error signal which is integrated to produce the frequency correction δf. See section III-B for details.

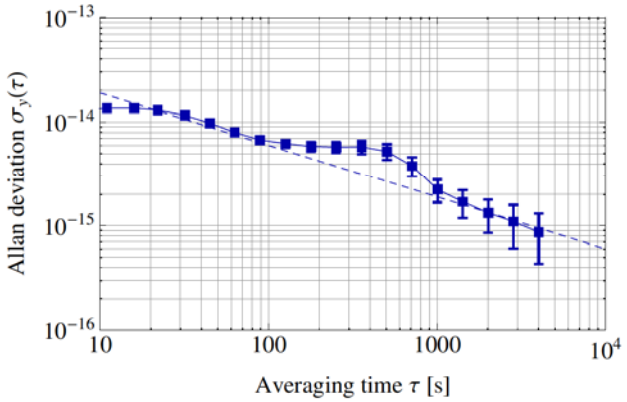


Fig. 6. Allan deviation of FOCS-2 compared to an active hydrogen maser. The dashed line indicates $6 \times 10^{-14}\,\tau^{-1/2}$. The bump at 500 s is due to temperature variations (± 1 K) in the laboratory.

at 1 Hz of the microwave, and subsequent square wave synchronous detection of the transition probability, to generate the error signal which is then integrated by the loop filter to control the synthesizer frequency. The resulting frequency correction δf is a measure of the frequency deviation of the fountain with respect to the maser. Note that the synchronous detector is digital, the servo loop filter as well, and the synthesizer frequency is numerically controlled.

With this scheme, we obtain the Allan deviation shown in Fig. 6 which follows a $6 \times 10^{-14}\,\tau^{-1/2}$ behavior. The bump at 500 seconds is due to temperature fluctuations of ± 1 K in the laboratory and should disappear with a better stabilization of temperature. Let's note that the first 2 points are not relevant since they are out of the servo loop bandwidth.

C. State preparation for an improved frequency stability

State preparation is a key point to reach the high stability shown in Fig. 6. We use two-laser optical pumping to bring all the atoms in $|F = 3, m = 0\rangle$. A first laser excites the $F = 4$ ground state in order to pump the atoms toward $F = 3$ while a second π-polarized laser excites the $F = 3 \rightarrow F' = 3$ transition of the D2 line in order to produce Zeeman pumping toward $m = 0$. To avoid trap states, the first laser is implemented in a 2D optical lattice geometry, thereby creating polarization gradients. This configuration has the advantage of simultaneously producing Sisyphus cooling when the optical lattice laser is tuned on the blue side of the $F = 4 \rightarrow F' = 4$ transition, which is important to remove the heat produced by optical pumping [6].

Thanks to state preparation, we were able to bring approximately 60% of the atoms in $|F = 3, m = 0\rangle$ which produces a notable increase of the useful flux. As a consequence, we observe an increase of both the signal-to-noise ratio and the frequency stability by a factor > 4 [8]. The resulting Allan deviation is shown in Fig. 6.

IV. Evaluation of the Frequency Shifts

In this section, we discuss the evaluation of three frequency shifts which are specific to the continuous fountain and therefore required the development of new experimental methods.

A. Reduction of the light-shift from the source

As mentioned in section II, the lasers producing the atomic beam are always switched on in our continuous fountain, and therefore they produce a light-shift of the clock transition which has to be reduced. To this end, we installed a rotating light-trap to prevent the source light from reaching the interrogation zone. A photodiode installed a few centimeters above the light-trap (see Fig.1(b)) allows us to measure the amount of light coming from the source along the atomic beam trajectory in the direction of the microwave cavity. By measuring the detected light before and after the installation of the light-trap, we obtained an attenuation factor of 10^4.

Before installing the light-trap, we measured the light-shift by comparing FOCS-2 frequency to UTC via a GPS link. In order to identify the light-shift contribution, we measured both the fountain frequency and the amount of light along the atomic beam trajectory (with the photodiode mentioned above) for different values of the source fluorescence light. In order to change the source fluorescence light, we varied the atomic flux by changing the 2D-MOT laser power. The results are presented in Fig. 7. The linear dependence indicates that the source fluorescence light is the main contributor. Moreover, the value of the light-shift at nominal flux and without light-trap is

$$\Delta\nu_{\rm LS}/\nu_{\rm Cs} = -1.6 \times 10^{-12} \text{ without light-trap} \qquad (1)$$

After installation of the light-trap, we measured the amount of fluorescence light collected by the photodiode and obtained an attenuation factor of 10^4. Since the light-shift is proportional to the light intensity, we can conclude that the light-shift with light-trap is reduced down to

$$\Delta\nu_{\rm LS}/\nu_{\rm Cs} - -1.6 \times 10^{-16} \text{ with light-trap} \qquad (2)$$

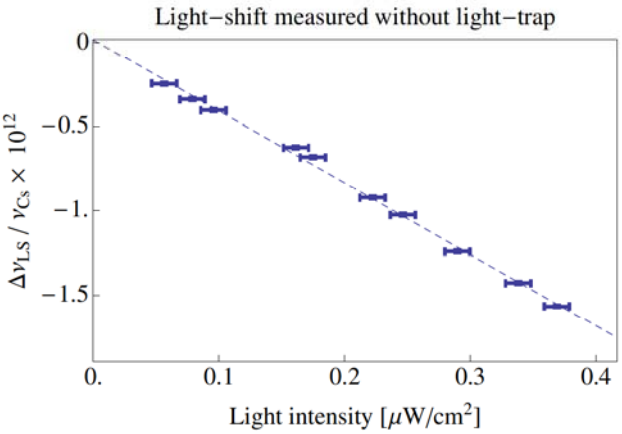


Fig. 7. Light-shift measured in FOCS-2 before the installation of the light-trap. The horizontal axis is the light collected by the photodiode shown in Fig.1(b). It is a measurement of the stray light emitted by the atom source toward the microwave cavity. This stray light is attenuated by a factor 10^4 with the light-trap. See section IV-A for details.

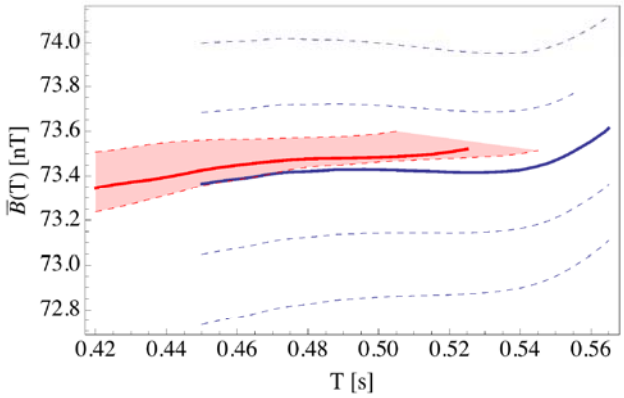


Fig. 8. Time-averaged magnetic field $\overline{B}(T)$ as a function of the transit time T. The blue solid line was obtained from the phase of the Fourier transform of the Ramsey fringes measured on $m_F = -1$. Due to the 2π ambiguity of the phase, this function is not uniquely determined. The blue dashed curves represent possible realizations which differ by a multiple of $1/(K_z T)$. The red band has been obtained from the time-resolved Zeeman spectroscopy measurement of the magnetic field profile. It allows us to identify one of the curves from the Fourier analysis. See Section IV-C for details.

B. Microwave cavity end-to-end phase shift

As mentioned in section II, the microwave cavity has two distinct interaction zones, and thus a cavity end-to-end phase shift $\Delta\phi$ results in a frequency bias $\Delta\nu_\phi = -\Delta\phi/(2\pi T)$ where T is the effective free evolution time. In order to measure this shift, the microwave cavity is mounted on a rotation stage which allows us to rotate it by 180^o around the vertical axis, thereby exchanging the two interaction zones, and thus changing the sign of $\Delta\phi$. This high quality rotation stage enables to reproduce the absolute position of the microwave cavity with a precision better than 0.01 mm. In order to measure $\Delta\nu_\phi$, we successively compare the frequency of FOCS-2 in positions A (original position) and B (rotated by 180^o) with the hydrogen maser during a short period of time such that the maser frequency drift is negligible. The cavity end-to-end phase shift is then given by $\Delta\nu_\phi/\nu_{Cs} = y(\text{position A} - \text{position B})/2$.

We performed a frequency measurement of eight hours, alternating between positions A and B every 30 minutes, and after averaging the frequency differences as explained above, we obtain the cavity end-to-end phase shift

$$\Delta\nu_\phi/\nu_{Cs} = (1.7 \pm 0.9) \times 10^{-15} \qquad (3)$$

C. Second-order Zeeman shift

Evaluation of the second order Zeeman shift $\Delta\nu_Z = K_0\overline{B^2}$ where $K_0 = 42.745$ GHz T^{-2} and $\overline{B^2} = (1/T)\int B^2(t)dt$ requires a precise knowledge of the magnetic field in the free evolution zone. The methods developed in pulsed fountains to measure the magnetic field in the resonator are based on throwing balls of atoms at different altitudes. This is not applicable to our continuous fountain since the atomic trajectory is not vertical and therefore the launching velocity is limited by geometrical constraints. Moreover, the atomic beam longitudinal temperature is higher in our continuous fountain (75 μK) than in pulsed fountains (1 μK) and as a consequence the distribution of apogees is wider. The effect of this large distribution of transit time is to modify significantly the Ramsey pattern, reducing the number of visible fringes but also increasing its sensitivity to magnetic field inhomogeneities. As a consequence, identification of the central fringe on a Zeeman sensitive Ramsey pattern showed to be difficult in FOCS-2, and thus we developed novel methods to determine the magnetic field.

The first method makes use of the Fourier transform of Ramsey fringes measured on Zeeman sensitive components ($m \neq 0$). It is based on the observation that Ramsey fringes observed in an atomic fountain are formed by the superposition of the individual atomic signals. Due to the atomic beam residual temperature, the atoms have slightly different trajectories and thus are subject to a different average magnetic field. As a consequence, both the velocity distribution and the magnetic field profile are imprinted on the Ramsey patterns observed on Zeeman sensitive transitions. In Ref. [9] we demonstrated that one can retrieve the time-averaged magnetic field $\overline{B}(T)$ associated with different trajectories (of transit times T) from the phase of the Fourier transform. By applying the procedure described in [9] to FOCS-2, we obtain the time-averaged magnetic field shown in Fig. 8. As discussed in Ref. [9] there is a 2π ambiguity in the determination of the phase which generates an ambiguity in $\overline{B}(T)$ of $\pm n/(K_z T)$ where $K_z = 7\,\text{Hz/nT}$ and n is any integer number (see the dashed lines in Fig. 8). In principle, it would be possible to identify the correct curve ($n = 0$) by extending the domain of T values, but this is not possible in FOCS-2 since the transit time values are limited by geometrical constraints. As a consequence another method is necessary in order to solve this ambiguity.

The second method consists in measuring the magnetic field profile $B(z)$ in the free evolution zone by using time-resolved Zeeman spectroscopy. The basic principle is to measure the Zeeman frequency f_z by exciting $\Delta F = 0, \Delta m = \pm 1$

TABLE I

PRESENT STATUS OF THE EVALUATION OF FOCS-2.

Physical effect	Bias	Uncertainty	Unit
Light-shift from the source	-0.16	0.08	10^{-15}
Cavity end-to-end phase shift	1.7	0.9	10^{-15}
Second-order Zeeman	25.0	0.2	10^{-15}
Blackbody radiation	-18.9	0.1	10^{-15}
Gravitational shift	54.66	0.13	10^{-15}
Second-order Doppler	-0.012	0.001	10^{-15}
Rabi pulling	< 0.005	0.005	10^{-15}
Cavity pulling	-	-	-
Collisional shift	-	-	-
Distributed cavity phase shift	-	-	-
Power shifts	-	-	-
Electronics shifts	-	-	-
Background gas	-	-	-
Majorana	-	-	-

transitions and then deduce the magnetic field from $B = f_z/K_z$. Finally, from the magnetic field profile $B(z)$ and with the knowledge of the atomic beam trajectory, one can calculate the time-averaged magnetic field $\overline{B}(T)$. The result is presented in Fig. 8 together with the measurement obtained from the Fourier analysis of Ramsey fringes. In the Zeeman spectroscopy measurement, the main contribution to the error bar comes from the uncertainty on the atomic beam trajectory. Therefore, starting with the same experimental data, we repeated the same analysis but with trajectories differing by $\pm 1\sigma$ from the nominal trajectory, and we obtained the $\pm 1\sigma$ error band shown in Fig. 8.

With this knowledge of the magnetic field, combined with the distribution of transit times, we obtain the second-order Zeeman shift

$$\Delta\nu_Z/\nu_{\mathrm{Cs}} = (25.0 \pm 0.2) \cdot 10^{-15} \tag{4}$$

V. PRESENT STATUS OF THE INACCURACY BUDGET

The present status of the evaluation of the frequency shifts in our continuous fountain FOCS-2 is shown in table I. In section IV, we discussed in detail the frequency shifts which are specific to the continuous fountain and therefore required the development of new experimental methods: the light-shift from the source, the cavity end-to-end phase shift and the second-order Zeeman shift. The blackbody radiation shift, the gravitational shift, the second-order Doppler shift and Rabi pulling were treated like in pulsed fountains. The second part of the table shows the shifts that we plan to evaluate next.

VI. CONCLUSION

In conclusion, the continuous cesium fountain frequency standard FOCS-2 is now in the metrological evaluation phase. Thanks to an innovative atomic beam source followed by state preparation into $|F = 3, m = 0\rangle$, we reached a high-level of frequency stability with an Allan deviation of $6 \times 10^{-14}\,\tau^{-1/2}$. We started the metrological evaluation of FOCS-2 with a detailed investigation of the frequency shifts which are specific to the continuous mode of operation and/or required the development of new experimental techniques. More precisely, we measured the light-shift from the source and showed that it is reduced to a negligible level with the help of our home-made rotating light-trap. The microwave cavity end-to-end phase shift has been evaluated with the cavity reversal technique. It represents the first direct measurement of a dephasing of the field between two distinct points of a microwave cavity in an atomic fountain. The methods developed for the evaluation of the second-order Zeeman shift in pulsed fountains are not applicable to our continuous fountain because of the difference in atomic beam geometry and temperature. We thus developed a new experimental technique involving the Fourier transform of Ramsey fringes which could find applications in other similar experiments. Finally, all the frequency shifts evaluated so far show that the continuous fountain FOCS-2 is compatible with an inaccuracy budget at the level of 10^{-15} or below.

ACKNOWLEDGMENT

This work was supported by the Swiss Federal Office of Metrology (METAS), the Swiss National Science Foundation (grant 200020-121987/1 and Euroquasar project 120418), and the University of Neuchâtel.

REFERENCES

[1] R. Wynands and S. Weyers, "Atomic fountain clocks," *Metrologia*, vol. 42, no. 3, p. 64, 2005.

[2] A. Joyet, G. Mileti, G. Dudle, and P. Thomann, "Theoretical study of the dick effect in a continuously operated ramsey resonator," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, pp. 150–156, 2001.

[3] J. Guéna, G. Dudle, and P. Thomann, "An experimental study of intermodulation effects in an atomic fountain frequency standard," *Eur. Phys. J. Appl. Phys.*, vol. 38, no. 2, pp. 183–189, 2007.

[4] N. Castagna, J. Guéna, M. D. Plimmer, and P. Thomann, "A novel simplified two-dimensional magneto-optical trap as an intense source of slow cesium atoms," *Eur. Phys. J. Appl. Phys.*, vol. 34, pp. 21–30, 2006.

[5] P. Berthoud, E. Fretel, and P. Thomann, "Bright, slow and continuous beam of laser-cooled cesium atoms," *Phys. Rev. A*, vol. 60, pp. R4241–R4244, 1999.

[6] G. Di Domenico, L. Devenoges, C. Dumas, and P. Thomann, "Combined quantum-state preparation and laser cooling of a continuous beam of cold atoms," *Phys. Rev. A*, vol. 82, no. 5, p. 053417, Nov 2010.

[7] F. Fuzesi, A. Jornod, P. Thomann, M. Plimmer, G. Dudle, R. Moser, L. Sache, and H. Bleuler, "An electrostatic glass actuator for ultra-high vacuum: A rotating light trap for continuous beams of laser-cooled atoms," *Rev. Sci. Instr*, vol. 78, p. 102109, 2007.

[8] L. Devenoges, G. D. Domenico, A. Stefanov, A. Joyet, and P. Thomann, "Improvement of the stability of an atomic fountain clock by making use of optical pumping," *in preparation*, 2011.

[9] G. D. Domenico, L. Devenoges, A. Stefanov, A. Joyet, and P. Thomann, "Fourier analysis of ramsey fringes observed in a continuous atomic fountain for in situ magnetometry," *arXiv:1104.0263*, 2011.

Improved Accuracy Evaluation of the NPL-CsF2 Primary Frequency Standard

Ruoxin Li, Kurt Gibble
Department of Physics
Pennsylvania State University
University Park, 16802 USA

Krzysztof Szymaniec
Time and Frequency Group
National Physical Laboratory
Teddington, UK
Email: ks1@npl.co.uk

Abstract— **Presented is quantitative evaluation of two leading uncertainties in the NPL-CsF2 fountain frequency standard. The distributed cavity phase shift evaluation is based on recent theoretical model where the cavity field is decomposed into a series of 2D Fourier components in azimuthal angle in the cylindrical cavity. Predictions of the model are reproduced experimentally. The microwave lensing effect is caused by dipole forces originating from radial variation of the microwave field amplitude and cavity apertures. The new evaluation of the two effects together with other recent improvements reduce the total type B uncertainty of NPL-CsF2 to 2.3×10^{-16}.**

The first accuracy evaluation of the primary frequency standard NPL-CsF2 was performed in 2009 [1]. The largest contributions to the total type B uncertainty were distributed cavity phase shifts (DCP) and the microwave lensing of the atomic cloud. Here we quantitatively evaluate these leading uncertainties.

Our DCP uncertainty evaluation is based on a recent theoretical model [2]. The difficulty of accurate three-dimensional finite element modeling of the cavity field's phase is overcome by decomposing the field into a series of two-dimensional Fourier components $\cos(m\phi)$; ϕ is an azimuthal angle in the cylindrical cavity. After measuring a number of cavity resonances to precisely determine the cavity geometry, the fields and atomic response were then calculated for typical operational parameters of the NPL-CsF2 fountain. Frequency errors due to $m \geq 3$ modes are negligible since the atomic trajectories are localized near the cavity axis. Predictions of the amplitude dependence (Rabi frequency for the interaction in the cavity) for $m = 0$ and $m = 1$ were reproduced experimentally.

The agreement of the power dependence for $m = 0$ DCP shifts validates the model, which allows us to make a correction of 1.6×10^{-17} for the predicted shift at optimal power; we take half of this value as an uncertainty. We experimentally evaluate the $m = 1$ DCP uncertainty by tilting the fountain [3], operating at optimal power. This gives an $m = 1$ DCP uncertainty of 7.7×10^{-17}, with a 0.3 mrad uncertainty of the fountain tilt from vertical along two orthogonal axes. For $m = 2$, we model the imaging of the atomic cloud at the detection and use the model to set a DCP uncertainty at 7.8×10^{-17}. We also apply a correction of 13.8×10^{-17} for the $m = 2$ DCP frequency shift [4]. The total DCP uncertainty of 1.1×10^{-16} is dominated by the $m = 1$ tilt sensitivity and $m = 2$ uncertainty.

With a reduced DCP uncertainty, the largest uncertainty is then the yet unobserved microwave lensing of the atomic cloud by the dipole forces of the microwave clock field. The dipole forces occur due to the radial variation of the microwave field amplitude and a frequency shift results from the system apertures [5]. We evaluate the shift to be 6.2×10^{-17} for the circular apertures and the 2D dipole forces of our cylindrical cavity in NPL-CsF2. The shift is corrected, with an assigned uncertainty of 50% of its value. Future measurements will allow this to be further reduced.

Further improvements recently introduced to NPL-CsF2 fountain standard optimize the operation near the zero-collisional shift point [6] and include a re-evaluation of microwave leakage. These improvements and the quantitative evaluation of the DCP shift and the microwave lensing reduce the type B uncertainty of NPL-CsF2 by about a factor of 2, to 2.3×10^{-16}.

ACKNOWLEDGMENT

Contributions of Sang Eon Park and Akifumi Takamizawa to the NPL experimental work are gratefully acknowledged. We acknowledge financial support from the NSF, Penn State, and the UK National Measurement Office.

REFERENCES

[1] K. Szymaniec, S.E. Park, G. Marra, W. Chalupczak, "First accuracy evaluation of the NPL-CsF2 primary frequency standard", Metrologia, vol. 47, pp. 363-376, 2010.

[2] R. Li and K. Gibble, “Evaluating and minimizing distributed cavity phase errors in atomic clocks”, Metrologia, vol. 47, p. 534-551, 2010.

[3] F. Chapelet et al., “Investigation of the distributed cavity phase shift in an atomic fountain” Proc. 20th European Frequency and Time Forum, p. 160, 2006.

[4] R. Li, K. Gibble and K. Szymaniec, 2011 (submitted).

[5] K. Gibble, “Difference between a photon’s momentum and an atom’s recoil”, Phys. Rev. Lett.,vol. 97, p. 073002, 2006.

[6] K. Szymaniec and S.E. Park “Primary Frequency Standard NPL-CsF2: Optimized Operation Near the Collisional Shift Cancellation Point”, IEEE Trans. Instrum. Meas., 2011 (in press).

Measurements with Multiple Operational Fountain Clocks

Steven Peil, Scott Crane, James Hanssen, Thomas B. Swanson and Christopher R. Ekstrom
Clock Development Division
United States Naval Observatory
Washington, DC 20392
e-mail: steven.peil@usno.navy.mil

***Abstract*—Performance of the first two operational rubidium fountains at the USNO is presented using relative measurements and comparisons against other timescales. Recent results with four fountains indicate frequency agreement at the 10^{-15} level and good agreement with the primary frequency standards contributing to TAI.**

I. Introduction

The U. S. Naval Observatory (USNO) maintains an ensemble of atomic clocks for generation of a Master Clock, the physical realization of UTC(USNO). As many as 20 hydrogen masers and 70 commercial cesium clocks can contribute to the ensemble. The timescale used to generate UTC(USNO) relies on the short-term frequency stability of the masers and the long-term stability of the cesium clocks. In order to advance timekeeping activities, 7 rubidium fountain clocks are being added to the clock ensemble.

The fountains will serve as a long-term frequency reference, similar to but better than the lower-stability cesium clocks, but will not provide any tie to the SI second. In order for a fountain to be integrated into the USNO clock ensemble like any other clock, it generates a continuous 5 MHz output that can be measured against the Master Clock. The 5 MHz is generated from a precision synthesizer referenced to a maser that also serves as the LO reference for a fountain; the fountain adjusts the synthesizer's output based on its measurement of the maser frequency. The maser can act as a flywheel to minimize the effect of interruptions to fountain operation, but the fountains have been designed to operate continuously for long periods of time.

II. NRF2 and NRF3

The design and operation of the first operational fountains, NRF2 and NRF3, have been described in [1]. The fountains were installed in a new clock building at the end of 2008, and have been under evaluation for about 2 years. During that time, much has been learned about obstacles to continuous, uninterrupted operation. Imperfections in the environmental controls in the clock building and faulty commercial electronics have disrupted fountain performance. Additionally, the semiconductor laser systems that we use need to be serviced at intervals of 1 to 2 years.

A. Relative Comparison

Despite these obstacles, several good periods of evaluation were obtained. Data from an uninterrupted 100-day comparison between NRF2 and NRF3 are shown in Fig. 1. The plot in Fig. 1(a) is the relative phase between the fountains after a single relative phase and frequency have been subtracted. The relative frequency removed from the data is 7.4×10^{-16}. A linear fit to the corresponding frequency data returns a slope of $4.2(2.7)\times10^{-18}$/day. The uncertainty is derived from the result in [2] for determining errors on frequency drifts in the presence of white frequency noise. In general, we measure frequency drifts that are consistent with zero (see section III. D.)

The Allan deviation between the two fountains, shown in Fig. 1(b), exhibits white frequency noise out to an averaging time of at least 23 days, with an average stability per fountain of $1.6\times10^{-13}/\tau^{1/2}$. During this 100-day comparison, each fountain operated virtually hands-free.

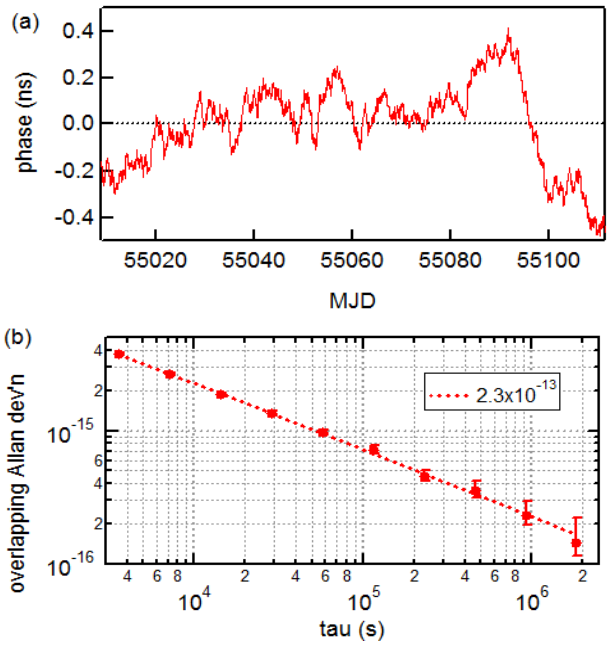


Figure 1 Data from 100-day comparison of NRF2 and NRF3. (a) Relative phase after removing single relative frequency and phase. (b) Overlapping Allan deviation versus integration time.

Funded by SPAWAR PMW 120

B. *Comparison to USNO Timescales*

Fig. 2 shows a plot of the relative phase of NRF3 versus the USNO cesium mean and maser mean. The cesium mean is a timescale based solely on the cesium clocks and represents a long-term frequency reference; the maser mean tracks the cesium mean because each maser is characterized against the cesium timescale before the maser mean is generated. The plot in Fig. 2 demonstrates a relative drift between the long-term reference and the two rubidium fountains measured to be $2.1(0.5)\times10^{-17}$/day.

While the comparison between NRF2 and NRF3 was limited to 100 days, NRF3 ran uninterrupted for a period of 6 months. Over this period of time, NRF3 served as an internal UTC predictor; because the Circular T provides UTC some number of days in the past, the present value of UTC needs to be estimated.

III. Fountain Ensemble

A. *Fountain Mean*

Recently a second pair of fountains, NRF4 and NRF5, has begun operation, and all four fountains have been running for about 2 months. One of the ways the rubidium fountains will be used for timekeeping is to characterize the frequencies and frequency drifts of individual masers, which is currently carried out with the cesium clocks. With four fountains running, a fountain mean is in the process of being implemented as part of standard timescale generation.

Fig. 3 shows the relative phase of a preliminary fountain mean over 60 days versus the maser mean, as well as the relative phase of the cesium mean and the maser mean. No relative frequencies have been removed, and the masers are characterized by the cesium mean in each case. The plot demonstrates the dramatic improvement in short-term noise of a timescale generated with only four fountains compared to that generated with dozens of commercial cesium clocks, which directly translates to improved efficiency with which the masers can be characterized, as well as the relative drift between the fountain mean and cesium mean.

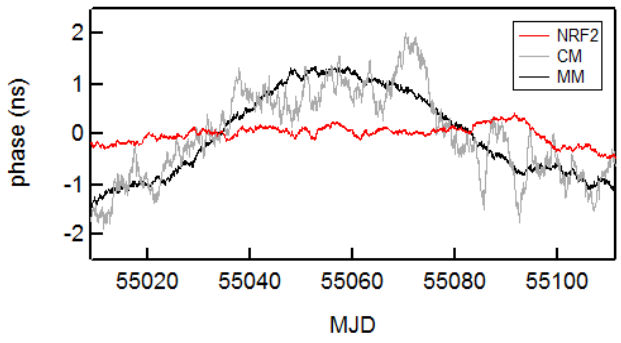


Figure 2 Relative phase of NRF3 and the cesium mean (grey), maser mean (black) and NRF2 (red), versus MJD. Each set of data has had its own relative phase and frequency removed. The NRF2 curve is the same data as in Fig. 1(a).

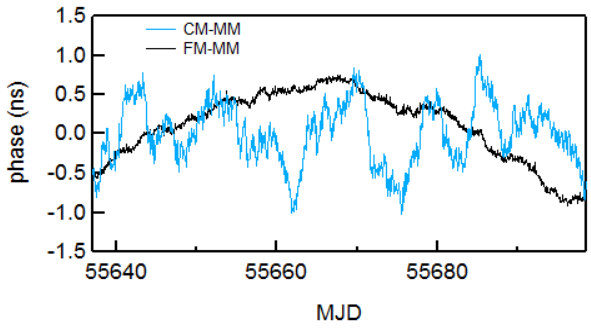


Figure 3 Plot of the relative phase of the fountain mean and the maser mean (black) and the cesium mean and the maser mean (blue) versus MJD.

B. *Relative Frequencies*

How well the frequencies of the four fountains agree is an interesting question. This is a question of the reproducibility of building and fielding these systems; in addition, the two pairs of fountains (NRF2 and NRF3 versus NRF4 and NRF5) have some different design elements.

The most significant potential source of frequency differences between fountains is the second-order Zeeman shift. The size of this shift is $575B^2$ (Hz/G^2). The fountains' magnetic fields are adjusted to be very close in value, to better than 1 μG, using the frequency of the central fringe of the magnetic field-sensitive $1\rightarrow1$ transition. Any error in determining this central fringe could lead to a frequency difference between fountains. The other most significant frequency biases, the black-body and gravitational shifts, should be common mode to a very high degree for the four systems. Other potential biases that could introduce frequency differences among the fountains are the cold-collision shift, cavity pulling, distributed cavity phase shift, and AC stark shift from stray laser light. The parameters determining these biases vary among some of the fountains.

Using data from the first 60 days when all of the fountains were running, the relative frequencies of all fountain pairs are measured. The results are shown in Fig. 4, where the error bars represent statistical uncertainties corresponding to the Allan deviation for a given comparison assuming white frequency noise. The raw frequencies (no biases adjusted for) all agree at the level of 10^{-15} or better.

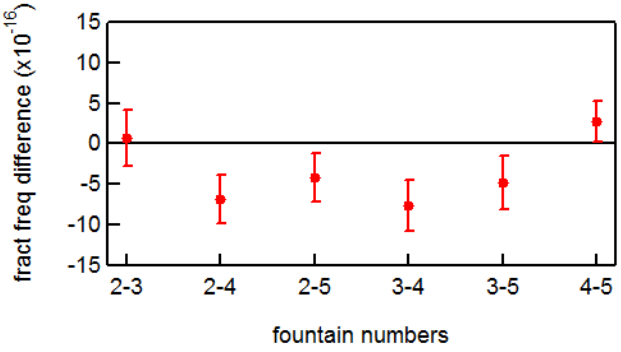


Figure 4 Fractional frequency difference between two fountains for all fountain pairs.

C. Frequencies Compared to Primary Standards

A rough measurement can be made of the fountain frequencies versus the frequencies of the primary standards used in determining International Atomic Time (TAI). The master clock is USNO's physical realization of UTC, which has the same frequency as TAI; a measurement of a fountain's frequency against the master clock, then, constitutes a measurement of the fountain compared to TAI. The frequency difference between TAI and the primary frequency standards is provided by the Circular T. The difference between these terms gives the frequency difference between a fountain and the primaries, which should correspond to the frequency bias of a given fountain.

We estimate the three most significant contributions to the frequency bias for a given fountain and remove them from the raw frequency difference between the fountain and the primaries. The results are shown in Table I. The values of the biases accounted for are -1.17×10^{-14} for the black-body shift for rubidium at T=295 K, 9.2×10^{-15} for the gravitational shift at an elevation of 84 m, and 6.97×10^{-14} for the second-order Zeeman shift at a bias field of 0.910 mG. In each case the uncertainty of the comparisons to the world's primary standards is dominated by the tie through TAI and UTC(USNO), which is estimated conservatively at 3×10^{-15}.

fountain number	Δv ($\times10^{-15}$)
NRF2	−.7
NRF3	−.8
NRF4	0
NRF5	−.3

Table I Average frequency difference between fountains and the primary frequency standards contributing to TAI after removing the frequency bias for each rubidium fountain.

D. Relative Frequency Drifts

The plot in Fig. 5 shows the measured relative frequency drift between all possible pairs of fountains over 60 days. All measured values are consistent with zero. The uncertainty is derived from the result in [2] for determining errors on frequency drifts in the presence of white frequency noise.

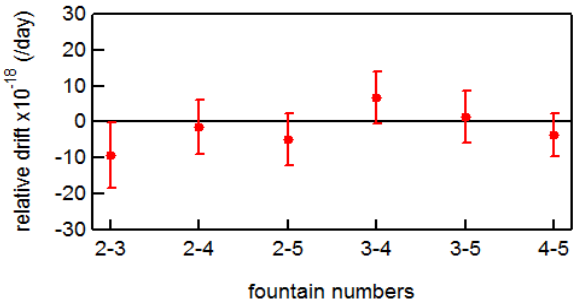


Figure 5 Plot of relative frequency drift for each pair of fountains.

IV. Scientific Application of Operational Fountains

Atomic clocks have found applications to fundamental science in addition to their roles in precise time and frequency metrology. According to current physical theories, which may need to be modified in order to arrive at a theory of quantum gravity, the relative frequencies of two atomic clocks using different atoms should be constant in both time and space. In particular, the principle of Local Position Invariance, an element of Einstein's Equivalence Principle, applied to atomic clocks dictates that the frequency difference between two different atomic clocks should stay constant as the clocks experience the same, varying gravitational potential. The varying gravitational potential that has been used to test this principle is provided by the elliptical orbit of the earth about the sun. This test therefore requires looking for an annual oscillation in the relative frequencies of two different types of atomic clocks.

The best test of this type to date has been carried out by comparing cesium primary frequency standards to hydrogen masers over the span of 7 years [3]. The discrete frequency measurements obtained with primary standards, typically several measurements per year, necessitates a long period of time to look for an annual variation. The authors of [3] relied on several primary standards to increase the rate of available measurements. Operation of a continuous fountain provides a high rate of frequency measurements against the reference maser (or other clock) that enables an annual variation to be searched for more efficiently. Estimates using a 6 month data set from NRF3 indicate that the statistical uncertainty achievable on a search for an annual term is comparable to the total uncertainty on the best current test of Local Position Invariance [3].

References

[1] C. Ekstrom, S. Peil, T. Swanson, S. Crane, in *Frequency Standards and Metrology: Proceedings of the Seventh Symposium*, Lute Maleki, ed., World Scientific, New Jersey, pp. 308-313 (2008).

[2] C. A. Greenhall, *Proceedings of the IEEE International Frequency Control Symposium*, Orlando, FL, pp. 428-432 (1997).

[3] N. Ashby *et al*., Phys. Rev. Lett. 98, 070802 (2007).

Quantitative Evaluation of Distributed Cavity Phase Shifts to Improve the Accuracy of SYRTE FO2

J. Guéna, S. Bize and A. Clairon
LNE-SYRTE, Observatoire de Paris, UMR CNRS 8630, UPMC
Paris, France
jocelyne.guena@obspm.fr

R. Li and K. Gibble
Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA

***Abstract*— We report measurements and *ab initio* calculations of the distributed cavity phase (DCP) shifts of the LNE-SYRTE FO2 atomic fountain clock. The measurements and validated model provide the first complete and quantitative evaluation of the DCP shift in an atomic fountain, reducing the FO2 DCP uncertainty to $\pm 8.4 \times 10^{-17}$. Here we emphasize the experimental techniques that precisely determine the fountain parameters to stringently test the DCP model.**

I. INTRODUCTION

The distributed cavity phase (DCP) shift is a first-order Doppler shift in atomic fountain clocks. The field in the microwave cavity used to interrogate the atomic transition is not a pure standing wave since the conductivity of cavity's copper walls is finite. Power flows from the feeds to the cavity walls, producing a small traveling wave and hence spatial variations of the phase of the field, the distributed cavity phase [1]. In a fountain, the atoms do not experience the same time-dependent phase variations on the two cavity traversals because of their transverse velocities, leading to this 1st order Doppler shift.

Our results demonstrating the first quantitative agreement between measurements and calculations of DCP shifts appear in [2]. In this paper, we do not repeat the results in [2] but instead present the supporting data that tightly constrain the fountain parameters, the position and velocity distributions of the atomic cloud, spatial variation of the atomic detection, and the truncation of the atomic cloud by apertures. We also illustrate the importance of using DCP shifts to balance the cavity feed amplitudes. This resolved a discrepancy in our first measurements. With the validated model, the DCP uncertainty of the SYRTE FO2 Cs fountain improves by a factor of 3.5, to $\pm 8.4 \times 10^{-17}$.

II. RABI OSCILLATIONS AS TEST OF THE MODEL FOR THE FOUNTAIN GEOMETRY

DCP shifts depend sensitively on a number of fountain parameters. Therefore, sufficiently stringent and independent confirmation of the parameters is helpful to critically validate the DCP calculations. We demonstrate that the Rabi flopping probability on the upward and downward fountain cavity traversals and the Ramsey contrast as a function of microwave amplitude provide sensitive and independent constraints.

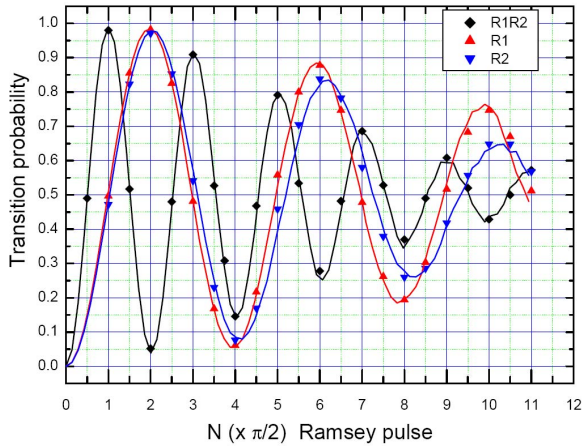


Figure 1. Comparisons between measured and predicted (curves) transition probabilities of resonant Ramsey pulses (R1R2) or a Rabi pulse on the upward (R1) and downward (R2) cavity traversals, as a function of the microwave amplitude, up to 11 $\pi/2$ pulses.

We begin with a knowledge of the fountain geometry and measurements of the initial atomic cloud size, position, and velocity distributions. A CCD image of the cloud, with a background of mechanical parts, and time of flight measurements are used. The detection response is determined by the detection laser beam power, spatial distribution, and detuning from the 852 nm atomic cycling transition, and a previous characterization of the spatial sensitivity of the fluorescence detection system. In Fig. 1 we show R1R2 in black, the usual Ramsey pulses, and then only Rabi pulses on the way up, R1 in red, and down, R2 in blue. For each of these, the microwave frequency is tuned to resonance. The points are measurements and the solid lines are from our model. The Ramsey R1R2 data essentially defines the amplitude scale, R1 measurements are most sensitive to the position distribution of the atoms during their upward passage, and R2 measurements probe the position distribution of the atoms during their downward passage, including the spatial inhomogeneities of the atomic detection. The excellent

This research was funded in part by CNRS, LNE, the NSF, and The Pennsylvania State University.

agreement is achieved with minor tuning of the detection curvature in one direction and the initial cloud transverse position. All other parameters in the model are the measured values. The contrast of these Rabi oscillations is very well reproduced for all three curves and, in particular, for the nominal clock operation (black point at 1 π/2) where our contrast is 0.985.

III. COMPARISON BETWEEN 2009 AND 2010 MEASUREMENTS OF THE M=0 DCP SHIFT

The importance of verifying the fountain parameters as in Fig. 1 became clear in our first measurements of the m=0 DCP shifts. The DCP calculation, with no tuning of parameters, reproduced the shape of the experimental data but with a clear discrepancy of a scale factor of about 1.5 [3]. We could find realistic values of the parameters to reproduce the m=0 DCP shifts, but these parameters were inconsistent with the Rabi flopping data in Fig. 1. Further work showed that there was a contamination of the m=0 shift by a residual m=1 shift due to feed asymmetries along with a small residual fountain tilt of about 200 µrad. We now balance the feeds by insuring that there are no frequency shifts at large tilts. Subsequent m=0 measurements with fountain tilts less than ±100µrad agree with the DCP calculations. In Fig. 2 the two data sets, before and after properly balancing the feeds (black and red squares) are plotted along with the simulations for the pure m=0 term with no free parameters (red curve) and a combined m=0 and m=1 with weights given by the known feed asymmetry and residual tilt (blue curve).

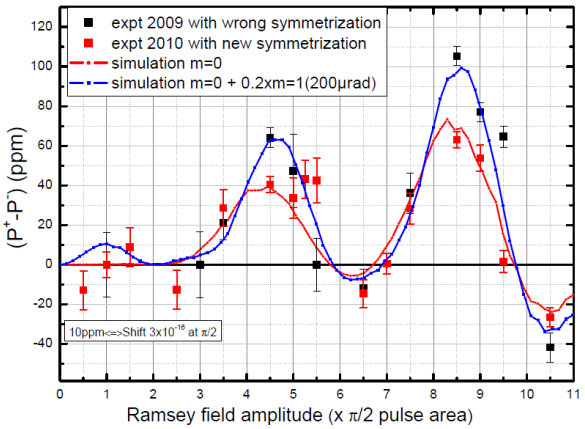


Figure 2. m=0 DCP shifts as a function of microwave power. The two data sets and calculations (curves) correspond to measurements with imbalanced feeds (black squares and blue curve) and with the feeds balanced to null the tilt sensitivity (red squares and curve).

REFERENCES

[1] R. Li and K. Gibble, Metrologia vol. 47, 2010, 534; R. Li and K. Gibble, Metrologia vol. 41, 2004, 376.

[2] J. Guéna, R. Li, K. Gibble, S. Bize and A. Clairon, "Evaluation of Doppler shifts to improve the accuracy of primary atomic fountain clocks", Phys. Rev. Lett , vol. 106, 2011, 130801.

[3] J. Guéna et al., "Characterization of the Distributed Cavity Phase Shift in FO2 for Improving the Accuracy of SYRTE Fountain Clocks", in Proceedings of the 2010 IEEE International Frequency Control Symposium (IEEE, New York, 2010), p. 307, and in Proceedings of the 24th European Frequency and Time Forum (Noordwijk, 2010).

Piston Temperature Measurement with SAW Sensors

T. Plum[1], S. Tourette[2], M. Loschonsky[3], M. Röbel[1]
[1] Deutz AG, Ottostr. 1, D-51149 Köln, Germany
[2] SENSeOR SAS, 694, Av du Dr. Maurice Donat, 06250 MOUGINS, France
[3] SENSeOR GmbH, Georges-Koehler-Allee 106, D-79110 Freiburg, Germany
Email: plum.t@deutz.com

***Abstract*— In this contribution we introduce a new approach for a piston temperature measurement system based on surface acoustic wave (SAW) resonators sensitive to temperature. The approach offers a new and competitive way to support the development of pistons and allows the research and improvement in combustion processes.**

I. INTRODUCTION

Piston temperature measurements face some major problems. The first measurements date back to the 1930s [1].They had been made on 12 inch bore diesel engines running at 250 rpm. One used sensors that were connected with cables to the measurement system. These cable led from the bottom of the piston over the moving crankshaft to the crankcase. The mechanical forces made the cables break, making long term measurements nearly impossible. A very commonly used method is based on active telemetric radio transmission. With the sender and receiver placed on the crankshaft. This approach also uses a wire connection between the piston, the transmitter and receiver unit with the threat of breaking wires. This measurement technique normally comes with a massive intervention in the engine parts.

Other measurement approaches are performed with mechanical linkage telemetry systems, where the sensor elements are connected directly to the data acquisition system with a metal arm with joints [2] as shown in Fig.1.

Further, the usage of metallurgical temperature sensors as shown on the leftside in Fig.1, often called as "Templugs", is a quite common approach [3]. Such sensors are made of steel alloys – serving as color indicators – which can be screwed or glued on parts that are exposed to temperature. The markers show an irreversible change in color depending on the temperature they are exposed to. The benefit of this temperature measurement approach is the easy handling and the low costs of the approach itself. Unfortunately, they can only serve as indicators for the maximum temperature they were exposed to without any further information of time of temperature gradients.

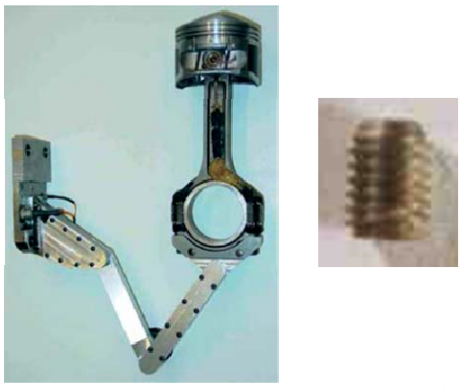


Fig. 1: Examples of piston temperature measurement approaches. *Leftside*: mechanical arm with joints. *Rightside:* metallurgical temperature sensor.

Regarding the requirements for measuring the temperature of a piston inside a running engine, surface acoustic wave temperature sensors seems to offer an easy to apply and low cost solution for an accurate validation of FEM simulated temperature profiles of pistons. Differential SAW resonators were used for achieving high accurate temperature measurements and minimizing effects like cross sensitivity due to acceleration, stress, frequency pulling or standing wave effects.

II. EXPERIMENTALS

A. Preparations before Testing

As first step to the measurement in the fired engine the optimal positions of the sensor and interrogation antenna with minor or no design changes to the engine were evaluated empirically. Therefore, sensor and antenna were mounted on potential positions and the signal link was evaluated by measuring the signal strength while turning the open engine manually until the optimal position for the

transmission with no design changes of the engine was found.
After having determined these optimal positions inside the engine, the piston and the crankcase were prepared for mounting the elements of the sensor system: SAW sensor, sensor and interrogation antenna to perform further reliability tests regarding the transmission on the open engine. Finally, a convenient signal strength up to a half stroke which resulted in an interrogation time slot of 10 ms at a speed of 2300 rpm was found.
The exact mounting was done as follows and shown in Fig.2. The machine parts were cleaned with acetone and isopropyl and an electric isolation layer for the sensor antenna was formed by an epoxy. On this layer the antenna and the cable to the SAW-sensor were screwed inside the piston and covered with the same epoxy suitable for high temperatures up to 300°C and resistant to engine fuels, oil and lubricants.
After tempering the glued piston in an oven with about 100°C over several hours, the piston was applied in the crankcase to verify the reception of the radio waves of the transceiver unit again.
The antenna in the crankcase was fixed with a hose clamp on the oil cooling injector. To support the cable with an additional protection it was placed in a small rigid pipe made of alumina.
To connect the antenna cable with the transceiver unit a hole in the crankcase was chosen that normally is used for the oil level gauge as shown in Fig. 4.

Fig. 2: Backside view of the prepared piston from the bottom of the oil pan. 1: transceiver antenna fixed to the oil cooling injector, a non moving part of the crankcase; 2: mounted sensor antenna inside on the bottom of the piston; 3: oil injector nozzle; 4: mounted SAW temperature sensor type TSEAS10 by SENSeOR; 5: Crankshaft.

The moving parts of an engine face beside the high temperature stress also forces cased by rough accelerations. Although the test engines are diesel engines that are typically used in heavy duty off-road applications that are operated at relatively low rotation speeds compared to spark ignited combustion engines that are used in passenger cars. The movement path of a typical piston in an typical consumer application sums up to 580 km operated at a speed of 1800 rpm for 24 hours [4].
The range of rotation speeds of interest were between 800 rpm and 2300 rpm. Within these ranges, a piston is exposed to rough acceleration values of approximately 400 g as shown in Fig. 3.

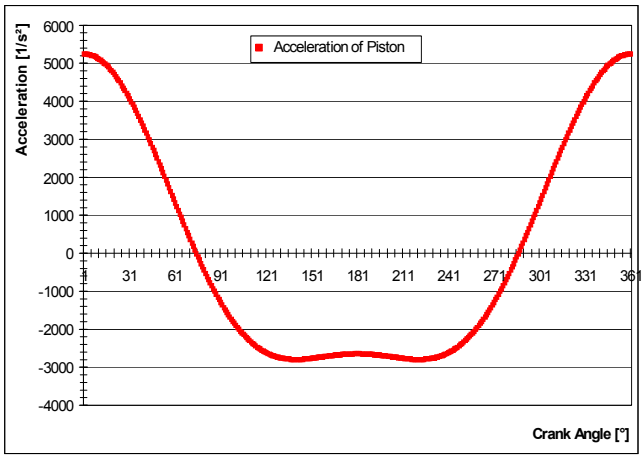


Fig.3: Acceleration about rotation angle

B. *Engine Test Set-Up*

Before the first run with the applied piston the transmission quality of the signal was tested with the mounted oil pan. No interferences were found with that setting.
These tests were repeated after installing the engine on the test cell (Fig. 4). Again the engine was turned by hand for safety reasons.

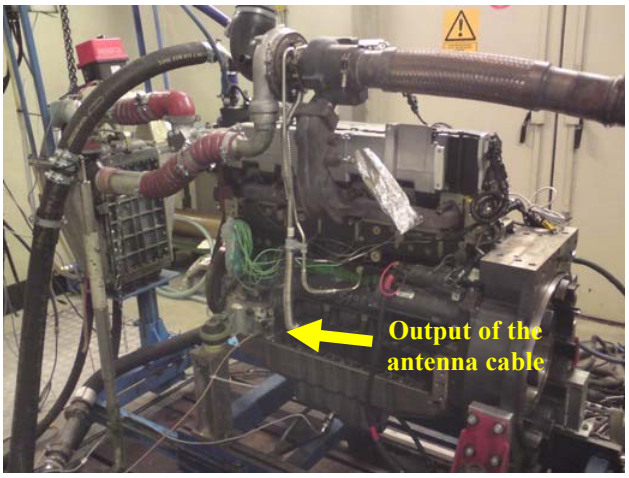


Fig.4: Engine TCD 7.8 in test cell before running.

C. *Engine Test Run*

The engine was fired and ran at 800 rpm. After a couple of minutes while reaching the optimal engine run conditions, the engine speed was changed to 900 rpm. The signal-strength was checked again to control the influence of oil dust. The risk of dumping the signal strength with the

increasing amount of oil dust was not observed and therefore the dust had no effects on the transmission of the radio waves and the measurements itself.
The engine speed was changed again after one could expect that the engine was in a stable work point. This implied that one did not expect significant temperature changes. Also the signal seemed to be stable enough to get a result for the temperature. The steps that were chosen for the engine speed change had been 100 rpm higher up to 2290 rpm as shown in Fig.8.
Occasionally the signal had to be optimized by changing the engine speed slightly for avoiding standing wave effects.
For all speed steps no load was put on the engine for this first run due to safety reasons.
After the test run the piston was removed of the engine and the screwed and glued antenna and the SAW-sensor were analyzed for damages or any relieves of the glue. The used adhesive component was robust enough to resist within the harsh environment inside a running and no damages of the sensor or the antenna was observed.

D. Signal Processing Strategies

Experimental data demonstrate that both resonances are rarely identified simultaneously as shown in Fig. 5 and Fig. 7. The differential measurement strategy is based on the principle that both a reference resonator with little temperature drift, and a measurement resonator with a strong temperature drift, patterned on a common piezoelectric substrate, are affected in a similar way by environmental disturbances (sensor aging, stress, radiofrequency parasitic components), and hence a frequency difference provides the needed long term stability and accuracy, independent of environmental conditions and interrogation unit aging.

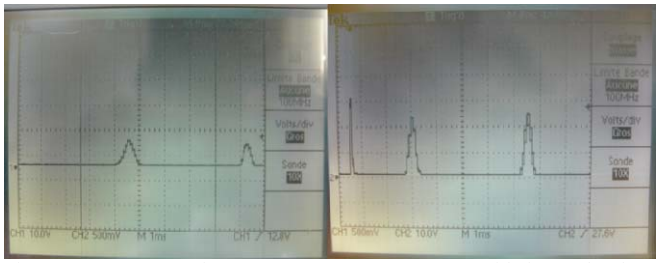


Fig. 5: Raw returned RF power measurement provided on the low frequency output of the reader. *Leftside*: the minimum acceptable signal level. *Rightside*: acceptable quality signal level. The output power is controlled using a proportional feedback loop in order to keep the returned power at the optimum level.

However, because the refresh rate is higher than several hertz, the reference and measurement frequencies are not needed during a same frequency sweep: measurements recorded during successive frequency sweeps are still close enough to validate the differential measurement strategy.
Hence, in order to increase the probability of a useful measurement despite the poor radiofrequency link budget, the last useful frequency measurement (either reference f_{ref} or measurement f_{mes}) were propagated to all successive unsuccessful measurement attempts. Nevertheless all frequency sweeps for which neither reference nor measurements frequencies were visible, were rejected.
Thanks to this strategy, a useful dataset has been gathered exhibiting a meaningful temperature evolution trend. The conversion from a frequency difference $\Delta f = f_{ref} - f_{mes}$ to a temperature requires the application of calibration coefficients individually identified during a preliminary calibration step. Indeed, based on a second order polynomial fit of the frequency f dependence with temperature T, the function

$$\frac{f_0 - f}{f_0} = CTF_1 \times (T - T_0) + CTF_2 \times (T - T_0)^2$$

classically used in the literature, with f_0 the frequency at reference temperature T_0 and CTF_1, CTF_2 the first order Coefficient of Temperature and second order Coefficient of Temperature respectively, a temperature dependence law was deduced as a function of frequency difference Δf following $T = A_0 + \sqrt{A_1 + A_2 \times \Delta f}$.
In the case of the sensor used during these experiments, A_0 (°C) = 109.6, A_1 (°C^2) = -118553 and A_2 (°C^2/Hz) = -0.1284. The calibration coefficients are classically identified in the -20 to +160°C range.
The application of calibration coefficients to the raw measurements of a sensor located in a piston, associated to the in-plane monopole antenna (35 mm interrogation range), yields the temperature measurement exhibited in Fig. 6.

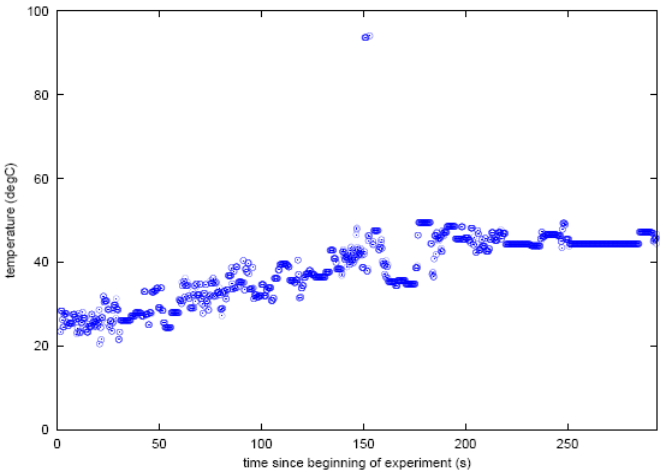


Fig. 6: Measurement of the temperature of a sensor located in the piston of an engine running from 800 to 2290 RPM. The frequency difference was converted to a temperature using calibration coefficients established prior to the experiment. Outlying points were filtered automatically.

This dataset was cleaned by removing the points below 433.3 MHz and above 433.7 MHz for the measurement resonance, and below 434.1 and above 434.44 MHz for the reference resonance. As seen in another measurement set (Fig. 7), obtained in the same rotation speed conditions and with the same antenna, these outlying points yield significant measurement dispersion. Our current

understanding of the source of these unwanted points is associated with the automatic gain control on the emitted power since the sensor is not seen by the interrogation antenna most of the time; the emitted power is maximum (+10 dBm). If the sensor happens to come close to the interrogation antenna – within reception range – while the emitted frequency is close to the resonance, then the received signal might appear as a local maximum even though it is not the global maximum of the resonance. This effect will be strongly reduced by extending the interrogation range (asymmetric dipole antenna), which will help the automatic gain control to remain locked most of the interrogation time, but most significantly also means that the unwanted points will remain far from the main measurement curve (since close the resonance, locking on a side-lobe has a low probability considering the frequency sweep rate compared to the duration the sensor is seen by the interrogation unit). Hence, a digital post-processing step will always easily remove these unwanted data.

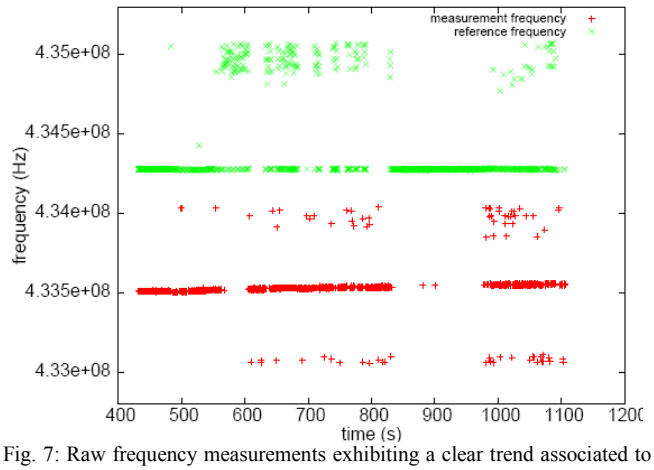


Fig. 7: Raw frequency measurements exhibiting a clear trend associated to the measurement (measurement frequency around 433.5 MHz, reference frequency around 434.2 MHz) and outlying points widely scattered above and below the actual measurement.

III. Results and Discussion

The temperature was calculated of the measured frequency values. The measurement showed frequency values which seemed to be induced by reflections of the signal.

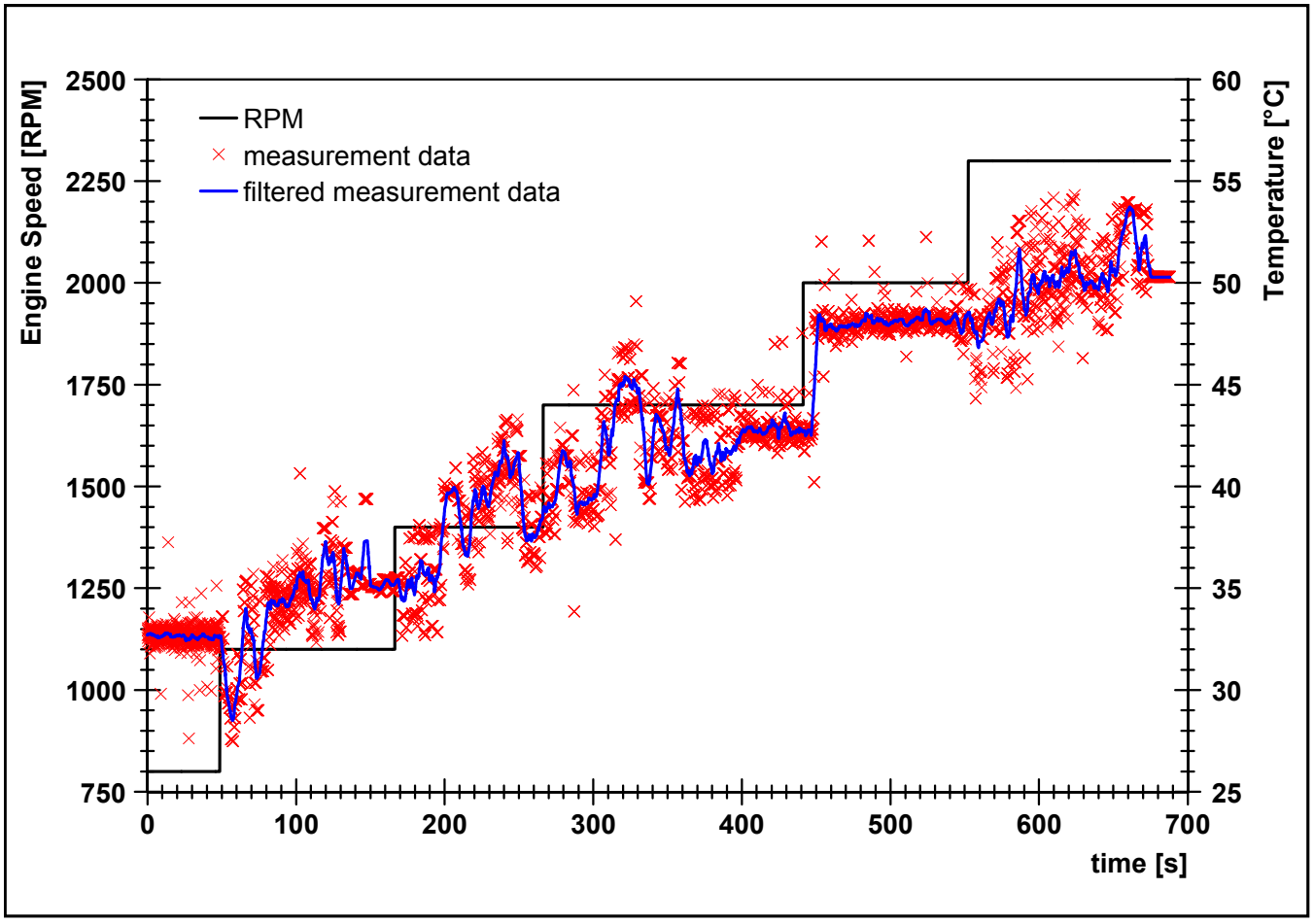


Fig.8: Measured and filtered temperature values in relation with the recorded engine speed during a test run.

After processing these data as described above, the temperature value at the piston could be determined as shown in dependency of the engine speed in (Fig. 8).

These first terms provide encouraging results since usable resonance frequency signals were obtained even with a poor interrogation range. Various antenna configurations have been identified which allow for an extended interrogation range, potentially up to the full displacement of the piston (yielding a continuous view of the sensor by the interrogation antenna). Further sensitive parameters have been identified during the tests on the site and have not necessarily been optimized yet. The signal processing steps to provide a consistent temperature reading were applied to the recorded signal successfully. As further steps, using an optimized antenna configuration and locating the sensor close to the final position appear as reasonable milestones, to be extended with the inclusion of multiple sensors which require an extended frequency interrogation range since each resonance occupies a different 900 kHz-wide frequency slot beyond the ISM range.

Therefore a new approach of a new and competitive way to support the development of pistons that allows the research and improvement in combustion processes had been identified. Successful tests with engine running at 800 rpm up to 2300 rpm had been demonstrated where SAW sensors able to withstand accelerations as high as 400 g inside running engines with an interrogation time of 10 ms.

ACKNOWLEDGMENT

R. B. G. thanks Prof. Dr.-Ing. P. Brychta, Prof. Dr. techn. L. M. Reindl, Dr. J.M. Friedt and Dipl.-Ing. D. Eisele, for scientific discussion, SENSeOR GmbH, Germany and SENSeOR SAS, France, both for technical support and DEUTZ AG, Germany for financial support.

REFERENCES

[1] H. W. Baker, "Operating Temperatures of Cast Iron and Aluminum Pistons in 12-Inch Bore Oil Engine," *Proceedings of the Inst. of Mechanical Engineers, ,* vol. 127, 1934.

[2] J. Chang, Z. Filipi, D. Assanis, T.-W. Kuo, P. Najt, and R. Rask, "Characterizing the thermal sensitivity of a gasoline homogeneous charge compression ignition engine with measurements of instantaneous wall temperature and heat flux " *Int. J. Engine Res.,* vol. 6, p. 289, 2005

[3] L. K.-H., C.-H. Shek, and J. K.-L. Lai, "Metallurgical Temperature Sensors," *Recent Patents on Mechanical Engineering,* vol. 1, p. 225, 2008.

[4] DEUTZ_AG, "Genuine DEUTZ Piston - Customer information, North America," 2011.

.

Characterization of WO_3 layers deposited on Quartz and Lithium Niobate SAW resonators for the design of gas sensors

M. David, M. Arab, J.R. Gavarri
Université du Sud Toulon Var
IM2NP UMR CNRS 6242
La Garde, France
marjorie.david@univ-tln.fr

D. Rabus[1], W. Daniau[1], T. Baron[1], J.M. Friedt[*], L. Delmas[*], S. Ballandras[1*]
[1] FEMTO-ST, UMR 6174 CNRS-UFC-ENSMM-UTBM
* SENSeOR, TEMIS Innovation
Besançon, France
sylvain.ballandras@femto-st.fr

***Abstract* — Surface acoustic wave (SAW) gas sensors generally require the use of a reactive layer for molecule adsorption. WO_3 has been identified for a long time as a high potential sensitive layer particularly for NH_3, NO_X, CO, etc. We report here on the characterization of elastic properties of such material using the dispersion behavior of SAW propagating under gratings passivated wich WO_3 films of various thicknesses. Quartz as well as $LiNbO_3$ SAW devices are used in that purpose, allowing for the derivation of a reliable data set. Complementary structural characterization using direct measurement techniques (DRX, AFM, TEM,) are reported to confirm the analysis deduced from SAW measurements. As a conclusion, the exploitation of WO_3 for SAW-based sensor is discussed.**

I. INTRODUCTION

Surface Acoustic Wave (SAW) sensors are presently receiving an increasing interest, although many works has been dedicated to operate to such developments. A particular effort is paid to develop devices capable to detect gas flows in common environment for health protection or security purposes. The arising of new materials and technologies actually pushes to re-investigate the capability of such devices to efficiently operate, particularly in harsh environment such as toxic atmospheres and high temperature. Hence, the possibility to remotely interrogate gas sensors based on resonator structures can be actually implemented and tested. In this work, we propose to characterize the operation of SAW resonator sensors coated with tungsten trioxide (WO_3), a layer which exhibits selective detection capabilities when heated near 350°C. The exploited principle consists in the change of dielectric properties of the film when adsorbing specific molecules (O_3, NO_2, NH_3 and ethanol). The design of such sensors requires the knowledge of the actual properties of the detection film as it is directly deposited atop the resonator electrodes, yielding modal dispersion which must be accounted for to guaranty the device operation according given requirements (such as those imposed by wireless interrogation in ISM bands). In that purpose, test devices have been fabricated, based AT Quartz and (YX*l*)/128° $LiNbO_3$ synchronous resonators operating near the 434-MHz-centered ISM band, coated with various WO_3 thicknesses (100, 200 and 300 nm). These thins films were characterized using X-ray diffraction/Rietveld refinement and the samples morphology were analyzed by Atomic Force and transmission electron microscopy. The diffract pattern showed the monophasic layers with the same crystallinity and they reveal a pseudo cubic phase (isotropic behaviour). We then have measured the reflection coefficient of the resonators allowing for fitting the coating elastic properties by model updating. This fit has been first achieved using quartz resonators and we have obtained the fitted coefficients (apparent elastic constants C_{11}, C_{12}, C_{44} and mass density ρ). They have been used to predict the coating-induced frequency shift on lithium niobate resonators. Although mass density is found to change dramatically versus WO_3 thickness (from 6500 to 3900 $kg.m^{-3}$), the elastic constants reveal more stable (C_{11} and C_{44} typically equal to 15, 8 and 2.5 GPa respectively). The max coating thickness actually is found near 300 nm, but it turns out that above 200 nm, the resonator response is too degraded to allow for any operation. A specific design then is proposed on $LiNbO_3$ for assessing the actual operation of the sensor.

In the first section of the paper, we report a direct analysis of the deposited material also has been achieved to complement and justify effective data derived using SAW–based measurements. Next, we briefly describe WO_3 coating properties, with a particular insight in physical parameter changes related to gas adsorption and the subsequent SAW-sensor application. As such devices need accurate design procedure to optimize their actual operation; the characterization of WO_3 elastic properties was a prerequisite and therefore the experimental measurements of the corresponding physical coefficients then is reported.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

Finally, exploitation of the corresponding results using $LiNbO_3$ SAW resonators is discussed.

II. WO_3 THIN FILM PROPERTIES FOR GAS DETECTION

The structural configuration of the WO_3 crystal lattice is the distorted rhenium trioxide (ReO_3) structure [1,2]. The crystal structure of tungsten trioxide is temperature dependent. The most common structure of WO_3 is monoclinic (P21/*n*). Based on the reversible changes in conductivity, WO_3 thin films have been used as sensitive layers towards O_3, NH_3, NO_x, CO, ethanol [3,4]. It was shown that the conductivity of WO_3 thin films changed in presence of NO_x gas with a dependence of temperature. At 375°C the resistance of WO_3 drop by a factor of 100 compared to its initial value at room temperature from 6.10^9 to 6.10^7 ohms respectively [5]. This behaviour can be explained by physisorptions and chemisorptions mechanisms and catalytic reactions between gases and material surface (phenomenon of adsorption and desorption) [6]. Due to its electrical properties, WO_3 is an ideal candidate for applications SAW gas sensors. Indeed, the interaction between gases molecules and tungsten trioxide causes a change in conductivity of thins films WO_3 resulting from in a variation of the velocity and propagation loss of the acoustic wave. By measuring the change in the frequency characteristics of the SAW sensor, it was possible to determinate different physical parameters of the sensitive layer WO_3 and sees the influence of the viscoelastic layer on the SAW device (dielectric loss, mass effect) [7,8]. To achieve this, it is imperative to have strong knowledge in the field of piezoelectricity and acoustic.

III. STRUCTURAL CHARACTERIZATION

A. Experimental

WO_3 thin films on the surface of a quartz substrate for SAW devices were prepared by reactive radio frequency magnetron sputtering at 13.56 MHz, using a 99.9% pure tungsten target. The synthesis setup and standard procedure have been described in more detail by previous work of the lab [6]. The obtained films were annealed at 400°C for 3 hours.

B. Characterization and results

The morphological analyses of WO_3 thin films were conducted by Atomic Force (AFM) and Transmission Microscopy (TEM, Philips TECNAI F20 FEG).

The structure properties of WO_3 films were determined by X-ray diffraction with Cu.Kα radiation (Siemens – Brucker D5000 diffractometer). X-ray patterns were compared with those of the Joint Committee on Powder Diffraction Standards (JCPDS). The DBWS Tools 2.16 software was used to perform Rietveld Refinement. Figure 1.a shown the state surface of the deposited WO_3 films on quartz substrates and reveals a weak roughness and a porous surface. This porosity varies from 10 nm to 100 nm and they extended in deep layers. The TEM image (fig 1.b) shows a cross section of the oxide layer deposited on Si substrate reference. For all studied thicknesses of WO_3: 50, 100, 200 and 300 nm, all films were homogenous.

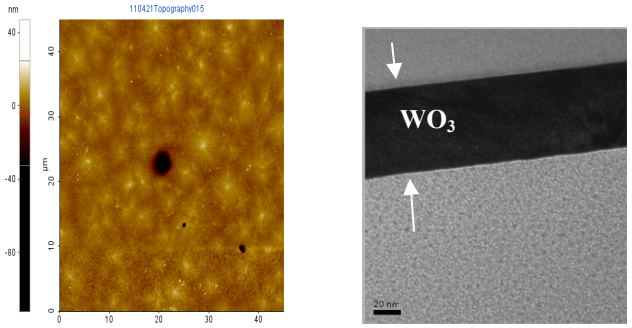


Figure 1. a) AFM image of WO_3 films on quartz substrate after annealing at 400°C for 3 hours; b) TEM cross section of WO_3 thin film thickness of 50 nm on Si substrate

The X-ray diffraction patterns and Rietveld refinement of trioxide tungsten are in good agreement (figure 2) showing a low crystallite size in accordance with TEM observation.

The XRD diagram show a polycrystalline structure identified as monoclinic phase and indexed from standard data sheet. The observed structural parameters are in good agreement with JCPDS (No. 83-0950) data [9].

From Rietveld refinement results, it is proved that WO_3 thin films have a monoclinic phase close to cubic one. The evaluated lattice parameter values of WO_3 are $a = 7.308$ Å, $b = 7.514$ Å, $c = 7.675$ Å, $\alpha = \gamma = 90°$, and $\beta = 90,632°$. This allows to hypothesize that WO_3 films behave as an isotropic system. All deposited layers (50 to 400 nm) were structured in the same phase and the calculated parameters remain unchanged.

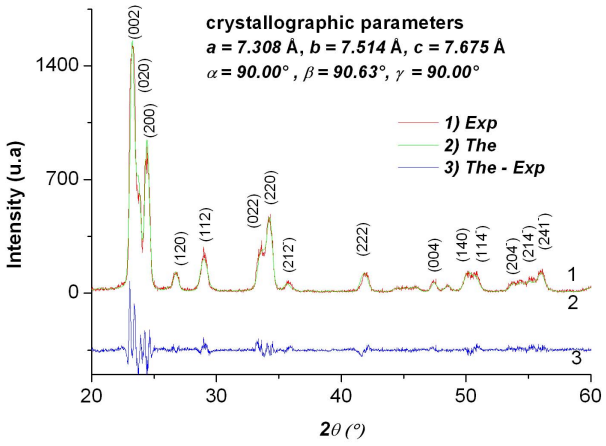


Figure 2. X-ray diffractogram (1.Exp) and theory Rietveld refinement (2.The) of WO_3 thin films with a thickness of 50 nm

IV. WO_3 Elastic constants characterization

In this section, the theoretical and experimental protocols for characterizing the WO_3 overlay properties are described, using Quartz SAW resonators (fig.3). The quality of the obtained physical paremeters is the assessed for $LiNbO_3$ (YZ) devices better suited for dielectric perturabtion detections.

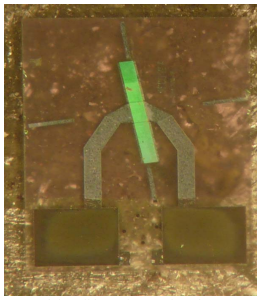


Figure 3. WO_3 thin film deposited on SAW devices

A. *Theoretical analysis*

As layered devices are concerned, with the WO_3 overlay covering the electrodes, one cannot make the economy for advanced simulation and analysis exploiting a combination of finite element and boundary element methods (FEM/BEM) developed in that purpose [10]. We here briefly recall the analysis approach, detailed description of the calculation fundamentals are reported in [11]. An harmonic admittance is derived from FEM computation of a single period of an infinitely long periodic transducer accounting for the inhomogeneous boundary conditions induced by the electrode/passivation surface conditioning. The acoustic contribution of the substrate is accounted for by Green's function based BEM. For an accurate simulation of the device behavior, one need a reliable set of physical constants, i.e. elasticity coefficients, mass density and dielectric parameters and associated intrinsic losses if available. Although some data can be found in the literature, providing a starting point for the simulation, a set of experiments has been achieved on SAW resonators of fig.1 built on a (ST,X+ψ) cut with ψ a small non zero propagation angle, therefore providing clear contributions at both edges of the stop-band and hence a direct evaluation of the reflection coefficient magnitude. This will be used for improving the theory/experiment agreement and consequently the accuracy of the above-mentioned constant determination.

B. *Experimental implementation*

As explained above, SAW resonators built on Quartz have been used first to determine the WO_3 overlay properties required for sign and analysis purposes. The asynchronous single-port resonators implemented in that purpose consist of one interdigited transducer (IDT) surrounded by two Bragg mirrors. They have been designed to operate at best in the 434-MHz centered ISM band. The IDT is composed of 82 Al electrode pairs with a mechanical period p of 3.627 µm and a metallisation ratio a/p of 0.8 to benefit for the reflection coefficient stabilitiy versus electrode width a. As suggested above, the mirror period differs from the IDT one, being fixed to 3.646 µm (same metallisation ratio) to force the resonance within the stopband of the mirrors and hence increasing their reflecvtion efficiency. The metal height h is 120 nm ($h/\lambda_{a\,c} \sim 1.5\%$). The acoustic apperture is 320µm (40 λ_{ac}) to avoid any possible diffraction effect. These devices have been then covered with WO_3 overlays of different thicknesses (50, 100, 200, 300 and 400 nm) to achieve the characterization discussed in the previous paragraph. Figure 4 shows the corresponding experimental electrical admittances compared with theoretical harmonic admittances derived with FEM/BEM. For each configuration, a set of material constants has been fitted to achieve the maximum agreement between both curves, reported in Table I. As one can easily see, the acoustic dissipation increases along the WO_3 thickness in such a way that above a 300 nm thick deposition, the experimental signal was no more exploitable (fig.5). Also the viscoelastic losses of the overlay has been fitted according to experimental observations, yielding an equivalent mechanical quality factor of 50 whatever was the WO_3 thickness.

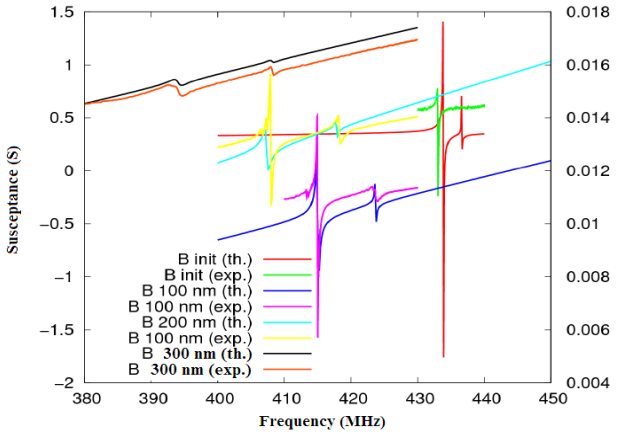


Figure 4. Theory/experiment assessment of SAW resonators on AT-Quartz cut for different WO_3 overlay thickness

In all the simulations, a relative dielectric constant of 230 has been fixed according to published material [12]. As shown by Table I, the mass density was changed along the WO_3 thickness for best fit purposes. Although one could also change the elastic constants in that matter, the reflection coefficient gave helpfull tracks for the fitting strategy, as it strongly depends on the combination of all the elastic constants. One could imagine keeping the mass density unchanged versus WO_3 thickness but the ratio between the elastic constants would then be kept unchanged. From a structural point of view, the film is assumed polycristalline and porous. This porosity vary with the deposition thickness. Within each grain, the elasticity behaviour is not changed but the overall of the layer actually reduces along the deposition height.

Materials	AT Quartz			(YX*l*)/128° $LiNbO_3$		
Physicals properties	***100nm***	***200nm***	***300nm***	***100nm***	***200nm***	***300nm***
Density ρ ($kg.m^{-3}$)	6500	4700	3900	6500	4700	3900
C_{11} (GPa)	14	15	12.5	14	15	15
C_{12} (GPa)	8	8	7	8	8	8
C_{44} (GPa)	2.5	2.5	2.5	2.5	2.5	2.5

Table I .Apparent elastic constants and density of the WO_3 versus overlay thickness. The initial value of the density was fixed to 7200 $kg.m^{-3}$ according to the literature and adjusted versus thickness for best fit purposes.

C. *Validation on $LiNbO_3$ SAW resonators*

The previous section has shown the way the physical parameters of the WO_3 overlay, more specifically the elastic constants and mass density, has been fitted. To confirm the reliability of these data, another set of experiments has been achieved using SAW resonators built on (YZ) $LiNbO_3$ cut with similar WO_3 deposition conditions and thickness (fig 3). As for Quartz resonators, a FEM/BEM analysis then was conducted using the fitted parameters to assess the experimental results. Figure 4 shows the corresponding experimental admittances compared with theory. For best fit purposes, slight changes have been applied to the elastic parameters but the overall behaviour of the device responses was well predicted using Quartz-resonator-derived data, yielding confidence in their physical meaning. Particularly, the frequency shift related to mass loading was found in good agreement here again and therefore, the above-proposed hypothesis concerning the porous surface and weak roughness was not infirmed. From a pragmatical point of view, these results show that the fitted apparent elastic constants and mass density can be used for the design of a next generation of sensors. This work is currently under development.

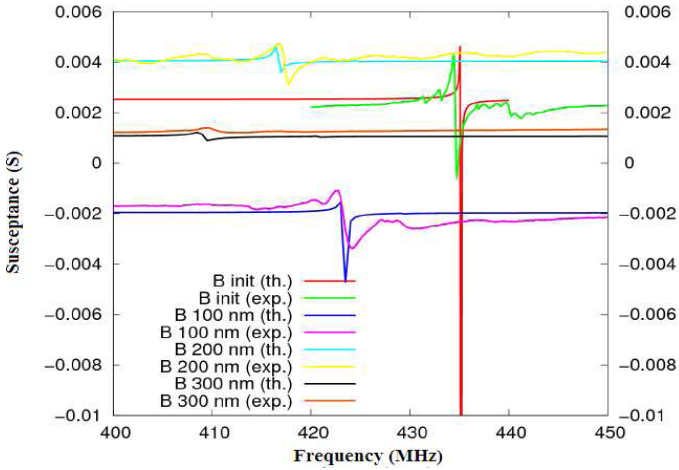


Figure 5. Theory/experiment assessment of SAW resonators on (YX*l*)/128° $LiNbO_3$ for same WO_3 overlay thickness

V. CONCLUSION

In this work, we investigated the design and the physical properties of WO_3 using SAW devices as a mean of characterzation. A pseudo cubic as isotropic system was identified for WO_3 sensitive films via structurals analysis. The measured reflection coefficient allowed to estimate the elastic properties and induced frequency shift on quartz/lithium niobate resonators as a function of the WO_3 layers thickness. Although mass density is found to change as a function of WO_3 thickness and their porosity, the apparent elastic constants reveal more stable. The max coating thickness actually is found near 300 nm, but it turns out that above 200 nm, the resonator response is too degraded to allow for any operation.

ACKNOWLEDGMENT

We gratefully acknowledge the PACA Regional Council, the General Council of Var, and the agglomeration community of Toulon Provence Mediterranean for their financial supports.

References

[1] P. Cox, "The Electronic Structure and Chemistry of Solids", Oxford University Press, Oxford, 1987.

[2] M.S. Whittingham, in: B.V.R. Chowdari, S. Radharkrishna (Eds.), Proceedings of the International Seminar on Solid State Ionic Devices, World Publishing Co., Singapore, 1988.

[3] M. D. Antonik, J. E. Schneider, et al., "Microstructural effects in WO, gas sensing film", Thin Solid Films, vol. 256 , 1995 , pp. 247-252.

[4] M. Stankova, X. Vilanova, et al, "Influence of the annealing and operating temperatures on the gas-sensing properties of rf sputtered WO_3 thin-film sensors" Sensors and Actuators, vol. B 105, 2005, pp. 271–277.

[5] Jong-In Yang, H. Lim, Sang-Do Han, "Influence of binders on the sensing and electrical characteristics of WO_3 based gas sensors," Sensors and Actuators vol. B 60, 1999, pp. 71–77.

[6] M. Bendahan, J. Guérin et al., "WO_3 sensor response according to operating temperature: Experiment and modelling", Sensors and Actuators, vol.B 124, 2007, pp. 24–29.

[7] S.J. Ippolito, et al., "Layered SAW hydrogen sensor with modified tungsten trioxide selective layer" Sensors and Actuators, vol.B 108, 2005, pp. 553 – 557

[8] A.Z. Sadek, "A polyaniline/WO_3 nanofiber composite-based ZnO/64◦YX $LiNbO_3$ SAW hydrogen gas sensor", Synthetic Metals, vol.158, 2008, pp.29–32

[9] L. Lozzi, L. Ottaviano, M. Passacantando, S. Santucci, C. Cantalini, Thin Solid Films 391, 2001,pp.224

[10] S. Ballandras, A. Reinhardt,et al, "Simulations of surface acoustic wave devices built on stratified media using a mixed finite element/boundary element integral formulation", Journal of Applied Physics, vol. 96, N°12 , 2004, pp.7731-7741.

[11] S. Ballandras, A. Reinhardt,et al, "Theoretical analysis of damping effects of guided elastic waves at solid/fluid interfaces", Journal of Applied Physics, vol. 99, 2006, pp.054907:1-9.

[12] Salje E, "Lattice dynamics of WO_3". Acta Crystallographica Section A, 1975, 31:360.

Langasite SAW temperature and oxygen multi-sensor

P. Zheng,[1,2] D.W. Greve,[1,3] I.J. Oppenheim,[1,4] and V. Malone[4]
[1]National Energy Technology Laboratory, Pittsburgh, PA, USA
[2]Department of Physics, Carnegie Mellon University, Pittsburgh, PA, USA
[3]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
[4]Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
Email: pengzhen@andrew.cmu.edu

***Abstract*—We report here a langasite SAW device suitable for sensing temperature and oxygen concentration, with applications such as monitoring the exhaust of oxy-fuel combustion systems. In this paper we report on the observed temperature dependence of oxygen sensitivity, explain the mechanism for this behavior, and present a scheme for temperature compensation.**

I. Introduction

Langasite ($La_3Ga_5SiO_{14}$) surface acoustic wave (SAW) devices have potential applications in harsh-environment sensing. Important applications include exhaust gas oxygen sensors for the control of oxy-fuel combustion systems. In such combustors it is possible to achieve a nearly pure carbon dioxide exhaust suitable for geologic sequestration. In this application oxygen sensing in the temperature range from 500 to 1000 ºC is needed. Langasite SAW devices have been operated as resonators up to 1000 ºC [1]; langasite SAW sensors for H_2 and C_2H_2 gas have been operated up to 550 °C [2]; and wireless temperature sensing has been demonstrated up to 850 °C [3,4]. Here we report recent results obtained for langasite SAW oxygen sensors. We particularly focus on the temperature dependence of these sensors, and a method for reducing the impact of temperature flucuations on gas sensing.

II. Temperature and Gas Sensing

The SAW sensors considered here utilitze the electroacoustic effect, where a conducting layer near the piezoelectric surface influences the surface wave velocity. Gas sensing is achieved by using a layer with a conductivity that depends on the gas concentration. The most suitable conducting layers for this high-temperature application are metal oxides, which exhibit a semi-conducting behavior with conductivity dependent on the oxygen defect concentration. In metal oxides with predominant electron conduction, a reversible reaction occurs with gas-phase oxygen of the form

$$O_o^x \leftrightarrow V_{\ddot{o}} + 2e' + \tfrac{1}{2}O_2^{(g)} \qquad (1)$$

where O_o^x is the concentration of oxygen atoms on an oxygen site, $V_{\ddot{o}}$ is the concentration of doubly-negative oxygen vacancies, e' is an electron, and $O_2^{(g)}$ is an oxygen molecule in the gas phase. Higher oxygen partial pressure in the gas phase drives the reaction to the right, decreasing electron concentration and increasing the electrical conductivity. It can be shown that the conductivity is given by

$$\sigma = Ae^{-E_A/kT}P_{o_2}^{1/m} \qquad (2)$$

where E_A is the activation energy for conduction, kT is the thermal energy, and P_{O_2} is the oxygen partial pressure. Increasing conductivity of the sensing layer decreases the surface wave velocity, as the redistribution of charge in the sensing layer reduces the stored energy in the electic field above the surface, reducing the piezoelectric stiffening. The change in velocity can be predicted analytically; Figure 1 shows the calculated change of surface velocity for a (0, 138.5, 26.6) langasite SAW device.

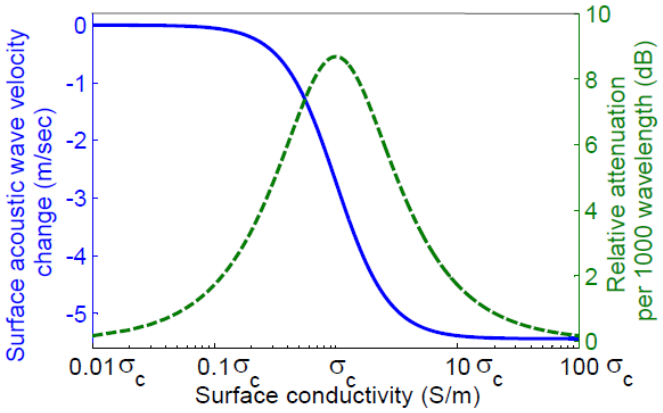


Figure 1. Calculated surface acoustic wave velocity changes (solid blue line) and relative attenuation (green dashed line) of langasite substrate with Euler angle (0, 138.5, 26.6) as a function of surface sensing film conductivity.

It is apparent from Figure 1 that significant velocity changes are only obtained for a limited range of sensing layer conductivity. As the conductivity is thermally activated according to (1), a particular sensing layer will have a limited temperature range where it exhibits useful response. Further, changes in temperature will also cause a change in surface wave velocity. We focus here are removing the impact of

small temperature changes on the surface wave velocity. We will also determine the temperature range over which on particular sensing layer is useful. In a practical sensor, the temperature and oxygen concentration need to be measured independently; this will be addressed in future work.

III. EXPERIMENTAL

SAW devices used in this work were fabricated on 76 mm diameter, 0.5 mm thick langasite wafers (Roditi International, UK) with Euler angles (0, 138.5, 27) using a lift-off technique. The metallization was e-beam deposited and consisted of 100 nm platinum with 10 nm titanium as an adhesion layer.

The SAW devices used in this work have the transmitter located in the center with open reflectors were located on two sides. The IDTs had 2 µm finger widths and gaps, yielding an 8 µm surface acoustic wavelength and a center frequency of 334 MHz at room temperature. For gas sensing, A 200 nm ZnO sensing film was deposited by reactive RF sputtering from a Zn target at a pressure of 4 mT and RF power 50 W, in 25% O_2/ 75% Ar. ZnO was masked from the IDT terminals using Mylar tape. The SAW device was then spin coated with polysiloxanes (IC1-200 and DC4-500 from Futurrex, Inc.). The polysiloxanes film was partially removed using a cotton tip with acetone and then annealed in air at 400 °C for 30 min to form a 100 nm SiO_2 insulating film on the IDT transmitter and terminals region.

XRD characterization of the sputtered ZnO film showed a wurtzite structure with a 30 nm grain size. In some experiments temperature sensing was performed at the same time using a second SAW device with longer path lengths, so that its reflections were not coincident with the first reflections of the gas-sensing device.

The SAW device(s) were placed in a furnace (Lindberg STF55433C) with computer-controlled gas flows. Connection to the measuring electronics was made using shielded thermocouple wire as a transmission line, connected to the SAW device using fine thermocouple wire and conductive ceramic paste. Measuring electronics consisting of a National Instruments (NI PXI-5670) and a vector signal analyzer (NI PXI 5661) [5] was used to measure the phase changes resulting from temperature and/ or gas concentration changes. Measurement of the phase changes requires a reference pulse; this reference pulse consisted of either another SAW reflection or the attenuated exciting pulse.

We first discuss separate measurements of temperature and gas concentration. Figure 2 shows the extracted effective velocity change as a function of temperature. The velocity decreases monotonically with temperature, consistent with previous reports [4,6,7].

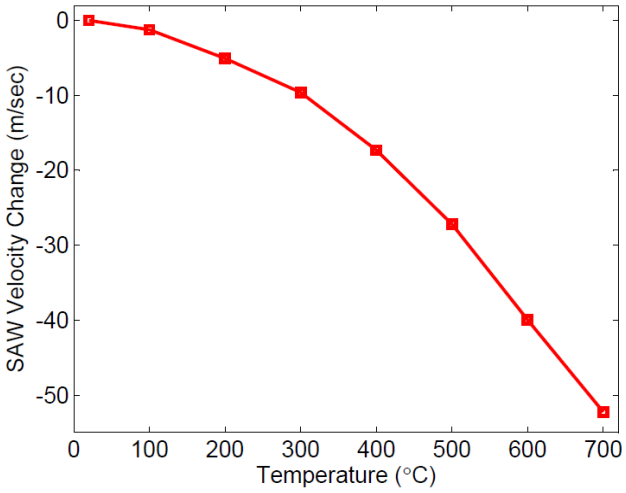


Figure 2. Extracted effective SAW velosity change as a function of temperature.

The effect of oxygen concentration was examined by measuring the phase change between \the exciting pulse and a reflection normalized for a 1 µs propagation time. The oxygen concentration was changed in a stepwise fashion with the balance of the gas being nitrogen. The oxygen concentration was varied by changing the mass flow controller setpoints while keeping the total mass flow constant. Figure 3 shows the results for three temperatures, where the phase changes consistent with *n* type conduction in the ZnO. While the phase changes are consistent with velocity changes due to the electroacoustic effect, they could also be caused by temperature variations in the furnace due to imbalanced gas flow. This is a particular concern at lower temperatures where the observed phase changes are small.

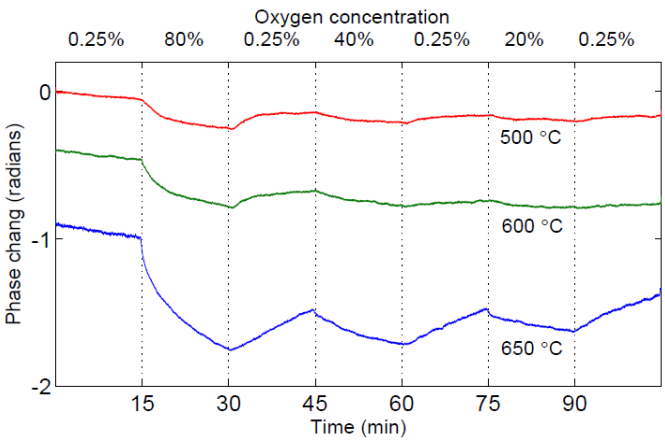


Figure 3. Phase change resulting from changes in the oxygen concentration.

The influence of small temperature variations can be removed by simultaneously measuring SAW reflections on two propagation paths, one with and one without without a ZnO sensing layer. This can be achieved by patterning the sensing layer so that some reflections are not influenced by the change in sensing layer resistivity. Alternatively, two separate SAW devices can be connected in parallel, one with and one without a ZnO sensing layer (Figure 4). In this case the SAW

devices must be different so that the reflections are distinguisable.

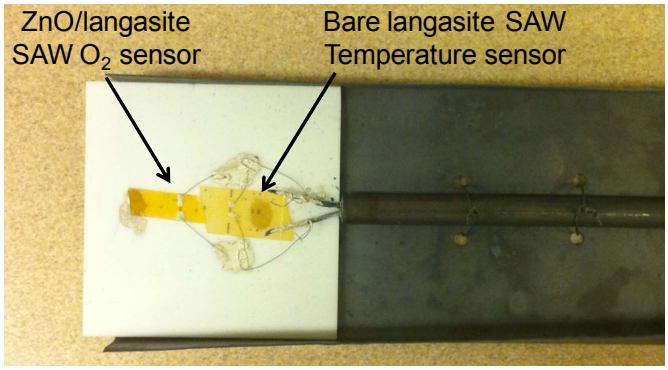


Figure 4. Two langasite SAW devices connected in parallel, one with and one without a ZnO sensing layer.

Figure 5 shows the observed reflections for two SAW devices connected in parallel at room temperature. Here the second series of reflections are from the temperature-sensing SAW device without a ZnO sensing layer. These reflections are strongest despite the longer path length; this is because the ZnO sensing layer causes additional attenuation [5].

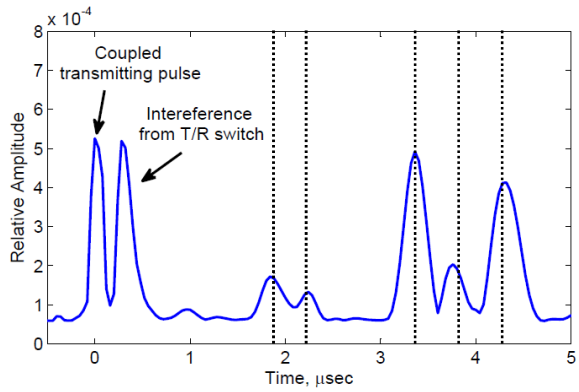


Figure 5. Reflections for two SAW devices wired in parallel at room temperature.

Figure 6 shows uncorrected phase measurements for the oxygen and temperature sensors (top two traces) and measurements compensated for temperature variations (bottom trace) at 650 ºC. Temperature compensation is achieved by subtracting the gas and temperature phase change measurements normalized for the propagation time. The data from the temperature sensor shows small perturbations (± 0.5 ºC) due to the gas flow switching. Removing these small perturbations shows that most of the phase change of the O_2 sensor was indeed resulting from the gas sensing layer.

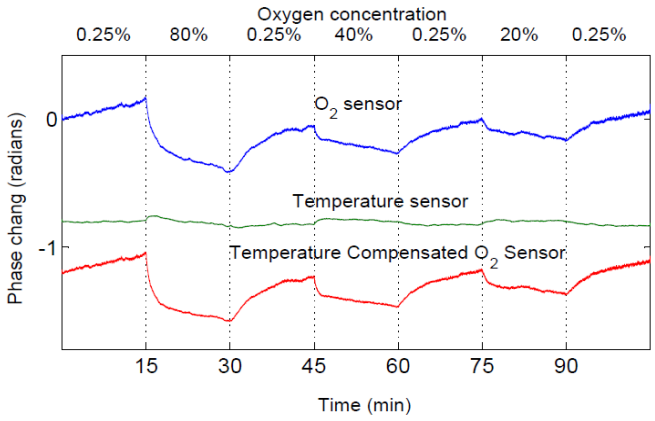


Figure 6. Compensation for temperature variations with a second SAW device.

Gas sensing performed over a wider range of temperature is presented in Figure 7. Data for 650 C and above has not been temperature-compensated due to failure of the temperature-sensing SAW device. Oxygen sensing however has been achieved up to 700 C.

This data shows clearly that sensing with a ZnO layer can only be performed over a limited temperature range. As shown in Figure 1, a significant change in propagation velocity is obtained only for a limited range of sensing layer conductivity. Due to the exponential dependence of conductivity on temperature, a peak in the sensitivity is expected at an intermediate temperature.

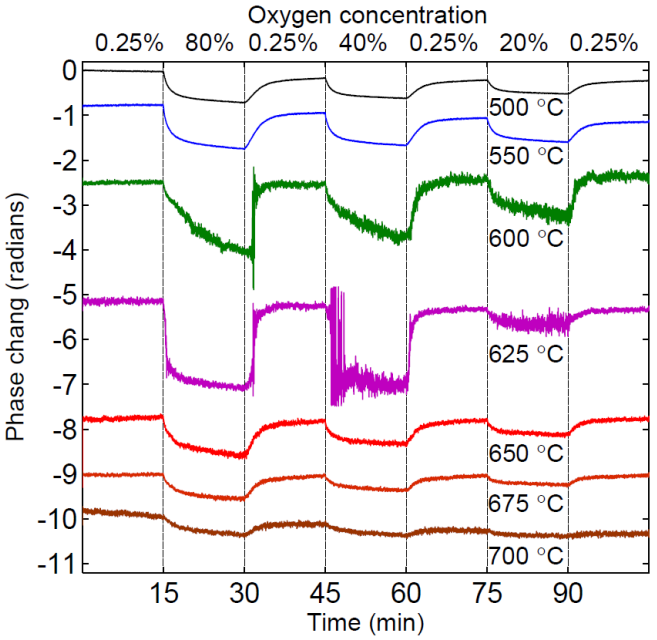


Figure 7. Oxygen sensing as a function of temperature.

The peaked behavior for the sensitivity as a function of temperature is an important issue in sensor design. In order to acheive accurate sensing over a range of temperature, sensing layers with little dependence on temperature are preferred. Recently some mixed-oxide sensing layers have been

investigated that have reduced temperature sensitivity [8]. Even in this case, the activitation energy is not exactly zero so simultaneous temperature measurement will still be needed.

Finally, we consider the issue of sensor stability. In early experiments, we evaluated the oxygen sensing as a function ot temperature at steadily increasing maximum temperatures of 630, 700, and 750 C. The total phase change resulting from an oxygen concentration change from 0.25% to 80% is plotted in Figure 8 for these three series of measurements. We see that the peak sensitivity shifts to lower temperatures with increasing temperature exposure. This change is consistent with increasing grain size in the ZnO layer, resulting in higher conductivity (increase in the factor A in equation (1)).

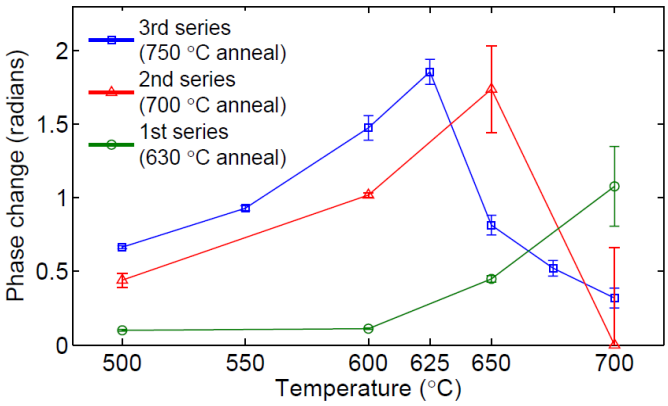


Figure 8. Measured phase change from 0.25% to 80% O_2 concentration at different temperature in three series of tests of the ZnO/langasite device.

In a subsequent series of experiments (not shown) closely similar results were obtained in three series of measurements. This suggests that the response of the sensing layer is stable when operated below the maximum anneal temperature.

IV. Conclusions

We have reported here on a langasite SAW sensor for simultaneous measurement of temperature and oxygen concentration. The sensitivity of the propagation velocity resulting to oxygen concentration depends on temperature because of the thermally activated conductivity of the sensing layer. As a result small perturbations in temperature will cause a change in the measured phase shift even if the oxygen concentration is constant. This effect can be minimized by using two SAW devices or two propagation paths, one with and one without a sensing layer.

We have also observed that the gas sensitivity can be influenced by aging of the sensing layer. This effect appears to be reduced when sensors are annealed at higher temperatures than used during operation.

Acknowledgement

This work was performed in support of research on carbon storage at NETL under RES contract DE-FE0004000.

References

[1] M. P. da Cunha, T. Moonlight, R. Lad, D. Frankel, and G. Bernhard, "High temperature sensing technology for applications up to 1000°C", in 2008 Proc. IEEE Sensors Conference, pp. 752 – 755.

[2] J.A. Thiele, and M. P. da Cunha, "High temperature SAW gas sensor on langasite", in 2003 Proc. IEEE Int. Conf. Sensors,Vol.2, pp. 769 – 772.

[3] A. Canabal, P.M. Davulis, G.M. Harris and M. P. da Cunha, "High-temperature battery-free wireless microwave acoustic resonator sensor system", Electronics Letters, vol. 46, issue 7, pp 471 – 472, 2010.

[4] P. Zheng, et al., "Pulse-mode temperature sensing with langasite SAW devices," in 2010 Proc. IEEE Int. Freq. Contr. Symp., pp. 297-300.

[5] P. Zheng, et al., "Langasite SAW Device with Gas-Sensitive Layer," in 2010 Proc. IEEE Inter. Ultra. Symp. (in press).

[6] J. Bardong, et al., "Precise Measurements of BAW and SAW Properties of Langasite in the Temperature Range from 25 °C to 1000 °C," IEEE International Frequency Control Symposium, 2008, pp. 326-301.

[7] S.Q. Wang, J. Harada and S. Uda, "A Wireless Surface Acoustic Wave Temperature Sensor Using Langasite as Substrate Material for High-Temperature Applications," Jpn. J. Appl. Phys. 42, pp. 6124–612, 2003.

[8] F. Neri et al., "$FeSrTiO_3$- based resistive oxygen sensors for application in diesel engines," Sensors and Actuators B: Chemical 134, pp. 647-653 (2008).

SAW Sensor Correlator System Performance Parameters

N.Y. Kozlovski
Mnemonics Inc.
Melbourne, FL 32934

M.W. Gallagher and D.C. Malocha
School of Electrical Engineering and Computer Science
University of Central Florida
Orlando, FL 32816-2450

***Abstract*—There has been little published data on spread spectrum SAW RFID correlator receiver performance, since most approaches published have used an FMCW system. The purpose of this paper is to discuss issues related specifically to correlator receiver performance parameters with respect to range, detection, and noise. The minimum detectable signal (MDS) at the ADC is used as the measure for prediction of the maximum range. The loop gain of the system, MDS, noise and processing gain bound the predicted achievable range for a correlator receiver. It will be shown that in a correlator receiver nano- to micro-joules of energy can obtain ranges greater than 50 meters, dependent on the key system parameters. A model is developed for prediction of the maximum range as a function of center frequency, output signal and power, MDS, synchronous interrogation, and loop gain. From the predictions, it appears feasible to have a sensor range of over 100 meters with modest interrogation energy in a pulsed correlator system.**

I. Introduction

The concept of using a correlator receiver for spread spectrum systems is well known. Results of a fully operational SAW orthogonal frequency coded (OFC) spread spectrum multi-sensor system have been recently reported [1]. The paper presented some of the important device and system considerations, such as the choice of operating frequency, device design, antenna size, loss, etc. The 915 MHz SAW temperature system presented was based on SAW OFC delay line devices that are interrogated with a pulsed RF signal. The temperature is extracted after signal reception through an analog-to-digital-converter (ADC) and the data points are post processed as in a correlator receiver, using an adaptive matched filter and associated software. The post processing is completely reconfigurable and is comparable to a software radio approach. Although an OFC SAW system will be used as an example, the correlator receiver analysis and approach should be applicable to any type of passive coded waveform system, SAW or others.

Surface acoustic wave (SAW) technology is beginning to seriously attract interest for a broad range of sensor applications, especially in aerospace and health monitoring applications [1] [2]. Many applications have very challenging requirements: zero maintenance (no battery), no external power (scavenging or external power source), reliable life-cycle (years in a wing structure or minutes in an engine exhaust), light and small, etc. A short list of system specifications may include simultaneous multi-sensor interrogation and reception, wireless, passive, radiation hard, and range of several centimeters to 10s of meters. The sensors should be small, rugged, provide RFID on chip, operate at cryogenic to high temperature, and differing embodiments should provide temperature, gas pressure, strain, chemo- or bio-detection and others.

Over the last 20 years there have been several proposed SAW embodiments for wireless, passive SAW RFID sensors, which include narrowband resonant devices, reflective delay line sensors, SAW chirp devices, external-sensor-SAW module, and CDMA [2]–[7]. Narrowband devices can provide an ID through differing resonant device frequency, while most delay line devices provide the coding through pulse position reflectors. The chirp sensor uses the correlation properties for enhanced sensor data extraction, but provides no effective multi-coding. Initial work on orthogonal frequency coded (OFC) SAW devices for RFID and communication began in 2000, and the first publication on SAW OFC was in 2004 [8] [9]. The research was begun to attempt to overcome the limitations of the previously described embodiments for device encoding, and then apply the approach to sensing.

II. OFC SAW Device Design

There have been a number of publications on the theory and approach to OFC based on communication theory, and then its application to SAW device embodiments [8] [9]. The basic embodiment for the OFC RFID tag and sensor is schematically shown in Fig. 1. A wideband transducer launches a SAW based on the interrogation signal, which is convolved with the OFC coded reflector array, and is re-radiated, via the transducer, to the receiver antenna. The chip orthogonality condition describes a relationship between the local chip frequencies and bandwidths, embodied in each SAW Bragg reflector. The reflector-chip frequency responses are a series of nearly ideal sampling functions with null bandwidths equal to $2 \cdot \tau^{-1}$. Each chip contains an integer number of carrier half cycles and the chip-Bragg center frequencies are separated by multiples of τ^{-1}. The key enabling technology is the fact that the nulls of adjacent Bragg reflectors align with all the peaks of the individual Bragg reflectors, indicating that the SAW signal

The authors acknowledge support from NASA STTR grants and the NASA Graduate Student Researchers Program (GSRP) Fellowship.

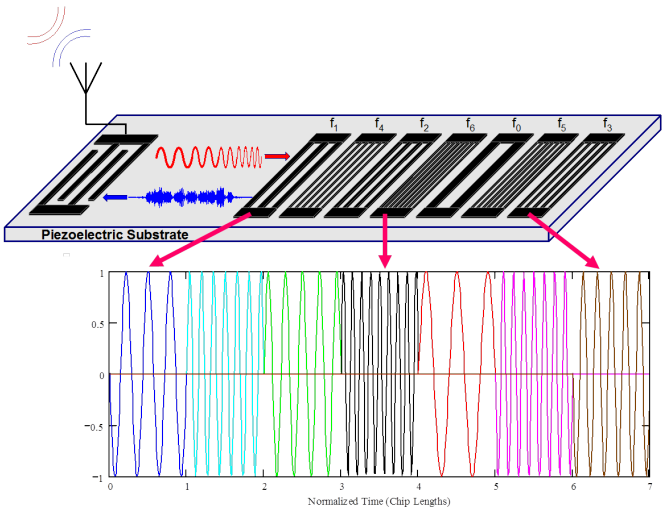


Fig. 1. Schematic of a 7 chip SAW OFC RFID tag which can be used as the platform for a sensor. The figure depicts a chirp input signal and the returned coded signal which is the convolution of the OFC code and chirp.

is nearly transparent to all Bragg reflectors, except near the center frequency and bandwidth of a single Bragg reflector. This allows the SAW to travel through the Bragg array without much attenuation, until a signal frequency coincides with the chip-Bragg reflector frequency. Coding is accomplished by shuffling the chips in time, which allows both frequency and time diversity. The OFC approach produces a wide or ultra-wide band spread spectrum device. The sensor information is encoded in the reflectors, time delay regions, or both. Dual tracks (in-line or parallel) can be used for enhanced coding or multiple sensor operations.

The OFC SAW devices used herein are at a center frequency of 915 MHz and used 5 chips per device with a bandwidth of approximately 92 MHz. This yields a device signal processing gain (PG) of 25.

III. Synchronous Transceiver Approach

The SAW OFC sensor device and system initial concept have been previously discussed [9]. The basic system concept, shown in Fig. 2 is composed of multiple SAW RFID-sensors (RFIDS) that can have various embodiments. The SAW and antenna compose the RFIDS. The transceiver, often called the reader for RFID systems, sends out an interrogation signal which is received by all the SAW sensors. The interrogation signal launches a SAW, which is encoded with the RFID and is appropriately modified to also encode the sensor information, and is rebroadcast to the receiver. The signal is demodulated and post processed to extract the RFID and the associated sensor information.

A conceptual diagram of the interrogation/receiver process is shown in Fig. 3, which uses a chirp interrogation signal and a correlator receiver. The chirp provides increased signal power over a single pulse and allows ultra wide band operation, if desired. The implementation of the actual reader hardware is more complex, but the operational principles remain the same. The near-baseband signal is post processed through an analog-to-digital converter and software.

An example of a 915 MHz pulsed synchronous, RF transceiver block diagram developed is shown in Fig. 4

IV. Signal and Noise Model

The system should attempt to maximize the detection probability, or minimize errors, as a function of system parameters. The system must overcome the signal path loss by amplifying the signal while minimizing all noise source effects .The following will define the signal-to-noise ratio based on the model parameters. From the model, with suitable approximations, the range for any synchronous correlator based system can be estimated.

The system critical operational parameters are:

- Center Frequency (f_o), Bandwidth (B)
- TxRx Antennas
- RFIDS Parameters
- Equivalent Isotropic Radiated Power ($EIRP$)
- Minimum Detectable Signal (MDS)
- Receiver Gain, Noise Figure, and Noise Sources

A. Parameter Definitions

The following elements within the model are defined. The elements do not consider details of implementation or optimization.

ADC	Ideal Analog-to-Digital Converter
MDS	Minimum Detectable Signal
S	Signal Power Measured at ADC
N	Noise Power Measured at ADC
kT	Thermal Noise Energy
$EIRP$	Equivalent Isotropic Radiated Power
G_{RFIDS}	RFIDS Gain (Less than unity for passive device)
G_{Rx-Ant}	Gain of Receiver Antenna
G_{Rx}	Receiver Gain from Antenna Output to ADC
PG	Signal Processing Gain of the System (τB)
PL	Path Loss
NF	Receiver Noise Figure
N_{ext}	External Noise Source at Antenna Output
N_{ADC}	ADC Equivalent Noise
N_{sum}	Number of Synchronous Integrations in ADC
PCG	Pulse Compression Gain

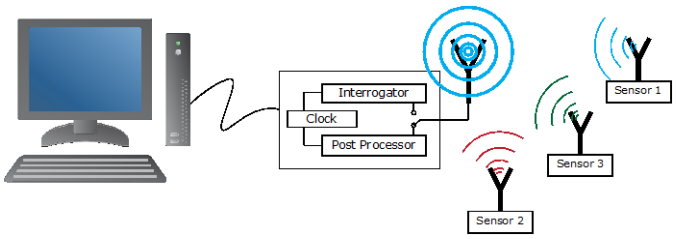


Fig. 2. Block diagram of a SAW wireless sensor system. The system will interrogate multiple sensors simultaneously. Receiving and identifying the RFID, the sensor information can be obtained via post processing of the received signal.

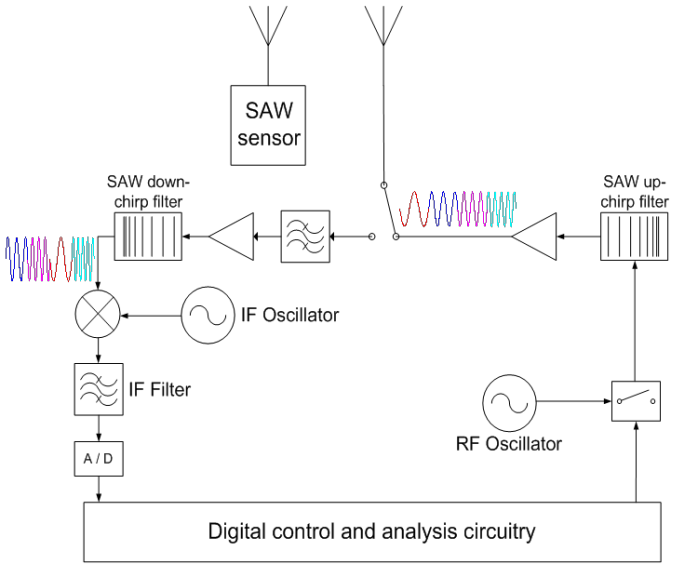


Fig. 3. A simple block diagram concept of the interrogation and correlator receiver approach as the sensor reader.

B. S/N Model

The signal is given as

$$S = EIRP \cdot G_{RFIDS} \cdot G_{Rx-Ant} \cdot PG \cdot PCG \cdot {N_{sum}}^2 \cdot PL. \quad (1)$$

The random noise is the summation of several possible sources, given as

$$N = [kTB \cdot G_{Rx} \cdot NF + N_{ext} \cdot G_{Rx} + N_{ADC}] \cdot N_{sum}. \quad (2)$$

To simplify the noise model and see the impact of excess noise compared to thermal noise, define a term

$$\gamma = [N_{ext} + (N_{ADC}/G_{Rx})]/kTB. \quad (3)$$

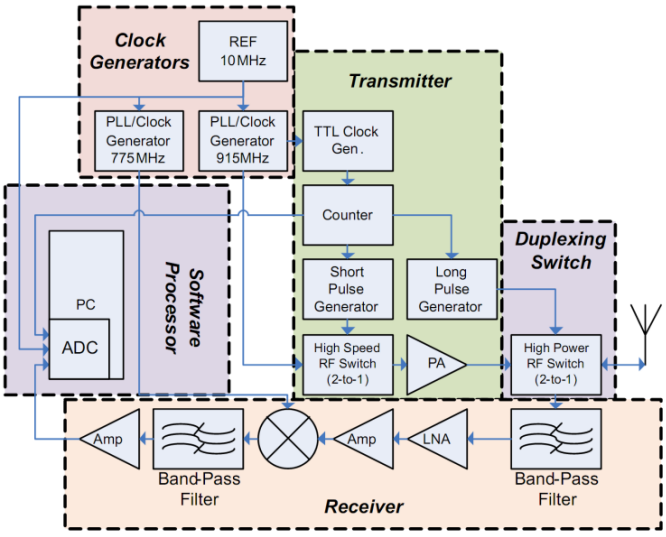


Fig. 4. A 915 MHz pulsed synchronous, RF transceiver block diagram developed for testing and analysis.

The noise can be rewritten as

$$N = [kTB \cdot G_{Rx} \cdot (NF + \gamma)] \cdot N_{sum} \quad (4)$$
$$= [kTB \cdot G_{Rx} \cdot NF^*] \cdot N_{sum}. \quad (5)$$

The signal to noise ratio at the ideal ADC, substituting, $PG = \tau \cdot B$ and regrouping, is given as

$$S/N = [EIRP/(NF^* \cdot kT/\tau) \cdot PCG \cdot PG] \cdot [(G_{RFIDS} \cdot G_{Rx-Ant})] \cdot [N_{sum} \cdot PL] \quad (6)$$

or

$$S/N = G_{PDL} \cdot PCG \cdot PG \cdot G_{sys} \cdot N_{sum} \cdot PL \quad (7)$$
$$= G_{loop} \cdot PL. \quad (8)$$

It is often convenient to write S/N in dB, which is given with dB terms in bold as

$$\mathbf{S/N_{dB}} = \mathbf{G_{PDL}} + \mathbf{PCG} + \mathbf{PG} + \mathbf{G_{sys}} + \mathbf{N_{sum}} + \mathbf{PL} = \mathbf{G_{loop}} + \mathbf{PL}. \quad (9)$$

The desired S/N establishes the required loop gain to overcome the path loss, in the presence of noise terms. Often, the MDS level is set at the noise power level and then

$$MDS = NF^* \cdot kT/\tau \quad (10)$$

and

$$G_{PDL} = EIRP/MDS \quad (11)$$

which represents a power gain.

As expected increasing the radiated power, reducing the thermal noise and noise figure, and increasing the encoded signal time length, all increase the S/N. For a passive sensor, such as a SAW device, the G_{RFIDS} is always less than unity. It is important to reduce the RFIDS loss as low as possible and to use as high gain antenna at the receiver as practical. Finally, coherent integration can, in theory, establish any desired S/N level, given a reduced RFIDS sampling rate. Interestingly, the receiver gain and the chip bandwidth do not impact the S/N, assuming a well designed receiver. The chip bandwidth does, however, impact the ability to code and detect, and therefore is an important variable for other system level considerations other than noise.

C. Range Prediction

For passive RFIDS, the range is given by Friis Equation as

$$Range = PL^{1/4}[v_{EM}/(4\pi f_o)] \quad (12)$$

where v_{EM} is the free space velocity.

A minimum S/N is determined for detection, and the maximum range, in meters, acheivable is obtained as

$$R_{max} = [(4\pi f_o)/v_{EM}]^{-1} \cdot [G_{PDL} \cdot PCG \cdot PG \cdot G_{sys} \cdot N_{sum}/(S/N_{min})]^{1/4} \quad (13)$$

or given in dB, dB terms in bold, as

$$\mathbf{R_{max-dB}} = (1/4) \cdot [\mathbf{G_{PDL}} + \mathbf{PCG} + \mathbf{PG} + \mathbf{G_{sys}} + \mathbf{N_{sum}} - \mathbf{S/N_{min}}] - 10 \cdot log[(4\pi f)/v_{EM}]. \quad (14)$$

V. System Design Considerations

The loop gain model developed predicts the maximum achievable range for the RFIDS synchronous correlator receiver system. The loop gain model components are separated into a series of terms that allow identification of important components for optimizing system range.

A. Path Loss

Assuming isotropic radiation, the received power from a transmitter decreases at 40 dB per decade as a function of increasing frequency and operational range. The thermal noise floor is constant with range but increases with bandwidth. The plots of Fig. 5 point to the use of the lowest operational frequency to minimize path loss and increase operational range. If target size is not a primary consideration, then operation at 250-500 MHz might be a good solution to increase operational range.

B. RFIDS-SAW and Antenna Target

However, to make the target (SAW plus antenna) small, it is desirable to work at relatively high frequencies. As operational frequencies increase, the SAW size typically decreases, the absolute operational bandwidth increases (for a given fractional bandwidth), the propagation losses increase, and the photolithographic resolution requirements increase. High velocity materials relax manufacturing process requirements, but constraints on fabrication and propagation loss have currently limited SAW operational frequencies to below 3 GHz. High coupling materials can provide low loss operation over wide bandwidths, but typically have large temperature coefficients of frequency. The SAW OFC devices developed thus far have used YZ-LiNbO3 since the material provides high coupling, broad bandwidths and minimal diffraction. A key parameter is the frequency dependent propagation loss constant, for YZ-LiNbO3 given as [10]

$$\alpha(f) = 0.19f + 0.88f^2 \quad (dB/\mu s), \quad (15)$$

with f in GHz.

The loss increases rapidly above 1 GHz, and it would be desirable to operate where the loss is not a dominant factor. Also, this loss term is optimistic, since thin films and other effects often increase expected device and material loss even greater with frequency. Antenna gain and achievable fractional bandwidths increase for a given volume for electrically small antennas (ESA) as frequency increases. The antenna gain and bandwidth can be estimated for an ESA given in Fig. 6, and shows that higher frequencies provide better performance with respect to both gain and bandwidth.

At 1 GHz, gain greater than 0 dB can be achieved for an antenna radius of an inch. Higher frequencies can further increase antenna gain and bandwidth, but the SAW device propagation loss counters the advantage, and the overall target performance will be optimized in the 850-1.5 GHz range, depending on other parameter factors.

Based on the previous arguments, the SAW OFC system currently under development has an operational frequency of 915 MHz; since many of the applications being considered have few or no restrictions on center frequencies and bandwidths. Regulatory constraints require consideration if operated in a restricted environment; placing additional constraints on the operating parameters. The 915 MHz frequency was chosen to balance the conflicting parameters of RFIDS small size, low loss, wide bandwidth, and fabrication process control. The devices at 915 MHz have a $\lambda/4$ line width of approximately 0.8 um on YZ-LiNbO3. The SAW OFC device results presented used 5 chips, with a processing gain (PG) of 25 and a device bandwidth of approximately 92 MHz. The surface mounted RFIDS was soldered to a PC board, open sleeve dipole antenna. The antenna, with the soldered SAW device, was 125 mm x 55 mm on a 32 mil FR4 copper clad PC board. The measured antenna gain, in a 50 ohm system, is approximately 1.8 dB and had a bandwidth of approximately 200 MHz [11]. The RFIDS had a loss of approximately 23 dB.

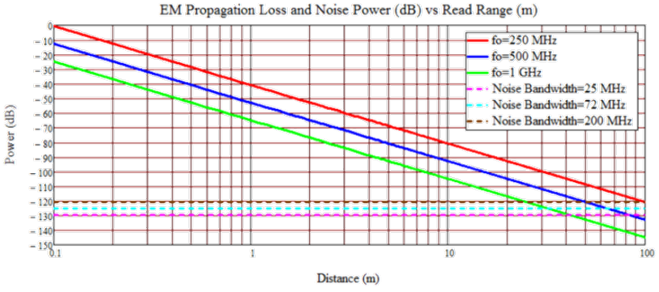


Fig. 5. Ratio of receive to transmit power, in dB, as a function of range, in meters, for a passive target for 3 differing operational frequencies, 0.25, 0.5 and 1 GHz - solid lines. The dotted traces are the thermal noise power levels (dBW) at 3 differing bandwidths, 25, 75, and 200 MHz.

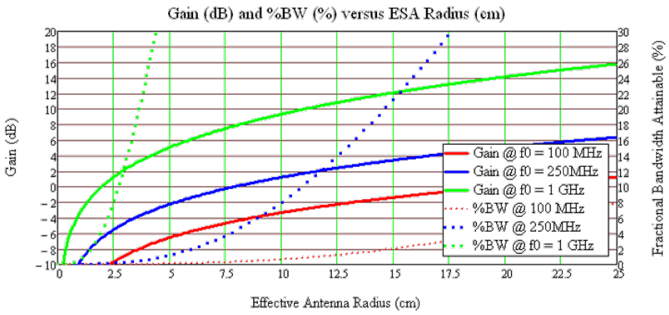


Fig. 6. Plots of the approximate gain and fractional bandwidth versus effective antenna radius for an electrically small antenna, from the analysis of Wheeler [11,12].

C. Transceiver

The following section presents results of measurements for two differing temperature SAW OFC systems at 915 MHz. The systems use an adaptive filter in a correlator based receiver to identify the SAW sensor and to extract temperature. Measurements are for a single sensor operation.

1) RF Pulsed Transceiver Parameters:

- f_0 = 915 MHz, $BW_{IF} > 100$ MHz
- Peak RMS Power = 30 dBm
- Pulse-Length = 10 ns
- PCG = -1 dB
- Antenna Gain = 9 dBi
- Bandwidth = 82 MHz
- G_{Rx} = 47 dB
- NF* = 18 dB
- N_{ext} = -22 dBm
- P_{FoM} = 30 - 1 - (-22) = 51 dB

The RF pulsed format is quite simple, but the radiated energy is small. For the pulsed system developed, a pulse length of 10 nsec gated a 915 MHz RF carrier. The peak power was 30 dBm and fed to a 9 dBi gain Yagi antenna, yielding an equivalent isotropic radiated power (EIRP) of 39 dBm. The radiated energy per pulse is approximately 79 nano-joules. Pulse compression gain in this case is approximately -1 dB, due to the bandwidth of the IF filter. Power figure of merit (P_{FoM}) can be expressed as sum of peak power to the antenna and PCG less noise floor of the system or the environment (N_{ext}).

Fig. 7 illustrates the range capabilities of an RF pulsed transceiver with the above parameters. Measurements were conducted of a single OFC sensor for identification of the code, as well as extraction of temperature. The brown line with markers is a plot of the fractional error in temperature to the range of the RFIDS. A false identification is indicated by a fractional error of 0.1 and occurs at a range of 3 meters. The red and black dot-dashed plots compare the predicted reflected RMS power measured at the ADC vs. range, and agree well until a large fractional error of temperature is measured.

2) Chirped Transceiver Parameters:

- Power to Antenna = 28 dBm
- Pulse-Length = 700 ns
- PCG = 17 dB
- Antenna Gain = 9 dBi
- Bandwidth = 78 MHz
- G_{Rx} = 45 dB
- NF* = 15 dB
- N_{ext} = -22 dBm
- P_{FoM} = 28 +17 - (-22) = 67 dB

Measurements were also made with a transceiver utilizing a chirped interrogation with the above parameters. The receiver IF bandwidth was 74 MHz, which reduced the apparent PG to 20. Pulse compression gain in this case is 15dB. The increase in P_{FoM} from pulsed to chirped system is 16dB. The range of the chirped transceiver is shown in Fig. 8, and the increase in range from 3 to 7.3 meters correlates well to the improvement

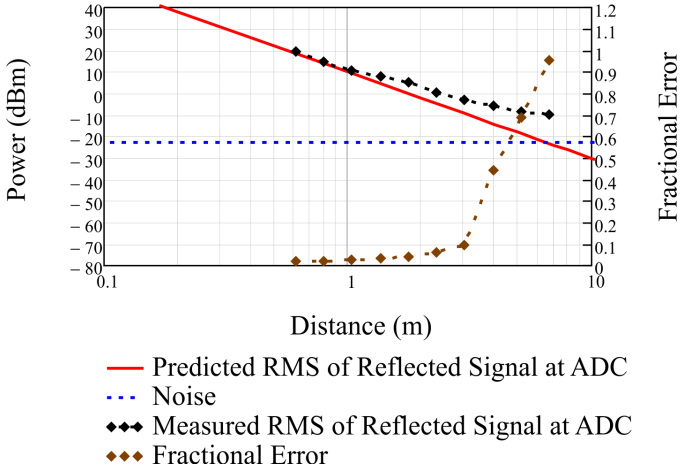


Fig. 7. Range measurement for RF pulsed system: RMS power reflected and received at ADC (red solid), noise floor (blue dashed), actual received reflection at ADC as measured using adaptive matched filter (black dot-dashed with markers), and fractional reading error (brown with markers)

in P_{FoM}. Similar to the pulsed system, the measured vs. predicted reflected RMS power strongly agree and validates the outlined range model given previously.

VI. Discussion

Major operational parameters of the correlator based interrogator have been established and a loop gain model has been developed for predictions of range. It has been shown that for certain key parameters, terms containing bandwidth of the system and the receivers gain drop out of the model. A power figure of merit has been defined. Two demonstrated systems implemented two different receiver architectures (RF pulse and chirp) with a maximum interrogatable range (for a single sensor) of 3m and 7.3m respectively, which correlates well with their P_{FoM}, which are 51 DB and 67 dB respectively.

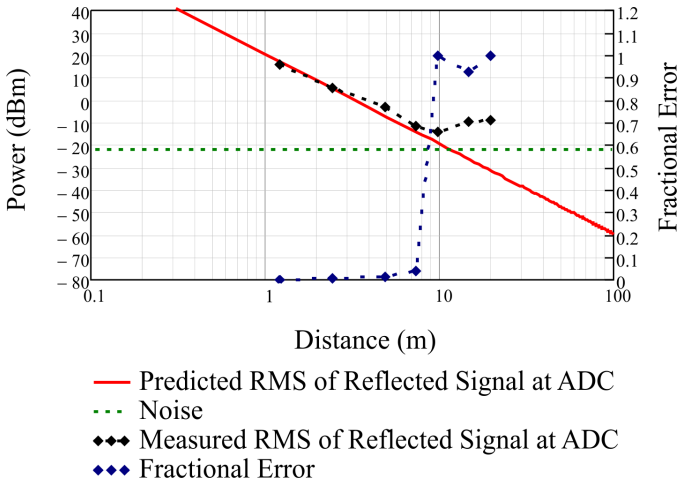


Fig. 8. Range measurement for chirped system: RMS power reflected and received at ADC (red solid), noise floor (blue dashed), actual received reflection at ADC as measured using adaptive matched filter (black dot-dashed with markers), and fractional reading error (brown with markers)

Based on practical improvements in the following system parameters:

- $\mathsf{NF}^* =$ -22 dB $\rightarrow$ -32 dB ($\Delta\mathsf{G} = 10$ dB)
- $\mathsf{G}_{\mathsf{RFIDS}} =$ -23 dB $\rightarrow$ -10 dB ($\Delta\mathsf{G} = 13$ dB)
- $\mathsf{PG}_{\mathsf{dB}} + \mathsf{N}_{\mathsf{sum-dB}} = 12$ dB $\rightarrow 22$ dB ($\Delta\mathsf{G} = 10$ dB)
- $\mathsf{PCG} = 17$ dB $\rightarrow 19$ dB ($\Delta\mathsf{G} = 2$ dB)
- Overall Gain Improvement: 35 dB

an increase in range of up to 80m is expected when using 10 synchronous averages or 40m with a single ping.

ACKNOWLEDGMENT

The authors wish to thank continuing support from NASA, and especially Dr. Robert Youngquist of Kennedy Space Center (KSC), for his encouragement. The authors are also grateful to Joni Richards, NASA-KSC, for her efforts in supporting a NASA Wireless SAW Workshop. The authors express thanks to all of the current and past students for their varied contributions to this research.

REFERENCES

[1] W. Bulst, G. Fischerauer, and L. Reindl, "State of the art in wireless sensing with surface acoustic waves," *Industrial Electronics, IEEE Transactions on*, vol. 48, no. 2, pp. 265–271, 2001.

[2] W. Wilson, D. Malocha, N. Kozlovski, D. Gallagher, B. Fisher, J. Pavlina, N. Saldanha, D. Puccio, and G. Atkinson, "Orthogonal frequency coded saw sensors for aerospace shm applications," *Sensors Journal, IEEE*, vol. 9, no. 11, pp. 1546–1556, 2009.

[3] W. Buff, F. Plath, O. Schmeckebier, M. Rusko, T. Vandahl, H. Luck, F. Moller, and D. Malocha, "Remote sensor system using passive saw sensors," in *Ultrasonics Symposium, 1994. Proceedings., 1994 IEEE*, vol. 1, 1994, pp. 585 –588 vol.1.

[4] G. Ostermayer, A. Pohl, C. Hausleitner, L. Reindl, and F. Seifert, "Cdma for wireless saw sensor applications," in *Spread Spectrum Techniques and Applications Proceedings, 1996., IEEE 4th International Symposium on*, vol. 2, 1996, pp. 795–799 vol.2.

[5] W. Seifert, W. Bulst, and R. C., "Mechanical sensors based on surface acoustic waves," *Sensors and Actuators A: Physical*, vol. 44, no. 3, pp. 231–239, 1994.

[6] F. Schmidt, O. Sczesny, C. Ruppel, and V. Magori, "Wireless interrogator system for saw-identification-marks and saw-sensor components," in *Frequency Control Symposium, 1996. 50th., Proceedings of the 1996 IEEE International.*, 1996, pp. 208–215.

[7] A. Pohl, G. Ostermayer, L. Reindl, and F. Seifert, "Spread spectrum techniques for wirelessly interrogable passive saw sensors," in *Spread Spectrum Techniques and Applications Proceedings, 1996., IEEE 4th International Symposium on*, vol. 2, 1996, pp. 730–734 vol.2.

[8] D. Puccio, D. C. Malocha, D. Gallagher, and J. Hines, "SAW sensors using orthogonal frequency coding," in *Proc. IEEE International Frequency Control Symposium*, 2004, pp. 307–310.

[9] D. Malocha, D. Puccio, and D. Gallagher, "Orthogonal frequency coding for saw device applications," in *Ultrasonics Symposium, 2004 IEEE*, vol. 2, 2004, pp. 1082–1085 Vol.2.

[10] A. Oliner and A. Slobodnik, "Materials and their influence on performance," in *Acoustic Surface Waves*, ser. Topics in Applied Physics. Springer Berlin / Heidelberg, 1978, vol. 24, pp. 225–303.

[11] M. Gallagher, B. Santos, and D. Malocha, "Wireless wideband saw sensor - antenna design," in *Frequency Control Symposium (FCS), 2010 IEEE International*, 2010, pp. 291–296.

Elasto-Acoustic Properties of $ReCa_4O(BO_3)_3$ (Re=La, Pr, Nd, Y, Gd) Piezoelectric Crystals

T. R. Shrout, F. P. Yu, S. J. Zhang
Materials Research Institute
Pennsylvania State University
University Park, PA, 16802
e-mail: trs16@psu.edu

Q. M. Wang
Dept. Mech. Eng. Mater. Sci., Uni. Pittsburgh, PA, 15260
Y. T. Fei, B. H. T. Chai
Crystal Photonics Inc.
5525 Benchmark Lane, Sanford, FL 32773

***Abstract*- Oxyborate single crystals, $ReCa_4O(BO_3)_3$ (Re=La, Pr, Nd, Y. Gd), have recently been demonstrated to offer superior dielectric, piezoelectric and electromechanical properties to quartz, langasite and other commonly used piezoelectric crystals. Of particular significance is the ability to function at temperatures > 1000°C, offering the potential for SAW, BAW and related sensors operational in harsh environments. In this work, the elasto-acoustic properties of oxyborate crystals were investigated as a function of rare earth cation. Crystallographic relationships were determined to be related to bond length and dependence of cation ordering, which was also reflected in the mechanical loss. Parameters related to high temperature operation, including electrical resistivity, dielectric loss, together with temperature dependent electromechanical and elastic properties were determined and discussed in relation to various applications.**

I. INTRODUCTION

Piezoelectric crystals, including quartz (α-SiO_2), lithium tetraborate ($Li_2B_4O_7$, LBO), gallium orthophosphate ($GaPO_4$), lithium niobate ($LiNbO_3$) and langasites ($La_3Ga_5SiO_{14}$, LGS), have been investigated for bulk acoustic wave (BAW) resonators, surface acoustic wave (SAW) devices and various sensors, including accelerometer(s) [1-2]. Recently, oxyborate crystals, $ReCa_4O(BO_3)_3$ (Re: rare-earth cation) have attracted considerable attention for various piezoelectric applications. As summarized in Table I, ReCOB borate crystals exhibit higher piezoelectric properties to that of quartz and $GaPO_4$, while offering the highest levels of electrical resistivity. Together with no temperature limiting phase transformation(s) prior to their melting points (Tm=~1500°C), borate crystals offer potential usage at temperatures exceeding 1000°C [3-8].

In this work, the elasto-acoustic properties of ReCOB crystals were investigated as a function of rare-earth cation. Crystal structure property relationships based on cation radii and structural disorder were established, leading to the identification of new borate crystals with enhanced piezoelectric properties. Parameters pertaining to high temperature operation, including electrical resistivity, dielectric loss, together with temperature independent electromechanical coupling and elastic properties were determined to establish the potential of borate crystals for piezoelectric resonators and sensors in harsh environments.

Table I Comparison of various high temperature piezoelectric crystals

Crystals	Symmetry	d_{eff} (pC/N)	Coupling k_{eff}	Resistivity (ρ) @500°C(Ω-cm)	Temperature usage range	Crystal growth method
Quartz[1]	32	2~3	~13.7%	~1×10^9	<350°C	Hydrothermal
LBO[2]	4mm	~20	~40%	/	<230°C	Bridgman
$GaPO_4$	32	~4	~18%	~1×10^{10}	<970°C	Hydrothermal
$LiNbO_3$[1,3]	3m	~21	~60%	~1×10^6	<600°C	Czochralski
LGS	32	6~7	~16%	~1×10^7	<800°C	Czochralski
Borates	m	5~7	~30%	~$\times10^{11}$	>1000°C	Czochralski

II. STURECTURE PROPERTY RELATIONSHIPS

$ReCa_4O(BO_3)_3$ crystals belongs to the point group '*m*'. For the determination of the complete set of elasto-acoustic related parameters, numerous sample geometries are required, as shown in Figure 1 [9].

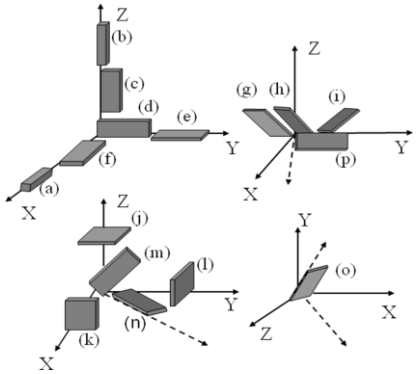


Figure 1. Sample orientations required for complete characterization of borate crystal with symmetry 'm': (*a*) *X* rod, (*b*) *Z* rod, (*c*) *XZ* bar, (*d*) *XY* bar, (*e*) *ZY* bar, (*f*) *ZX* bar, (*g*) (*XZ*w)45° bar, (*h*) (*XZ*w)30° bar, (*i*) (*XZ*w)-45° bar, (*j*) *Z* square plate, (*k*) *X* square plate, (*l*) *Y* square plate, (*m*) (*XY*t) 45° bar, (*n*) (*ZX*t)45° bar, (*o*) (*YZ*tw) 45°/-45° bar, (*p*) (*ZY*l)45° bar (after [9]).

This work supported by NSF under grant number ECCS 09-25586.

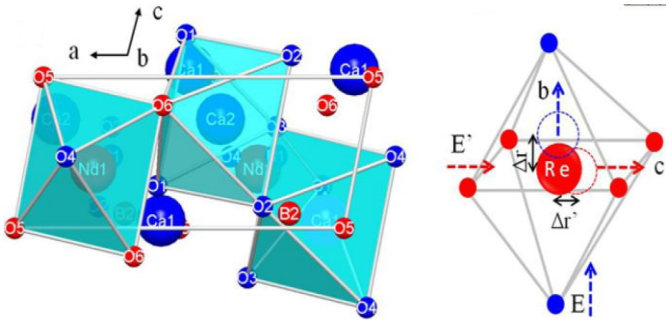


Figure 2 Crystal structure of NdCOB (left) and the schematic of Re- ion displacement movement (right) [9-10].

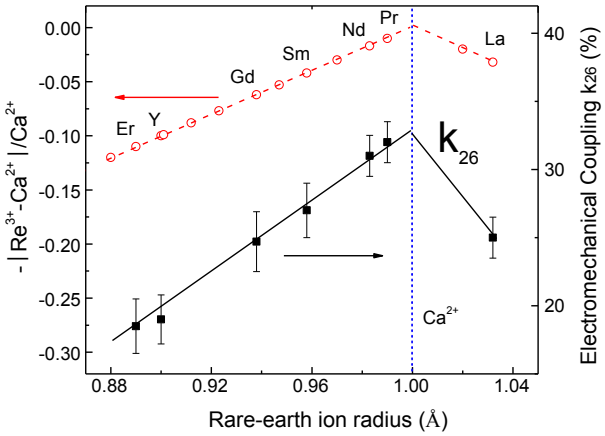


Figure 3 Variation of electromechanical coupling factor k_{26} as a function of Re-cation radius [10].

A schematic of the crystal structure showing the Re-cation is exhibited in Figure 2 [9-10]. Analogous to the langasite family ($R_3Ga_5SiO_{14}$, R= rare-earth elements) of crystals, where the electromechanical coupling factor was found to be related to the lattice parameter '*a*'[11], the elasto-acoustic properties of borates were found to be dependent on the Re-cation radii.

Figure 3 presents the piezoelectric shear properties as a function of Re-cation radii, revealing a maximum value in coupling k_{26} for ReCOB single crystals. The observed relationship was determined to be related to the Re-cation radii and the level of cation disorder. From observations, new borate crystals, including Nd and Pr borates were grown and optimum piezoelectric properties confirmed [10].

The level of cation disorder in the oxyborate crystals was reflected in the mechanical Q, as shown in Figure 4. Due to increased phonon scattering, the electrical resistivity was also found to decrease with increasing Re-cation radii, with the exception of LaCOB. The resistivity values for different ReCOB crystals at 1000°C were determined to be 1.0×10^6~1.0×10^8 Ohm·cm, as shown in Figure 5, where the small inset gives the extrapolated room temperature resistivities [8].

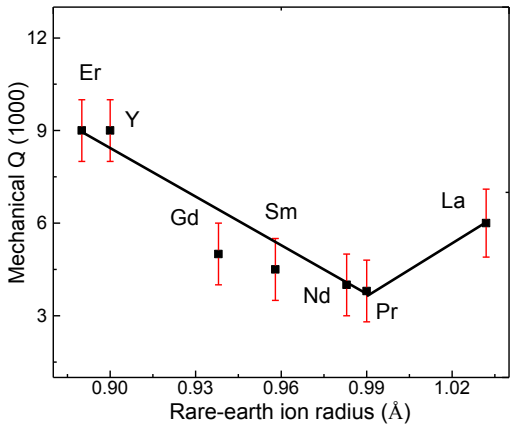


Figure 4 Variation of mechanical Q as a function of Re^{3+} radius in oxyborate crystals.

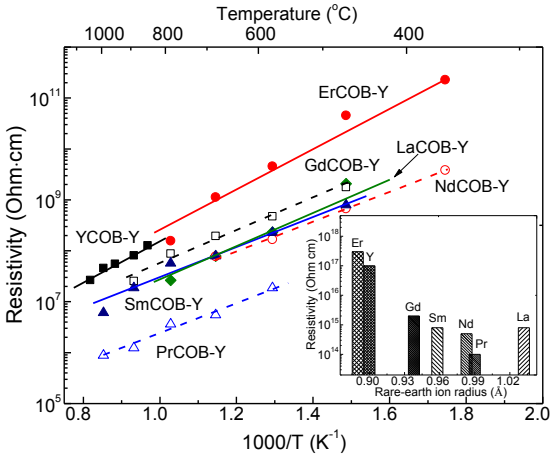


Figure 5 Variation of electrical resistivity for different ReCOB crystals along the Y-axis (after [8]).

III. HIGH TEMPERATURE PROPERTIES

Figure 6 shows the electromechanical coupling factor k_{26} for (ZXw)-θ cut YCOB as a function of temperature up to 950 °C. The largest k_{26} value was obtained in the (ZXw)-30°-cut sample, being 22% at room temperature, maintaining the same value till 950°C [7, 12]. Figure 7 presents the variation of resonance frequency f_r for (ZXw)θ (θ = –60°, 0°, 33°) YCOB crystal cuts. As can be seen, the resonance frequency shifts downward with increasing temperature, exhibiting linear temperature/ frequency characteristics over a wide temperature range of 20–950 °C. The highest first order TCF was found to be on the order of –91 ppm/K for the (ZXw)-33° cut. The impedance modulus for (ZXw)-0° cut samples as a function of temperature is shown in the inset of Fig. 7, indicating that the resonance and antiresonance peaks shift to lower frequencies with increasing temperature. Generally, sharp and narrow resonance frequency peaks reflect high

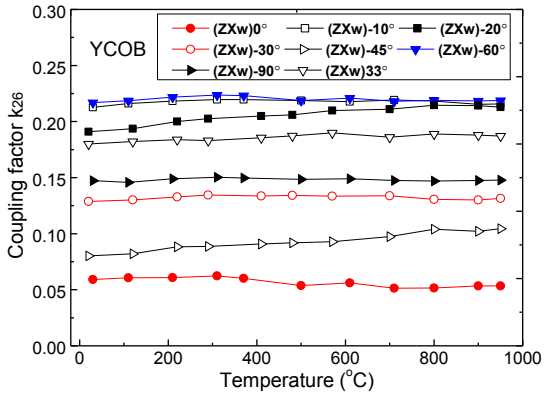


Figure 6 Variation of electromechanical coupling k_{26} for YCOB crystals as a function of temperature and crystal cut.

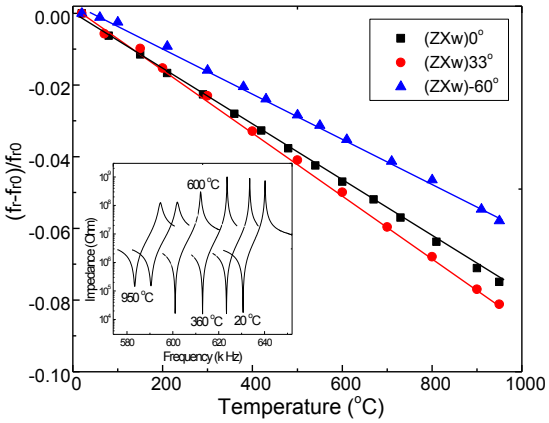


Figure 7 Frequency shift of YCOB crystals as a function of temperature and crystal cut. Inset is the impedance modulus for (ZXw)-0° cut YCOB as a function of temperature.

mechanical quality factors, Q. From the impedance modulus, the mechanical quality factors in air were calculated and found to be on the order of 16,000 at 20 °C and 2,000 at 950°C [7].

IV. HIGH TEMPERTURE APPLICATIONS

Potential high temperature applications of single crystals in the oxyborate family have been explored. As shown in Figure 7, the (ZXw)θ –cut YCOB crystals were found to possess linear frequency -temperature characteristics, with TCF on the order of -60 ~-90ppm/K, which are promising for temperature sensing applications at elevated temperatures. In contrast, bulk acoustic wave (BAW) resonators require temperature stability of the resonance frequency. In order to further investigate the potential of zero TCF crystal cut(s), NdCOB samples with different orientations as a function of rotation angle were prepared, as shown in the inset of Figure 8. The shear mode resonance frequency in the temperature range –140 to 200 °C for different crystal cuts (ZXw)θ (θ = –18°, –15°, 0°, 5°, 15°,

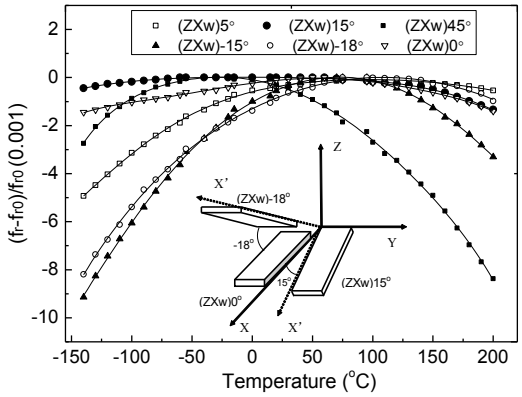


Figure 8 Resonance frequency shift of (ZXw)θ (θ = -18°, -15°, 0°, 5°, 15°, 45°) crystal cuts as a function of temperature for NdCOB crystals (after [13]).

Table II. Principle properties of NdCOB and PrCOB crystals, compared to commercial piezoelectric crystals used in frequency control devices [13-14].

Crystals	Cuts	Temperature range (°C)	TCF(1)*	TCF(2)*	d_{eff} (Pc/N)	k_{eff}
Quartz	AT-cut	-60~120	0	0	3~4	<8.8%
LBO	(YXl)51°	-30~110	0	-260	/	26%
$GaPO_4$	(YXl)-15.77°	-50~150	0	-10	4~5	13%
LGS	Y-cut	5~80	0	-58.9	~7	16%
CTGS	(YXl)-22.5°	-60~200	0	-39	~11	17%
PrCOB	(YXt)-1.5°	-140~200	0	-40	15.8	30.2%
NdCOB	(YXt) 15°	-140~200	0	-12	15.0	29.0%

* TCF (1): ppm/°C; TCF (2): ppb/°C²

45°) were measured, in which the (ZXw)15° crystal cut exhibited stable temperature frequency characteristics [13].

Table II lists the fitted first and second order TCF parameters for (YXt) θ-cut NdCOB and PrCOB single crystals, compared to commercial piezoelectric crystals. The piezoelectric d_{eff} and electromechanical k_{eff} are also listed in the table for comparison. The first order TCF reached zero for NdCOB crystals when the cut angle θ was 15°, while the crystal cut angle θ =-1.5° for PrCOB crystals was found to possess a zero TCF. Of particular significance is that both NdCOB and PrCOB single crystals have the highest piezoelectric coefficients d_{eff} ~ 15-16pC/N, with electromechanical coupling k_{eff} being on the order of 19-30%, higher when compared to other single crystals [13-14].

Other applications, such as vibration sensors (accelerometers) were also demonstrated with sensitivities being on the order of 2.5pC/g in a wide temperature range of 30-1000°C, achieved in a monolithic compression mode accelerometer, indicating high stability and reliability in the harsh environment [6].

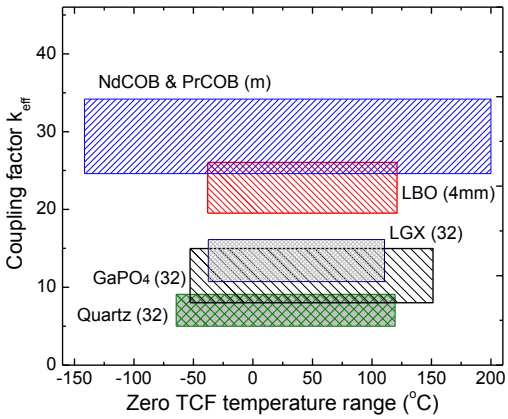


Figure 9 Comparison of NdCOB and PrCOB with commercial piezoelectric crystals used in BAW resonators (after [14]).

V. CONCLUSION

The electromechanical properties of piezoelectric crystals in the rare earth oxyborate family, including YCOB, GdCOB, ErCOB, SmCOB, NdCOB and PrCOB, were explored as a function of the rare earth cation radius, with maximum values observed for PrCOB crystals. The relationship between crystal structure and electromechanical properties for ReCOB crystals were believed to be related to the Re^{3+} ion radius and/or Re-O bond length and disorder distribution of Re^{3+} and Ca^{2+} cations. A consequence of increased disorder was the reduced electrical resistivity and mechanical Q, as confirmed in PrCOB crystals. YCOB crystals were found to possess linear characteristics of the frequency temperature behavior up to 1000°C, promising for temperature sensor applications, while NdCOB and PrCOB crystals exhibited nonlinear behavior in the temperature range of -140 ~200°C, leading to zero TCF characteristics for the potential in BAW resonator applications. Figure 9 summarizes the main properties of NdCOB and PrCOB crystals, compared to other commercial piezoelectric crystals with zero TCF and turnover temperature at 20°C. As shown, NdCOB and PrCOB offer significantly higher electromechanical coupling factors, together with the wide temperature usage range, make them candidates for BAW resonator applications [14].

REFERENCES

[1] R. C. Turner, P. A. Fuierer, R. E. Newnham and T. R. Shrout, "Materials for high temperature acoustic and vibration sensors: A review," *Applied Acoustics*, vol.41, pages 299-324, 1994.

[2] D. Damjanovic, "Materials for high temperature piezoelectric transducers," *Current Opinion in Solid State and Materials Science,* vol.3, pages 469-473, 1998.

[3] S. J. Zhang, E. Frantz, R. Xia, W. Everson, J. Randi, D. W. Snyder and T. R. Shrout, "Gadolinium calcium oxyborate piezoelectric single crystals for ultrahigh temperature (>1000°C) applications," *J. Applied Physics*, vol. 104, art no. 084103, 2008.

[4] S. J. Zhang, Y. T. Fei, B. H. T. Chai, E. Frantz, D. W. Snyder, X. N. Jiang and T. R. Shrout, "Characterization of piezoelectric single crystal $YCa_4O(BO_3)_3$ for high temperature applications," *Appl. Phys. Lett.,* vol. 92, art no. 202905, 2008.

[5] S. J. Zhang, T. R. Shrout, Y. T. Fei, B. H. T. Chai, E. Frantz and D. W. Snyder, "High temperature piezoelectric single crystal $ReCa_4O(BO_3)_3$ for sensor application," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control.* vol.55, pages 2703-8, 2008.

[6] S. J. Zhang, X. N. Jiang, M. Lapsley, P. Moses and T. R. Shrout, "Piezoelectric accelerometers for ultrahigh temperature application" *Appl. Phys. Lett.* Vol. 96, art no. 013506, 2010.

[7] F. P. Yu, S. J. Zhang, X. Zhao, D. R. Yuan, Q. M. Wang and T. R. Shrout, "High temperature piezoelectric properties of yttrium calcium oxyborate single crystals," *Phys. Status Stolid – RRL,* vol. 4, pages 103-5, 2010.

[8] F. P. Yu, S. J. Zhang, X. Zhao, D. R. Yuan, L. F. Qin, Q. M. Wang, T. R. Shrout, "Dielectric and electromechanical properties of rare earth calcium oxyborate piezoelectric crystals at high tempertures," *IEEE Trans. Ultrason. Ferroele. Frequency Control*., vol. 58, pages 868-73, 2011.

[9] F. P. Yu, S. J. Zhang, X. Zhao, D. R. Yuan, C. M. Wang and T. R. Shrout, "Characterization of neodymium calcium oxyborate piezoelectric crystals with monoclinic phase," *Crystal Growth & Design*, vol. 10, pages 1871-1877, 2010.

[10] F. P. Yu, S. J. Zhang, X. Zhao, S. Y. Guo, X. L. Duan, D. R. Yuan, T. R. Shrout, "Investigation on dielectric and piezoelectric properties of ReCOB (Re: rare earth elements) piezocrystals," *J. Phys. D.- Appl. Phys.* Vol. 44, art no. 135405, 2011.

[11] H. Takeda, H. Nakao, S. Izukawa, H. Shimizu, T. Nishida, S. Okamura, T. Shioaki, Growth and piezoelectric properties of $Re_3Ga_5SiO_{14}$ and $ReCa_4(BO_3)_3$ (Re: rare-earth elements) single crystals, *J. Alloys Compd*, vol. 408-412, pages 474-479, 2006.

[12] S. Zhang, F. Yu, R. Xia, Y. Fei, E. Frantz, X. Zhao, D. Yuan, B. Chai, D. Snyder, T. Shrout, "High temperature ReCOB piezocrystals: recent developments," *J. Crys. Growth,* vol. 318, 884-9 (2011).

[13] F. P. Yu, S. J. Zhang, D. R. Yuan, X. Zhao, Q. M. Wang and T. R. Shrout, "Crystal cuts with zero temperature coefficient of frequency in monoclinic piezoelectric crystal: NdCOB," *Phys. Status Solidi – RRL*, vol. 4, pages 185-187, 2010.

[14] F. P. Yu, S. J. Zhang, X. Zhao, D. R. Yuan, Q. M. Wang, T. R. Shrout, "6M Hz BAW resonators fabricated using new piezocrystals $PrCa_4O(BO_3)_3$ and $NdCa_4O(BO_3)_3$," *Phys. Status Sol.- RRL,* vol. 5, pages 47-9, 2011.

Modeling Phase Noise in Multifunction Subassemblies

Michael M. Driscoll*

Abstract – **Obtaining requisite phase noise performance in hardware containing multifunction circuitry requires accurate modeling of the phase noise characteristics of each signal path component, including both absolute (oscillator) and residual (non-oscillator) circuit contributors. This includes prediction of both static and vibration-induced phase noise. The model (usually in spreadsheet form) is usually refined as critical components are received and evaluated. Additive (KTBF) phase noise data can be reasonably estimated, based on device drive level and noise figure. However, accurate determination of component near-carrier (multiplicative) and vibration-induced noise usually needs to be determined via measurement. The model should also include the effects noise introduced by IC voltage regulators and discriminate between common vs. independent signal path residual noise contributors.**

I. INTRODUCTION

Accurate prediction of the spectral characteristics of signal generation and reception circuitry early in the hardware design phase is essential as it quickly identifies those circuit elements that are major noise contributors and performance limiters. Simplified, pre-design modeling is also helpful during proposal development in terms of identifying technology tradeoffs and making realistic assessments of attainable performance, risk, cost, etc. As oscillator noise reduction techniques and resonator characteristics have progressed, it can no longer be safely assumed that the phase noise contribution of the non-oscillator components can be largely ignored, especially in the presence of significant vibration and acoustic stress [1-5]. Modeling should include all possible or suspected major noise contributors in signal path circuitry and is time-consuming from both the standpoint of gathering data and entering it into the model. Nevertheless, it is considered a very necessary and worthwhile effort - one that is far less expensive in the long run, compared to the consequences fabricated hardware exceeding phase noise specification limits. The proposed model treats absolute (oscillator) phase noise separately from total residual (non-oscillator) signal path noise, with the two contributors added together at the hardware output signal ports.

*Written while employed by Sycom Services, a subsidiary of Ameriforce Craft Services, under contract to Northrop Grumman Electronic Systems, Baltimore, MD.

II. ABSOLUTE PHASE NOISE

Oscillator output signal phase noise can be inserted into the model as measured, vendor-advertised, or formulated data. Figure 1 shows a simplified block diagram of an oscillator incorporating a modular amplifier sustaining stage [5-6].

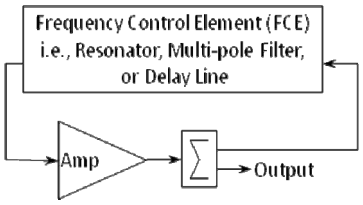


Figure 1. Feedback Oscillator Block Diagram

The data entered into the spreadsheet includes the amplifier open loop power input, noise figure, and multiplicative (usually 1/f) noise, and frequency control element (FCE) in-circuit group delay, and short-term frequency stability (self-noise). In a well-designed crystal or SAW oscillator, the FCE self-noise is usually the primary contributor to output signal near-carrier phase noise [6-9]. Amplifier multiplicative noise may be entered as a 1Hz intercept value or as measured data. Similarly, the white phase noise (floor) level may be computed from the noise figure and input drive level, with a "fudge factor" sometimes included to account for noise floor degradation with amplifier in-compression operation [10]. FCE self- noise may be entered, based on measured data or past experience. Sustaining stage measured data should be gathered at the actual amplifier operating point and powered from the intended voltage regulator, with the noise measurement system self noise accounted for. The output signal phase noise single sideband level can be calculated from these parameters as:

$$10 \log[((1/2\pi\tau f)^2+1)(K_1/f+K_2) + K_3/f^3] \qquad (1)$$

The numeric inside the outer brackets represents one half of the power spectral density (PSD) of the phase fluctuations. The $K_1/f+K_2$ terms represent the open loop PM noise of the non-resonator portion of the oscillator (the sustaining stage amplifier and other functional sub-circuits). K_1/f is the amplifier 1/f PM component. If the measured open loop 1/f PM noise at f=1Hz is –140dBc/Hz, then K_1=can be entered as 10^{-14}. K_2 is the white noise component. $K_2=10^{XX}/10$ where "XX" is the additive (i.e., KTBF) phase noise level in dBc/Hz = –177dBm/Hz-Pin(dB)+NF(dB). K_3 is a term that

accounts for the FCE self-noise that, in acoustic resonators and delay lines most often occurs as flicker-of-frequency noise in the oscillator. When measured in a phase bridge using an external signal generator, this type of instability induces flicker-of-phase (10dB/decade) noise onto the carrier signal. In oscillators comprising discrete component sustaining stages, the model must take into account the effect of the FCE frequency-selective impedance characteristics on the active device noise contributors [11-13].

IV. RESIDUAL PHASE NOISE

Like absolute noise, non-oscillator component, residual noise can be inserted into the model as measured, vendor-advertised, or formulated data. Usually the residual phase noise PSD contains a flicker-of-phase (K_1/f) component and a white noise (K_2) component as described in section II [1].

Measured data should be obtained with the UUT operated at the intended drive level and operated from the intended DC supply or voltage regulator. If/when the device phase shift is influenced by DC supply voltage fluctuations, the phase noise PSD characteristic may be degraded and deviate from the $K_1/f+K_2$ type characteristic, depending on the DC supply/voltage regular noise characteristics and internal device design (feedback control loop peaking, etc.)

V. COMMON vs. INDEPENDENT SIGNAL PATH NOISE

When more than one signal drives a signal path device like a mixer, accurate phase noise modeling requires that a distinction be made between phase noise that is common to the device input signals and that which is not. For example, consider the simplified block diagram shown in figure 2.

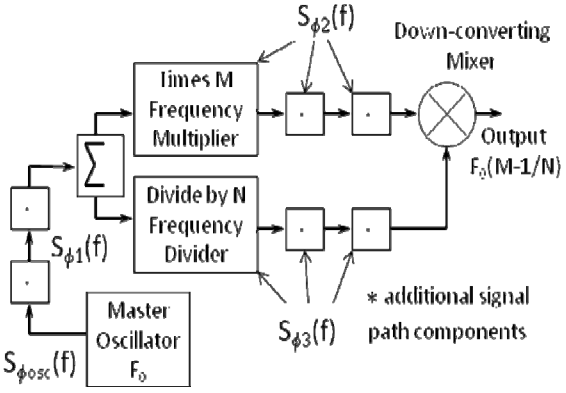


Figure 2. Common vs. Independent Signal Path Noise

$S_{\phi osc}(f)$ represents the oscillator *absolute* phase noise, $S_{\phi 1}(f)$ the *common* signal path *residual* phase noise, and $S_{\phi 2}(f)$, $S_{\phi 3}(f)$ represent the *independent* signal path *residual* phase noise, including the residual noise of the multiplier and divider, referenced to the device outputs. The phase noise PSD at the mixer output is:

$$(M\text{-}1/N)^2 S_{\phi osc}(f) + (M\text{-}1/N)^2 S_{\phi 1}(f) + S_{\phi 2}(f) + S_{\phi 3}(f) \qquad (5)$$

III. PHASE-LOCKED OSCILLATORS

The use of a phase-locked loop (PLL) is a widely used technique for imparting the near-carrier phase noise and long-term frequency stability of the PLL reference input signal to the locked oscillator (VCO), while retaining phase noise characteristics of the VCO at carrier offset frequencies in excess of the loop bandwidth [14-15]. In addition, the PLL is also used in indirect frequency synthesizers in order to provide frequency agility. Often, the entire PLL is available in integrated circuit chip form where the loop frequency/phase detector and frequency dividers are digital logic level devices. Figure 3 shows a simplified block diagram of a phase-lock loop.

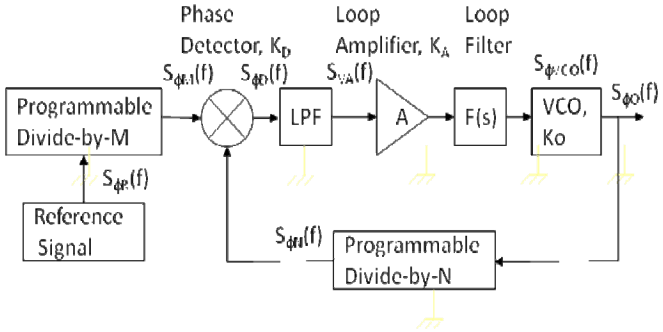


Figure 3. Phase-lock Loop Functional Elements

When the PLL is configured in discrete component form, classic formulas apply for modeling the output signal phase noise. In Figure 3, $S_{\phi R}(f)$, $S_{\phi VCO}(f)$, $S_{\phi M}(f)$, $S_{\phi N}(f)$, and $S_{VA}(f)/2K_D^2$ are the phase noise PSDs of the reference signal, VCO, frequency dividers (at the divider output), phase detector, and loop amplifier. The amplifier equivalent input noise voltage (volts^2/Hz) divided by two (50% AM, 50% PM) times the detector constant squared yields the equivalent amplifier phase noise, and includes input current noise flowing through the external input resistance. The open-loop gain is:

$$G_{OL} = (K_D)(K_A)(F(s))(K_O/s)(1/N) \qquad (2)$$

The phase noise PSD of the phase-locked VCO is:

$$(S_{\phi R}(f)/M^2 + S_{\phi M}(f) + S_{\phi N}(f) + S_{VN}(f)/2K_D^2) \text{ times } ((N)(G_{OL})/(1+G_{OL}))^2 \text{ plus } S_{\phi VCO}(f)/((1+G_{OL})^2) \qquad (3)$$

The phase noise of the components depicted in figure 3 can be modeled as described in sections II and III. From a phase noise standpoint, the optimum PLL bandwidth is usually at/near the frequency where the free-running VCO phase noise level is equal to that of all the non-VCO noise contributors times N^2. In phase lock loops incorporating logic level frequency dividers with high N divide ratios, the divider noise often limits output signal

phase noise performance. Manufacturers of PLL ICs often specify PLL phase noise performance in terms of a "normalized phase noise floor" or Figure of Merit (FOM). The FOM has been defined in the literature as:

$$L_{1Hz} = L_{floor} - 20\log(N) - 10\log(F_c) \quad (4)$$

L_{floor} is the "in-band" (inside the PLL bandwidth) single sideband, white phase noise level, N is the loop feedback division ratio, and F_c is the phase (detector) comparison frequency [16].

VI. VIBRATION-INDUCED NOISE

Vibration-induced phase noise is very often the dominant contributor to hardware output signal phase noise in the frequency offset region where vibration is impinging on sensitive components. Although this region often does not extend beyond 1-2kHz, the resulting phase noise degradation can extend to many times that frequency if when mechanical non-linearities exist. Examples of non-linearities include: loose particles from machining or galling, loose parts from assembly i.e. washers), lightly sprung electromechanical relays, and improperly secured electrical components, wires, cables, covers, and inner enclosed subassemblies. Even a slight amount of scraping between parts can cause a problem, as well as lack of flatness between large area enclosure surfaces. Mechanical resonances will cause amplification of the vibration level at the resonance frequency, where the vibration PSD level increase is proportional to the resonant Q factor squared [3].

Oscillator vibration induces changes in the oscillator loop phase shift and/or resonator resonant frequency that induces *frequency modulation* (FM) onto the oscillator output signal. Since the oscillator FCE represents the highest group delay component in the oscillator, it's vibration sensitivity is the dominant source of vibration-induced FM. Oscillator and oscillator resonator vibration sensitivity (Γ_f) is usually expressed on a fractional frequency basis, with $\Gamma_f = \delta f/f_0$ per g. The "f" subscript is used here to denote fractional frequency sensitivity to vibration. If the oscillator FCE is a delay line whose phase shift at the operating frequency varies according to a *phase* sensitivity to vibration (Γ_ϕ) in units of radians per g, the resultant oscillator fractional frequency vibration sensitivity can be expressed as:

$$\Gamma_f = 2\pi\tau\Gamma_\phi/f_0 \quad (6)$$

where τ equals the delay line delay or multipole filter group delay in seconds [3]. If the vibration is random, the vibration PSD is usually denoted as $S_g(f)$, and the vibration-induced, oscillator phase noise PSD is given by:

$$S_\phi(f) = S_g(f)((\Gamma_f f_0/f)^2) \quad (7)$$

where f is the carrier offset (vibration) frequency.

Unlike oscillator vibration, vibration in non-oscillator, signal path components induces *phase modulation* onto the carrier signal. Sensitive signal path components include narrow band (high group delay) filters and other narrowband circuits used for impedance matching, etc. Also included are broadband components like coaxial cables, optical fibers, and connectors [3, 17]. Cables are sensitive because they are usually free to move between points where they are secured and can scrape or hit against printed boards and enclosures and because their movement can induce changes in effective length and characteristic impedance, and can also induce motion in coaxial connectors – especially bayonet and "push-on" styles. When vibration results in signal PM, the use of Γ_ϕ is recommended for use in noise modeling, and the vibration-induced phase noise is:

$$S_\phi(f) = S_g(f)(\Gamma_\phi^2) \quad (8)$$

For coaxial cables, Γ_ϕ is usually varies with the vibration frequency. Filter vibration sensitivity can be represented in fractional (center) frequency or phase sensitivity form. If represented as a fractional frequency term, the vibration-induced phase noise is:

$$S_\phi(f) = S_g(f)((2\pi\tau\,\Gamma_f f_0)^2) \quad (8)$$

where τ is the filter pass-band group delay. Band-pass filter vibration-induced phase noise is subject to the filter transmission selectivity response, and vibrating filter resonators can induce a non-uniform effect on vibration-induced filter PM [18].

The most common method for reducing the spectrally-degrading effects of vibration is the use of passive, mechanical vibration isolators. These devices attenuate vibration levels at frequencies above their resonant frequency, amplify vibration at the resonant frequency, and exhibit unity transmission below the resonant frequency. The expression defining the transmission response of a mechanical isolation system is:

$$T=\mathrm{sqrt}((1+((f/(Qf_n))^2))/(((1-((f/f_n)^2))^2)+((f/(Qf_n))^2))) \quad (9)$$

where Q and f_n are the Q-factor and resonant frequency of the isolation system or vibration isolators. Figure 4 shows the theoretical vs. "realistic" isolator transmission response characteristics. For random vibration, the output vibration PSD is equal to the input vibration PSD times Γ^2. The dashed line in figure 4 is in recognition that most isolation systems do not provide attenuation in excess to 20 to 30dB. Figure 5, which shows the measured transmissibility of a low frequency isolation system, clearly demonstrates this effect [19].

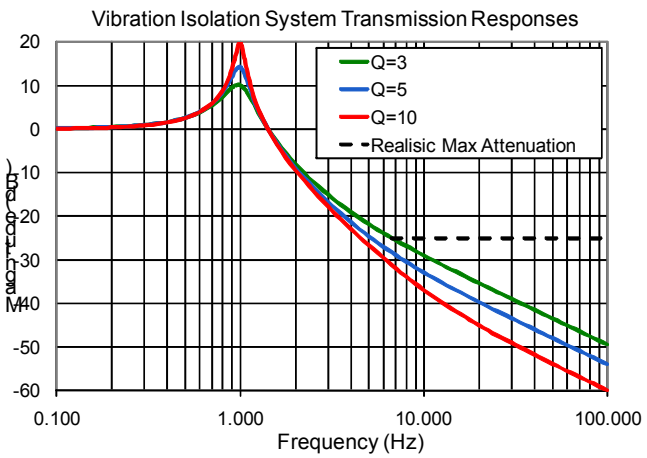


Figure 4. Theoretical Mechanical Isolation System Responses

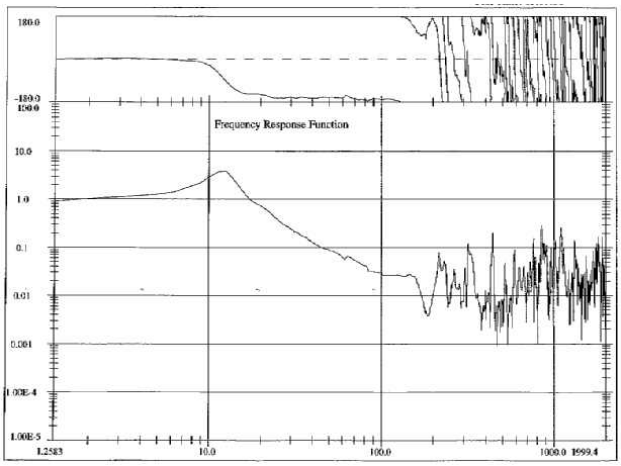


Figure 5. Measured Isolation System Response

VII. STATISTICAL VARIATION

When modeling both absolute and residual noise, variations in individual device or subassembly signal path noise levels (especially near-carrier phase noise) must be taken into account. Typical variability must usually be gleaned from measured data. Although some suppliers do provide typical vs. guaranteed levels, statistical data is usually not supplied. For example, short-term stability variations on the order of 10dB or more are typically observed for "identically" fabricated, VHF quartz crystal resonators [20]. Figure 6 shows the results of measurement of identical part number, tape and reel HBT amplifiers operated at UHF.

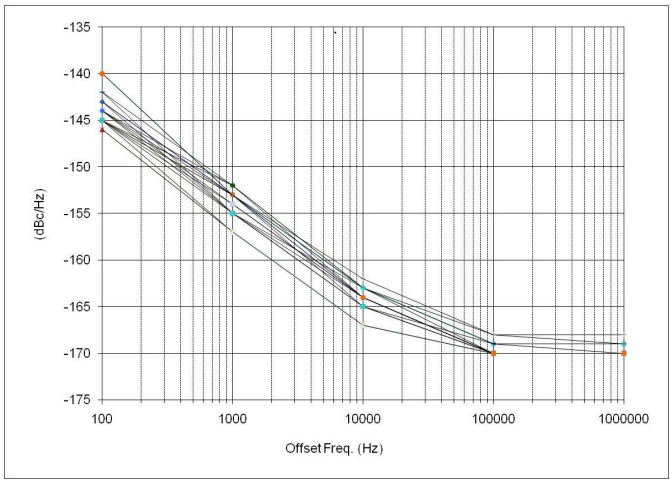


Figure 6. Measured PM Noise for Identical Part Number, UHF HBT Amplifiers

VIII. SPREADSHEET EXAMPLES

Examples of noise modeling spreadsheets are shown in the Appendix.

IX. CONCLUSIONS

Obtaining requisite phase noise performance in hardware containing multifunction circuitry requires accurate modeling of the phase noise characteristics of each signal path component.

Accurate prediction of spectral characteristics early in the hardware design phase quickly identifies those circuit elements that are major noise contributors and performance limiters.

It cannot be assumed that the phase noise contribution of non-oscillator components can be ignored, especially in the presence of significant vibration and acoustic stress.

Modeling is time-consuming, but is far less expensive in the long run, compared to the consequences fabricated hardware exceeding phase noise specification limits.

X. REFERNCES

[1] A. Hati, et. al., "Vibration-Induced PM and AM Noise in Microwave Components", IEEE Trans. on UFFC, Vol. 56, No. 10, Oct. 2009, pp. 2050-2059.

[2] M. Bloch, O. Mancini, T. McClelland, L. Terracciano, "Acceleration "G" compensated quartz crystal oscillators", IEEE Freq. Contr. Symp. 2009 Joint & 22nd European Freq. & Time Forum, April, 2009, pp. 175-180.

[3] J. B. Donovan and M. M. Driscoll, "Vibration-Induced Phase Noise in Signal Generation Hardware", Tutorial Session, 2009 Joint IEEE Freq. Contr. Symp. & 22nd European Freq. & Time Forum, April, 2009.

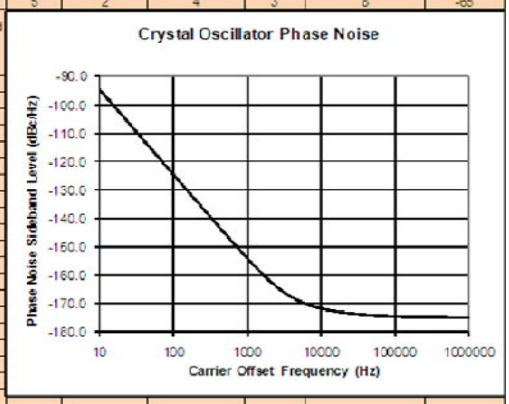
[4] C. McNeilage, et.al., "Review of Feedback and Feedforward Noise Reduction Techniques", Proc. 1998 IEEE Freq. Contr. Symp., May, 1998, pp. 146-55.

[5] Driscoll, M.M., "Reduction of Quartz Crystal Oscillator Flicker-of-Frequency and White Phase Noise (Floor)Levels and Acceleration Sensitivity via Use of Multiple Resonators" IEEE Trans. on UFFC, Vol. 40, No. 4, July, 1993, pp. 427-430.

[5] Driscoll, M.M., "Low Noise Crystal Oscillators Using 50-ohm, Modular Amplifier Sustaining Stages", Proc. 40th Freq. Contr. Symp., May 1986, pp. 329-335.
[6] T. E. Parker, "Extremely Low Phase Noise SAW Resonator Oscillator Design and Performance," Proc. 1987 IEEE Ultras. Symp., pp. 47-52.
[7] M. M. Driscoll, "Spectral Degradation in VHF Crystal Controlled Oscillators Due to Short-Term Instability in the Quartz Resonator", Proc. 1985 IEEE Ultras. Symp., pp. 340-345.
[8] Ferre-Pikal, E.S., et.al., "Experimental Studies on Flicker Noise in Quartz Crystal Resonators as a Function of Electrode Volume, Drive Current, Type of Quartz, and Fabrication Process", Proc. 50th IEEE Freq. Contr. Symp., May, 1996, pp. 844-851.
[9] M. M. Driscoll and W. P. Hanson, "Measured vs. Volume Model-Predicted Flicker of Frequency Instability in VHF Quartz Crystal Resonators", Proc. 1993 IEEE Freq. Contr. Symp., pp. 186-192.
[10] A. Hati, et. al., "Merits of PM Noise Measurements over Noise Figure: A Study at Microwave Frequencies", IEEE Trans. on UFFC, Vol. 53, No. 10, Oct. 2006, pp. 1889-1893.
[11] R. F. McDonald, P. J. Edwards, W. N. Cheung, "Shot Noise-Suppressed Operation of Bipolar Junction Transistors", Centre for Advanced Telecommunications and Quantum Electronics Research, PO Box 1, Belconnen, ACT 2616.
[12] V. H. Estrick, "A High Performance VHF Oscillator with Optimized Crystal Drive Power", Proc. 1993 IEEE Freq. Contr. Symp., pp. 209-215.
[13] Driscoll, M.M., "Low Noise Microwave Signal Generation" IEEE Trans. on UFFC, Vol. 35, No. 3, May, 1988, pp. 426-434.
[14] Young Wan Kim and Jae Du Yu, "Phase Noise Model of Single Loop Frequency Synthesizer", IEEE Trans. on Broadcasting, Vol. 54, No. 1, March 2008, pp. 112-117.
[15] Floyd M. Gardner, Phaselock Techniques, Second Edition, John Wiley and Sons, New York, 1979.
[16] P. White, "Understanding Phase Noise From Digital Components in PLL Frequency Synthesizers", Applied Radio Labs Design File DN006, Dec. 20, 2000.
[17] D. A. Howe, et.al., "Active Vibration-induced PM Noise Control in Optical Fibers: Preliminary Studies", Joint 2007 Freq. Contr. Symp. & 21st Europ. Freq. & Time Forum, pp. 552 – 556.
[18] R. C. Smythe, "Acceleration Effects in Crystal Filters: a Tutorial", IEEE TRANS. on UFFC,Vol. 39, No. *3,* May 1992, pp. 335-340.
[19] J. B. Donovan, "Maximum Attenuation of Vibration Isolation Systems", Northrop Grumman Technical Memo, Jan, 2003.
[20] Driscoll, M.M., "Spectral Degradation in VHF Crystal Controlled Oscillators Due to Short-Term Instability in the Quartz Resonator", Proc. 1985 IEEE Ultrasonics Symposium, 1985, pp. 340-345.

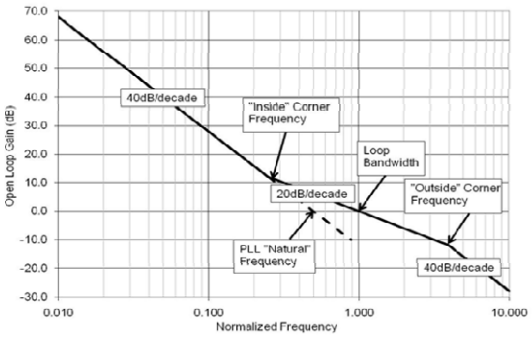
APPENDIX

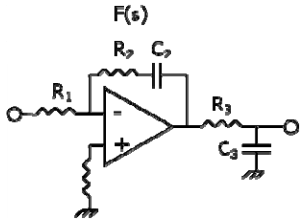
	A	B	C	D	E	F	G	H	I	J	K	L
1	VHF QUARTZ CRYSTAL OSCILLATOR PHASE NOISE MODEL											
2	Fo (MHz)	Crystal Loaded Q	Amp Steady State Po (dbm)	Amp PM Noise at 1Hz (dBc/Hz)	Amp Steady State Gain (dB)	Amp Steady State Pin (dBm)	Amp NF (dB)	Amp KTBF Noise Degradation (dB)	Power Divider I.L. (dB)	Crystal Mode Select Tuned Ckt I.L.	Crystal Network I.L. (dB)	Quartz Crystal Flicker-FM Noise at 1Hz
3	1.00E+08	6.00E+04	20	-135	15	5	5	2	4	3	8	-65
4	Freq	Amp S(ϕ)f)/2 (rad^2/Hz)	Amp PM Noise, dBc/Hz	Crystal 1/f FM Noise, Sϕ(f)/2 (rad^2/Hz)	Oscillator Closed Loop PM Noise Sϕ(f)/2 (rad^2/Hz)	Oscillator Closed Loop PM Noise (dBc/Hz)						
5	10	3.17E-15	-145.0	3.16E-10	3.4E-10	-94.7						
6	16	2.00E-15	-147.0	7.94E-11	8.5E-11	-100.7						
7	25	1.26E-15	-149.0	2.00E-11	2.1E-11	-106.7						
8	40	7.97E-16	-151.0	5.01E-12	5.4E-12	-112.7						
9	63	5.04E-16	-153.0	1.26E-12	1.3E-12	-118.7						
10	100	3.19E-16	-155.0	3.16E-13	3.4E-13	-124.7						
11	158	2.03E-16	-156.9	7.94E-14	8.5E-14	-130.7						
12	251	1.29E-16	-158.9	2.00E-14	2.2E-14	-136.7						
13	398	8.26E-17	-160.8	5.01E-15	5.5E-15	-142.6						
14	631	5.33E-17	-162.7	1.26E-15	1.4E-15	-148.5						
15	1000	3.48E-17	-164.6	3.16E-16	3.8E-16	-154.3						
16	1585	2.31E-17	-166.4	7.94E-17	1.1E-16	-159.6						
17	2512	1.58E-17	-168.0	2.00E-17	3.7E-17	-164.3						
18	3981	1.11E-17	-169.5	5.01E-18	1.7E-17	-167.8						
19	6310	8.17E-18	-170.9	1.26E-18	9.6E-18	-170.2						
20	10000	6.32E-18	-172.0	3.16E-19	6.7E-18	-171.7						
21	15849	5.16E-18	-172.9	7.94E-20	5.3E-18	-172.8						
22	25119	4.42E-18	-173.5	2.00E-20	4.4E-18	-173.5						
23	39811	3.96E-18	-174.0	5.01E-21	4.0E-18	-174.0						
24	63096	3.66E-18	-174.4	1.26E-21	3.7E-18	-174.4						
25	100000	3.48E-18	-174.6	3.16E-22	3.5E-18	-174.6						
26	158489	3.38E-18	-174.7	7.94E-23	3.4E-18	-174.7						
27	251189	3.29E-18	-174.8	2.00E-23	3.3E-18	-174.8						
28	398107	3.24E-18	-174.9	5.01E-24	3.2E-18	-174.9						
29	630957	3.21E-18	-174.9	1.26E-24	3.2E-18	-174.9						
30	1000000	3.19E-18	-175.0	3.16E-25	3.2E-18	-175.0						



cell B5=(10^(D$3/10))/A5+10^-((177+F$3-G$3-H$3)/10). cell C5=10*LOG(B5)
H3 is the additional noise floor degradation, if applicable, due to in-compression sustaining stage amplifier operation.
cell D5=(10^(L$3/10))/A5^3
cell E5=(B5*((A$3/(2*B$3*A5))^2+1)+D5). cell F5=10*LOG(E5)

Figure 7. Example of a Crystal Oscillator Noise Modeling Spreadsheet





"Inside" corner frequency $1/2\pi\tau_2 = 1/2\pi R_2C_2$ (open loop gain slope drops to 20dB/decade)

"natural" frequency $\omega_n = SQRT(K_0K_n/\tau_1)$, $f_n = \omega_n/2\pi$

"outside" corner frequency = $1/2\pi R_3C_3$, damping factor = $\zeta = \tau_2\omega_n/2$

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Freq. (Hz)	HF Input Phase Noise Level at HF (dBc/Hz)	PLL Detector, Divider, and Loop Amplifier Phase Noise, Referred to HF (dBc/Hz)	Total PLL Reference Noise, Referred to UHF (dBc/Hz)	UHF VCO Static Phase Noise Sideband Level (dBc/Hz)	"Inside" PLL Corner Frequency (Hz)	PLL Bandwidth (Hz)	"Ouside" PLL Corner Frequency (Hz)	Freq. (Hz)	Real Part of Open Loop Gain	Imaginary Part of Open Loop Gain	Open Loop Gain, GOL	Open Loop Gain, GOL (dB)
2	10.0	-128.5	-130.0	-90.1	-70.0	40.0	120.0	400.0	10.0	-4.8E+01	-1.1E+01	4.9E+01	33.9
3	11.2	-128.8	-130.5	-91.0	-71.5				11.2	-3.8E+01	-8.6E+00	4.0E+01	32.0
4	12.6	-130.6	-130.9	-91.8	-73.0				12.6	-3.1E+01	-6.6E+00	3.2E+01	30.0
5	14.1	-131.6	-131.4	-92.6	-74.5				14.1	-2.4E+01	-7.0E+00	2.5E+01	28.1

	N	O	P	Q	R	S	T	U
1	1/(1+GOL)	1/(1+GOL) (dB)	GOL/(1+GOL)	GOL/(1+GOL) (dB)	Frequency (Hz)	Contribution of PLL Reference, DetectorDivider, and Loop Amplifier (dBc/Hz)	Contribution of PLL VCO (dBc/Hz)	Estimated Phase Noise Sideband Level of Phaselocked VCO (dBc/Hz)
2	2.1E-02	-33.7	1.0E+00	0.2	10.0	-90.0	-103.7	-89.8
3	2.6E-02	-31.7	1.0E+00	0.2	11.2	-90.8	-103.2	-90.5
4	3.2E-02	-29.8	1.0E+00	0.3	12.6	-91.6	-102.8	-91.2
5	4.1E-02	-27.8	1.0E+00	0.3	14.1	-92.2	-102.3	-91.8

cell B2 = 10*LOG((10^-10.5)/A2^3+(10^-11)/A2^2+(10^-13)/A2+10^-15.8) polynomial fit
cell C2 = 10*LOG((10^-12)/A2+10^-15), cell D2 = 10*LOG(10^(B2/10)+10^(C2/10))+36 36dB=64=f_{vco}/f_{ref}
cell E2 = 10*LOG((10^-4)/A2^3+10^-15.9) flicker noise + KTBF noise
cell J2 = -(F$2*G$2*(H$2^2)+G$2*H$2*(I2^2))/((H$2*I2)^2+I2^4)
cell K2 = (F$2*G$2*H$2*I2-(H$2^2)*G$2*I2)/((H$2*I2)^2+I2^4), cell L2 = =SQRT(J2^2+K2^2)
cell M2 = =20*LOG(L2), cell N2 = 1/(SQRT((J2+1)^2+K2^2)), cell O2 = =20*LOG(N2)
cell P2 = =L2*N2, cell Q2 = 20*LOG(P2), cell S2 = D2+Q2, cell T2 = E2+O2
cell U2 = 10*LOG(10^(S2/10)+10^(T2/10))

Figure 8. An Example of a Phase-locked Oscillator Noise Modeling Spreadsheet

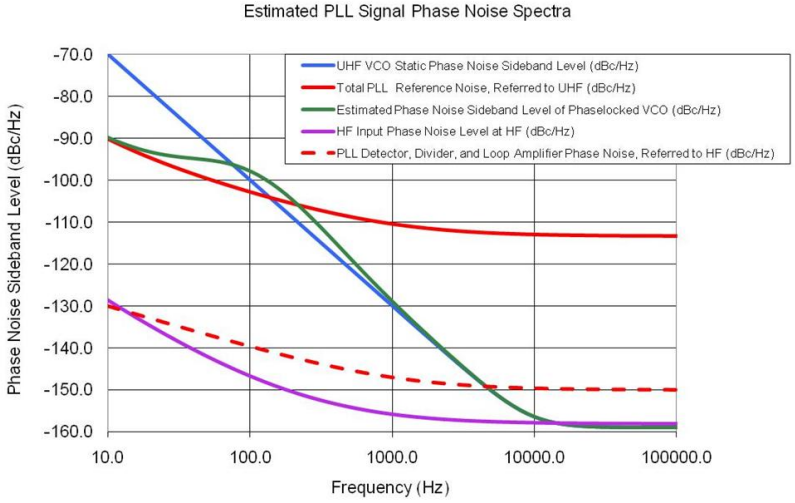
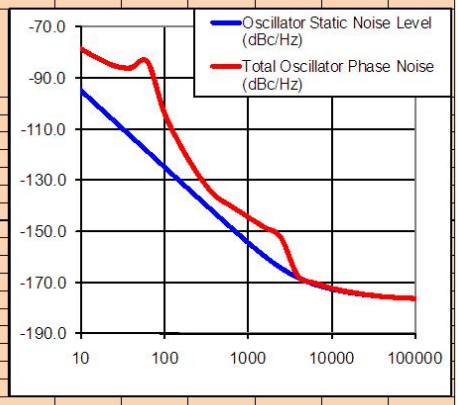


Figure 9. Typical PLL Modeling Results

	A	B	C	D	E	F	G
1	Effects of Vibration On Oscillator Phase Noise						
2	Freq	Carrier Frequency (Hz)	Oscillator Vibration Sensitivity (df/fo per g)	Vibration PSD at the Isolated Oscillator	Oscillator Static Noise Level (dBc/Hz)	Oscillator Vibration-Induced Phase Noise (dBc/Hz)	Total Oscillator Phase Noise (dBc/Hz)
3	10	1.0E+08	3.00E-10	3.17E-03	-94.7	-78.5	-78.4
4	16			3.46E-03	-100.7	-82.1	-82.0
5	25			4.39E-03	-106.7	-85.0	-85.0
6	40			9.04E-03	-112.7	-85.9	-85.9
7	63			3.99E-02	-118.7	-83.5	-83.5
8	100			1.06E-03	-124.7	-103.2	-103.2
9	158			1.19E-04	-130.7	-116.7	-116.5
10	251			2.29E-05	-136.7	-127.9	-127.3
11	398			7.50E-06	-142.6	-136.7	-135.7
12	631			7.50E-06	-148.5	-140.7	-140.1
13	1000			7.50E-06	-154.3	-144.7	-144.3
14	1585			7.50E-06	-159.7	-148.7	-148.4
15	2512			7.50E-06	-164.4	-152.7	-152.4
16	3981			1.00E-07	-168.1	-175.5	-167.4
17	6310			1.00E-07	-170.8	-179.5	-170.2
18	10000			1.00E-07	-172.6	-183.5	-172.2
19	15849			1.00E-07	-173.9	-187.5	-173.7
20	25119			1.00E-07	-174.8	-191.5	-174.8
21	39811			1.00E-07	-175.5	-195.5	-175.5
22	63096			1.00E-07	-176.0	-199.5	-176.0
23	100000			1.00E-07	-176.4	-203.5	-176.4



cells D3 - D23 copied from spreadsheet page with vibration input and vibration isolation system responses
cells E3 - E23 copied from spreadsheet page with oscillator static noise calculation
cell F3 = 10*LOG(((B$3*C$3/A3)^2)*D3/2)
cell G3 = 10*LOG(10^(E3/10)+10^(F3/10))

Figure 10. An Example of a Crystal Oscillator Static + Vibration-Induced Noise Model

	A	B	C	D	E	F	G	H	L	M	N	O
1	Frequency (Hz)	AR1 Phase Noise, $S_\phi(f)/2$ (rad^2/Hz)	AR1 Phase Noise, (dBc/Hz)	Z1 Output Phase Noise, $S_\phi(f)/2$ (rad^2/Hz)	Z1 Output Phase Noise, (dBc/Hz)	Cumulative Phase Noise (rad^2/Hz)	Cumulative Device Static Phase Noise (dBc/Hz)	Vibration Level (g^2/Hz)	BPF Vibration-Induced Phase Noise, $S_\phi(f)/2$ (rad^2/Hz)	BPF Vibration-Induced Phase Noise (dBc^2/Hz)	Coax Vibration-Induced Phase Noise, $S_\phi(f)/2$ (rad^2/Hz)	Coax Vibration-Induced Phase Noise (dBc^2/Hz)
2	10	3.17E-15	-144.99	1.27E-15	-148.97	1.39E-14	-138.56	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
3	15.85	2.00E-15	-146.99	8.00E-16	-150.97	8.81E-15	-140.55	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
4	25.12	1.27E-15	-148.98	5.05E-16	-152.97	5.57E-15	-142.54	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
5	39.81	8.01E-16	-150.97	3.19E-16	-154.96	3.52E-15	-144.53	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
6	63.10	5.07E-16	-152.95	2.02E-16	-156.95	2.23E-15	-146.51	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
7	100.00	3.23E-16	-154.91	1.28E-16	-158.93	1.42E-15	-148.48	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
8	158.49	2.06E-16	-156.86	8.14E-17	-160.89	9.05E-16	-150.43	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
9	251.19	1.32E-16	-158.79	5.19E-17	-162.84	5.81E-16	-152.36	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
10	398.11	8.57E-17	-160.67	3.34E-17	-164.77	3.76E-16	-154.24	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
11	630.96	5.64E-17	-162.49	2.16E-17	-166.65	2.47E-16	-156.07	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
12	1000.00	3.79E-17	-164.21	1.42E-17	-168.46	1.66E-16	-157.80	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
13	1584.89	2.63E-17	-165.81	9.57E-18	-170.19	1.15E-16	-159.41	3.00E-03	1.92E-15	-147.2	3.8E-16	-154.3
14	2511.89	1.89E-17	-167.24	6.63E-18	-171.79	8.22E-17	-160.85	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
15	3981.07	1.43E-17	-168.46	4.77E-18	-173.22	6.18E-17	-162.09	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
16	6309.57	1.13E-17	-169.46	3.60E-18	-174.44	4.89E-17	-163.11	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
17	10000.00	9.47E-18	-170.24	2.86E-18	-175.44	4.07E-17	-163.90	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
18	15848.93	8.30E-18	-170.81	2.39E-18	-176.22	3.56E-17	-164.48	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
19	25118.86	7.57E-18	-171.21	2.10E-18	-176.79	3.24E-17	-164.90	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
20	39810.72	7.10E-18	-171.49	1.91E-18	-177.19	3.03E-17	-165.18	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
21	63095.73	6.81E-18	-171.67	1.79E-18	-177.46	2.90E-17	-165.37	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
22	100000.00	6.63E-18	-171.79	1.72E-18	-177.65	2.82E-17	-165.49	1.00E-06	6.40E-19	-181.9	1.3E-19	-189.0
23	1/f PM Noise at 1Hz (dBc/Hz)	-135		-145					BPF Center Freq. (Hz)	3.00E+08		
24	Pin (dBm)	0		17					BPF Group Delay (sec)	3.00E-08		
25	Noise Figure (dB)	5		10					BPF Fractional Frequency Vibration Sensitivity (1/g)	2.00E-08		
26	Effective Gain (dB)	17		-10								
27	In-Compression? 0=No, 2=Yes	0		0							Coax Vibration Sensitivity (rad/g)	5.0E-07
28	Multiplication Factor	1		2								

Figure 11. An Example of Residual Static + Vibration-Induced Noise Modeling
Note the degrading effects of both vibration in the Band-Pass Filter and Coaxial Cable(s) by comparing the data in columns G, M, and O

Vibration Isolation of Acceleration Sensitive Devices

Joseph B. Donovan
Electronic Systems
Northrop Grumman Corporation
Linthicum, Maryland, USA

Abstract— **Significant vibration exists in many applications such as aircraft, ground vehicles, ships, and even in the laboratory due to cooling fans, footsteps, etc. The acceleration sensitivity of frequency control devices often requires that they be isolated from their vibration environment in order to maintain the required stability. Even the latest advances in acceleration compensation of oscillators require high-performance vibration isolation systems to reach their full potential. This paper presents an approach to the design of vibration isolation systems tailored to the needs of the package designer for an acceleration-sensitive device. The approach starts with the vibration environment, the maximum allowable device acceleration, and the space available and arrives at sufficient information to select candidate off-the-shelf isolators or specify custom ones. Limitations and trade-offs inherent in practical isolation systems are presented as well as common pitfalls awaiting the designer.**

I. INTRODUCTION

The designer responsible for packaging of an acceleration-sensitive device will likely find that the required stability cannot be achieved without taking additional measures. Reference [1] describes in detail how to perform the calculation of the maximum allowable level of vibration for the device whether its sensitivity is best characterized by fractional frequency sensitivity or phase sensitivity and whether the noise generated is broadband or discrete in nature. Broadband noise places an upper limit on the power spectral density of acceleration, S_g, in units of g^2/Hz, whereas discrete noise places an upper limit on the peak acceleration, G, in units of g or multiples of the acceleration of earth's gravity. Both S_g and G may vary with frequency offset from carrier, corresponding to the modulation, or vibration, frequency. For the purposes of this paper, only broadband noise will be considered. However, the approach is similar, and somewhat simpler, for discrete noise.

Once the allowable level of vibration has been determined, it must be compared to the vibration environment to determine whether additional measures need be taken. Knowledge of the applicable vibration environment may come from design specifications or it may be inferred from what is known about the intended end-use environment.

II. VIBRATION ENVIRONMENT

Example vibration levels associated with a few end-use environments are shown in Fig. 1. The appropriate level, combined with any influence of intervening structure, forms the excitation power spectral density, S_e. Thus, additional measures are required if $S_e > S_g$. The transmissibility, T, or ratio of output to input motion needed, is given by

$$T = \sqrt{\frac{S_g}{S_e}} \tag{1}$$

such that the attenuation required corresponds to $T < 1$ and the amplification permitted corresponds to $T > 1$.

III. ATTENUATION

If required, attenuation is most often achieved through vibration isolation. A vibration isolation system consists of several passive isolators, each having sufficiently low stiffness to produce the desired attenuation at frequencies of interest.

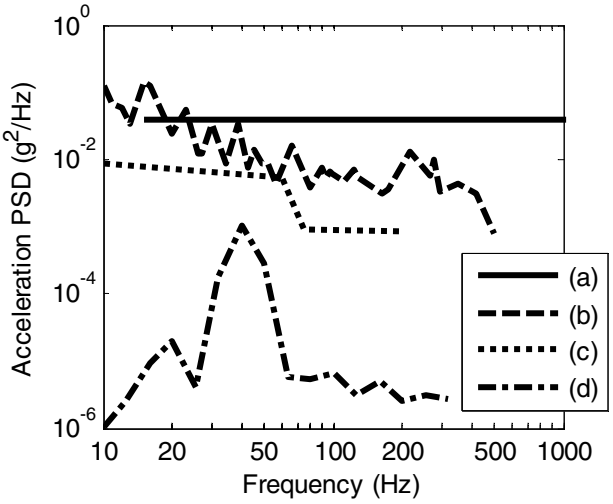


Figure 1. Example vibration levels for (a) aircraft[1], (b) wheeled vehicles[2], (c) ships[3], and (d) the laboratory of an orbiting space station[4].

[1] Jet at 0.9 Ma, sea level, far from engines [2] Method 514.6
[2] Composite wheeled vehicle, vertical [2] Method 514.6
[3] Unlisted ship types [3] Category V
[4] Non-isolated racks [4]

978-1-61284-110-6/11/$26.00 ©2011 IEEE

Note that a passive isolation system will amplify vibration at frequencies below those of primary interest. Thus it is critical to properly define the minimum frequency of interest or, at least, provide for a range of frequencies over which amplification is permitted.

Some of the unique aspects of isolation system design when its purpose is primarily to provide for adequate frequency or phase stability are

- Device fragility is usually not a driving consideration.
- Individual isolator size is of greater concern than the number of isolators.
- The isolation system design is just one of several inter-related inter-disciplinary systems which must be optimized simultaneously, such as electrical design and mechanical packaging.

The first aspect focuses the approach on the attenuation characteristics of the system with transmitted acceleration a lesser concern. The second aspect permits isolator size to be driven by displacement requirements and not the force due to gravity, etc. The third aspect places an emphasis on designation of isolator stiffness, size, and sway space as dependant variables more so than the design of isolation systems for other purposes.

The fundamental characteristics of a vibration isolation system are its natural frequency, f_n, and sharpness of resonance, Q. These characteristics can be determined by assuming the isolation system behaves as a single degree-of-freedom under-damped system. This is a simplification of realistic isolation systems such as the one diagramed in Fig. 2.

The transmissibility of such a system is given by [1]

$$T = \sqrt{\frac{1+\left(\frac{f}{f_n Q}\right)^2}{\left[1-\left(\frac{f}{f_n}\right)^2\right]^2+\left(\frac{f}{f_n Q}\right)^2}} \qquad (2)$$

A plot of (2) is shown in Fig. 3 for Q=3, a typical value for isolators.

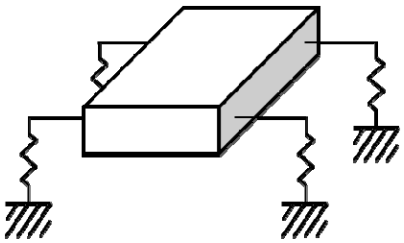


Figure 2. Diagram of an isolation system.

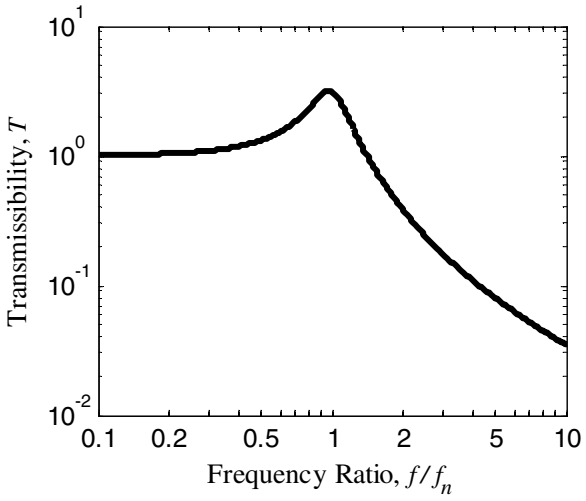


Figure 3. Transmissibility of a single degree-of-freedom system for Q=3.

Solution of (2) for the upper limit on natural frequency yields

$$f_{nul} = \sqrt{\frac{-b-\sqrt{b^2-4ac}}{2a}} \qquad (3)$$

where

$$a = T^2 - 1$$

$$b = \left(\frac{T^2-1}{Q^2} - 2T^2\right) f^2$$

$$c = T^2 f^4 .$$

Equation (3) is plotted in Fig. 4, again for Q=3, for several values of T.

Anticipation of T<0.1 is not realistic due to the inevitable presence of multiple degree-of-freedom behavior [1]. If such attenuation is required, multiple cascaded stages of isolation must be considered. If there are several offset frequencies for which attenuation is required, (3) can be used to determine the lowest value of f_{nul}. This is the upper limit on natural frequency under conditions where performance is required. However, as will be shown in the following section, there is a lower limit on natural frequency as well, dictated by the need for the isolators to survive.

Given f_{nul} and the isolated weight, w, the upper limit on system stiffness, k_{ul}, is obtained from [5]

$$k_{ul} = \frac{w}{g}\left(2\pi f_{nul}\right)^2 \qquad (4)$$

where g is the acceleration of earth's gravity.

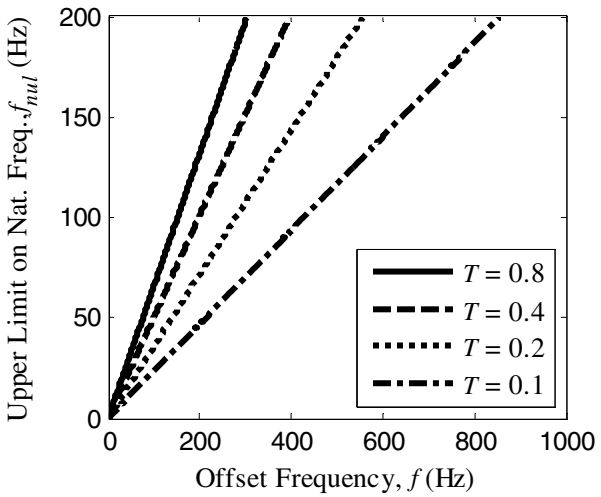


Figure 4. Relationship of the upper limit on natural frequency to the transmissibility required at a particular offset frequency for Q=3.

Given the isolated mass, the upper limit on system stiffness rules out combinations of isolator stiffness and number of isolators which result in a greater system stiffness. For this to be true, the center-of-gravity (CG) and center-of-elasticity (CE) must be coincident or very nearly so. The CG is, of course, the point on which the mass will balance without rotation. Conversely, the CE is the point on which a force can be applied resulting in equal deflection of all isolators. If the CG and CE are not coincident, so called "coupling" will occur and it becomes necessary to consider rotational response to translational excitation. Furthermore, device sensitivity to rotational acceleration is a relatively unexplored area. Substitution yields

$$k_{ul} = \frac{w}{g}(2\pi f)^2 \left[\frac{\left(2T^2 - \frac{T^2-1}{Q^2}\right) - \sqrt{\left(2T^2 - \frac{T^2-1}{Q^2}\right)^2 - 4T^2\left(T^2-1\right)}}{2\left(T^2-1\right)} \right] \quad (5)$$

IV. Producibility

It is unrealistic to expect an isolation system to exhibit the same natural frequency over time and from unit-to-unit. Thus, the upper limit on natural frequency, f_{nul}, will be accompanied by a lower limit on natural frequency, f_{nll}, separated from the upper limit to some extent. The separation may be defined by a factor F as follows

$$f_{nll} \le \frac{f_{nul}}{F} \quad (6)$$

The actual value of F is a result of at least

- Unit-to-unit variability due to manufacturing tolerances on natural frequency of up to 30% [6].
- Temperature sensitivity exhibited by elastomers resulting in natural frequencies varying by 15% to 60% depending on material and temperature range [7].
- Anisotropy, especially for wire rope, resulting in natural frequencies varying by 60% to 200% depending on the configuration and direction of displacement [8].
- Strain sensitivity exhibited by both elastomers and wire rope result in natural frequencies varying by 30% to 70% depending on strain amplitude [7].

The values of F in Table I are appropriate for rough sizing of isolators to help ensure that they are producible.

TABLE I. Practical Separation of Natural Frequency Limits

Type of Isolators	*F*
Off-the-shelf isolators	3.0
Custom isolators	2.0
Custom isolators operating under limited conditions	1.2

V. Isolation System Size

A. Sway Space

The lower limit on natural frequency typically defines the sway space required. Damage to the device could occur if the sway space were not adequate for the deflection, resulting in a collision. Assuming single degree-of-freedom behavior, the rms displacement of the isolation system is

$$d_{rms} = \frac{g}{(2\pi f_n)^2} G_{rms} \quad (7)$$

where G_{rms} is the rms acceleration response. In general

$$G_{rms} = \sqrt{\int \left(T^2 S_e\right) df} \ . \quad (8)$$

It can be shown that the rms acceleration, when S_e is reasonably flat (white), is [9]

$$G_{rms} \cong \sqrt{\frac{\pi}{2} f_n Q S_e\left(f_n\right)} \quad (9)$$

where $S_e(f_n)$ is the excitation PSD evaluated at the isolation system natural frequency. A common assumption is that the

sway space, s, needed to accommodate the maximum displacement is

$$s = 3d_{rms}\,. \tag{10}$$

This assumption is based on the fact that due to Gaussian random vibration, displacements which exceed three times the rms value occur less than 0.3% of the time and can be ignored. Thus, the sway space needed to equal the maximum displacement of the isolation system is

$$s = \frac{3gF^{\frac{3}{2}}}{(2\pi)^2 f_{nul}^{\frac{3}{2}}}\sqrt{\frac{\pi}{2}QS_e\left(\frac{f_{nul}}{F}\right)}\,. \tag{11}$$

It is apparent that the lower the natural frequency, the larger the sway space required. The mechanical package, then, must leave sufficient room to include the sway space plus any desired clearance.

B. Isolator Size

The lower limit on natural frequency also defines the size of the isolator required. An isolator not large enough to accommodate the maximum displacement would be damaged even if the sway space were available. Thus, any isolator considered must have its maximum permissible displacement compared to the maximum displacement calculated in the previous section. Even isolators large enough to accommodate the maximum displacement could be damaged if the displacements were repeatedly applied, eventually resulting in fatigue failure. A useful assumption is that the relationship between displacement and the number of cycles to failure at that displacement follows the so-called "power law" [10]

$$N = \left(\frac{d_0}{d_{CA}}\right)^{2m} \tag{12}$$

where d_{CA} is the constant amplitude displacement and N is the number of cycles to failure which generally occur at the rate dictated by the system natural frequency, f_n, such that

$$N = f_n t \tag{13}$$

where t is the time to failure in seconds. The constants d_0 and m depend on the make-up of the isolator.

The exponent, m, determines the relative damage done by displacements of differing amplitudes. Published values of m range from approximately 1.6 to 2.0 for elastomers [9, 11, 12] and similar values can be found for wire rope in bending [13, 14]. A value of m=1.6 is a conservative value useful for rough sizing.

The factor, d_0, is the value of d_{CA} that corresponds to $N=10^0=1$ when extrapolated from much larger numbers of cycles. However, this is not literally the displacement that would cause failure in one cycle. The maximum permissible displacement of the isolator can be quite different than d_0. Published values of d_0 indicate that it is roughly proportional to the height of the resilient portion of the isolator, h, for the most compact isolator types. That is

$$d_0 = Ch \tag{14}$$

where the constant of proportionality C, ranges from approximately 20 to 40 for elastomers [9, 11, 12]. Similar values can be found for wire rope in bending [13, 14]. A value of C=20 is a conservative value useful for rough sizing. A plot of the power law relationship is shown in Fig. 5.

The data that is available for determining these constants are based on tests of isolators conducted with constant amplitude displacement as indicated by the definition of d_{CA}. To meet our needs it is necessary to determine the constant amplitude displacement which is equivalent to the random displacement experienced by isolators subjected to broadband random vibration.

Miner's rule provides a means of accumulating the damage due to displacement at various amplitudes [9]. By accumulating the damage due to Gaussian distributed amplitudes using the value of m given above, it can easily be shown that a constant amplitude displacement of approximately two times the rms displacement produces equivalent damage. Thus

$$d_{CA} = 2d_{rms}\,. \tag{15}$$

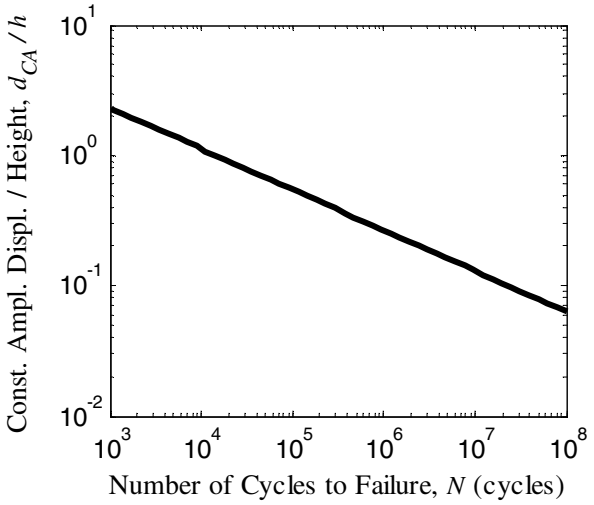


Figure 5. "Power law" relationship between displacement and number of cycles to failure.

Making substitutions and solving for *h* gives

$$h = \frac{2gt^{\frac{1}{2m}} F^{\left(\frac{3}{2}-\frac{1}{2m}\right)} \sqrt{\frac{\pi}{2} Q S_e \left(\frac{f_{nul}}{F}\right)}}{(2\pi)^2 C f_{nul}^{\left(\frac{3}{2}-\frac{1}{2m}\right)}}. \quad (16)$$

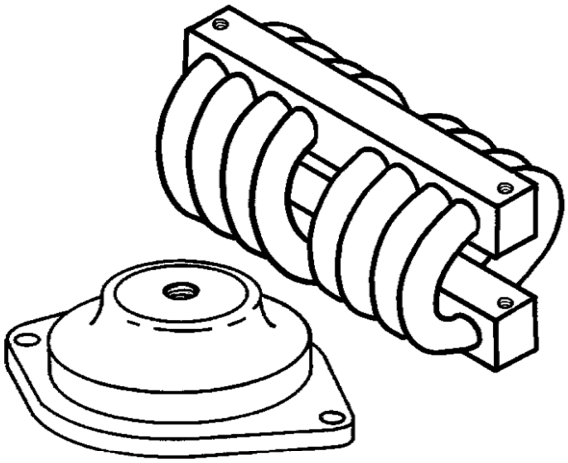
Using (16), it is possible to roughly estimate the size of each isolator required to meet the durability requirements.

VI. Isolator Selection

Examples of vibration isolators for use with acceleration sensitive devices are shown in Fig. 6. The typical resilient part is an elastomer or wire rope. Examples of elastomers used are natural rubber and silicone rubber. Wire rope is typically corrosion resistant steel. In either case, metal flanges or other features are incorporated which permit the isolator to be mechanically attached to both the isolated mass and the supporting structure.

A. Off-the-Shelf Isolators

Off-the-shelf isolator selection begins with the height determined from (16). From among all candidate isolators, only those whose resilient portion has a height in excess of h need be considered. Next, comparison of the required sway space from (11) with the maximum permissible deflection will rule out more. Finally, determine the number of isolators of each type which results in the total stiffness which is close to, but does not exceed, the stiffness obtained from (5). From the remaining candidates, choose the one that has an agreeable number of isolators (typically four, six, eight, or more) as well as maximizes design margin in other areas.



(a) (b)

Figure 6. Typical vibration isolators for used with acceleration sensitive devices. The resilient element is often (a) an elastomer or (b) wire rope.

B. Custom Isolators

There are often gaps between series of isolators which can be filled with custom designs. Isolators with resilient parts made of elastomer lend themselves to customization in more ways than wire rope. With either type of isolator the dimensions of the resilient part and of the attachment features can be customized. With elastomers, the material properties can be adjusted. Wire rope can be exercised and/or lubricated to adjust its properties.

VII. Common Pitfalls

There are several pitfalls that the package designer for an acceleration sensitive device is likely to encounter. First, there are a number of assumptions inherent in the method shown or ones like it.

- Although the sharpness of resonance, Q=3, has been assumed, it can vary between 2 and 10 depending on the resilient material, strain, temperature, and chemical exposure.
- The stiffness indicated here must be the dominant stiffness. "Shorting" stiffness in parallel with the isolators, such as from wires or cooling air ducts, must be much less.
- Applications which have significant constant acceleration or shock levels, have survival vibration levels which exceed the performance vibration levels, or must endure conditions which cause significant creep, will require more sway space than indicated here. However, it is possible to include bumpers which limit sway to that indicated here with the penalty of higher transmitted acceleration.
- Estimates of fatigue life are extremely rough. They do not consider the cumulative damage from vibration in multiple directions. At best they are a prediction for the median of the population. In order to account for scatter in the data, a safety factor 3 to 10 on life may be appropriate depending on the criticality of the application [14].

Second, for the reasons given above and others, it is absolutely necessary to confirm all conclusions formed by this approach with the isolator supplier. The intent of this method is to get the isolation system design in the right "ball-park" so that major re-design is unnecessary at later stages of the design process.

Finally, creative packaging engineers often fall prey to the temptation to design their own vibration isolators. This almost always ends in disappointment as the lessons learned by experienced designer/manufacturers cannot be leveraged. If even just one of the several inter-related design constraints is ignored, performance, durability, and/or producibility will suffer.

VIII. EXAMPLE

Consider the design of packaging for an acceleration sensitive device weighing 1 lb which has an upper limit on power spectral density of acceleration of 4×10^{-4} g^2/Hz at 300 Hz offset. The device is to be deployed on the aircraft of Fig. 1 (S_e=4×10^{-2} g^2/Hz) for 10 hours. The required transmissibility from (1) is

$$T = \sqrt{\frac{4\times10^{-4}}{4\times10^{-2}}} = 0.1 .$$

Assuming Q=3, (3) provides the upper limit on natural frequency f_{nul}=70 Hz and (4) limits the system stiffness to be no greater than

$$k_{ul} = \frac{1}{386}(2\pi 70)^2 = 501 \text{ lb/in.}$$

or 125 lb/in. each for a four isolator system.

Assuming off-the-shelf isolators are desired (F=3.0) the sway space from (11) is

$$s = \frac{3\cdot 386\cdot 3.0^{\frac{3}{2}}}{(2\pi)^2 70^{\frac{3}{2}}}\sqrt{\frac{\pi}{2}\cdot 3\cdot 0.04} = 0.11 \text{ in.}$$

Furthermore, from (16), in order to have sufficient durability, the resilient portion of the isolator must have a height at least

$$h = \frac{2\cdot 386\cdot (10\cdot 3600)^{\frac{1}{2\cdot 1.6}} 3.0^{\left(\frac{3}{2}-\frac{1}{2\cdot 1.6}\right)}\sqrt{\frac{\pi}{2}3\cdot 0.04}}{(2\pi)^2\cdot 20\cdot 70^{\left(\frac{3}{2}-\frac{1}{2\cdot 1.6}\right)}}$$

$$= 0.27 \text{ in.}$$

Examination of isolator catalogs, following the steps outlined in section VI of this paper, readily leads to candidate isolators employing either elastomers or wire rope. These candidates can form the basis for packaging concepts.

IX. CONCLUSION

This paper presents an approach to the design of vibration isolation systems tailored to the needs of the package designer for an acceleration-sensitive device. This approach is not rigorous or all-inclusive but it does establish realistic expectations for the size and performance of the isolation system early in the mechanical design process. Armed with only the required stability and vibration environment, one can perform rough sizing of the sway space and isolators. The resulting stiffness can be used to identify candidate off-the-shelf isolators or to begin the process of specifying custom ones. The approach respects limitations imposed by producibility while warning of common pitfalls.

ACKNOWLEDGMENT

The author is grateful to colleagues Michael M. Driscoll and Herman Rossman, both formerly of Northrop Grumman Corporation. Their mentorship was essential to this work and to the author's professional development.

REFERENCES

[1] J.B. Donovan and M.M. Driscoll, "Vibration-induced phase noise in signal generation hardware," EFTF-IFCS 2009 Joint Conference, Besancon, France, April 2009.

[2] MIL-STD-810G, *Environmental Engineering Considerations And Laboratory Tests*, October 2008.

[3] R.H. Chalmers, "Tailoring shipboard environmental specifications", Technical Document 2071, Naval Ocean Systems Center, San Diego, California, April 1991.

[4] R. DeLombard, K. Hrovat, E. Kelly, and K. McPherson, "Microgravity Environment on the International Space Station", NASA/TM—2004-213039, AIAA–2004–0125, National Aeronautics and Space Administration, Cleveland, Ohio, April 2004.

[5] W.T. Thompson, *Theory of Vibration with Applications*, 4th ed., Prentice Hall, New Jersey, 1993, p. 19.

[6] MIL-M-17185, *Mounts, Resilient; General Specificationsand Tests For (Shipboard Application)*, August 1956.

[7] Lord Corporation Aerospace Catalog.

[8] Enidine Wire Rope Technology Catalog.

[9] C.M. Harris, *Shock and Vibration Handbook*, 4th ed., McGraw-Hill, New York, 1996. p 11.10, p. 34.16, p. 35.18

[10] R.G. Lambert, "Accelerated test (time-compressed) methodologies for elastomeric isolators under random vibration", Proceedings of the Institute of Environmental Sciences, 1994.

[11] A.N. Gent, *Engineering With Rubber*, 2nd ed., HanserGardner, Cincinatti, Ohio, 2001 .

[12] P.B. Lindley, *Engineering Design with Natural Rubber*, Malaysian Rubber Producers Research Association, Hertford, England, 1978.

[13] J.E. Shigley and L.D. Mitchell, *Mechanical Engineering Design*, 4th ed., McGraw-Hill, New York, New York, 1983.

[14] C.R. Chaplin and A.E. Potts, *Wire Rope Offshore*, OTH 91 341, HSE Books, Sudbury, England, 1991, p. 167.

Limited Live-time Measurements of Frequency Spectra

D. A. Howe, N. Ashby, D. Lirette, A. Hati, and C. Nelson
Time and Frequency Division
National Institute of Standards and Technology
Boulder, CO 80305

***Abstract*—Frequency-difference-of-arrival (FDOA) can be used to monitor and track an emitter's location by observing Doppler frequency shifts of the carrier during short pulsed transmissions. This specific application needs to consider two frequency-stability levels: the short τ_{on}-average frequency measurements and the sampling time interval between measurements, denoted by τ_s. We show the advantage of using "dynamic" ThêoH for τ_{on} frequency measurements. Between τ_{on} transmissions while powered off, the emitter's reference-oscillator frequency will change due to power-start thermal variations, vibration, stress, and other frequency disturbances. If we regard long-term frequency instability as an uncertainty on an oscillator's expected or designated frequency, then the dominant error of frequency prediction is likely to be the error due to frequency drift and/or random walk FM, indicating the non-stationary behavior or disturbances. A two-sample variance is devised, called "Psi-variance," that has desirable statistical properties similar to those of the Allan variance. From this, we compute the power spectral density of frequency fluctuations, $S_y(f)$, from which phase noise, L(f), can be derived.**

I. Introduction and Summary

We characterize an oscillating signal for applications in which the oscillator itself is either only "on" or measured for short times (τ_{on}) between long, periodic or "stride" intervals (τ_s), during which the oscillator is off for $\tau_s - \tau_{on}$ seconds. A statistic called Psi-deviation, $\Psi(\tau_{on},\tau_s)$, is created that estimates the frequency-prediction error from the last frequency measurement. The variance of the prediction is $\Psi^2(\tau_{on},\tau_s)$, which is a time-averaged, two-sample variance and provides desirable properties similar to the Allan variance. The application is the tracking of an emitter, transmitting only for short periods of time, by means of frequency difference of arrival (FDOA). While the emitter's oscillating signal is on and transmitting (or being measured), receivers calculate tracks based on Doppler shifts, that in turn provide a navigation solution. On a two-dimensional plane, circular error probability (CEP) due to the emitter's oscillator noise is minimized, given low enough phase noise and frequency error during the interval τ_{on}. Because τ_{on} is short and we want to optimally characterize frequency stability, we compute ThêoH [1]. Using 10 sequential segments of τ_{on}, we display each segment's ThêoH in a "waterfall" or surface plot, to characterize the oscillator's turn-on transient. Dubbed "Dynamic ThêoH", or DThêoH, it allows us to compare the time taken by different oscillators to obtain consistent ThêoH. Oscillators having low enough cost, size, weight, and power appropriate for power cycling are OCXOs and TCXOs or oven and temperature compensated MEMs oscillators, respectively. Section II details the measurement setup and an example of DThêoH for an OCXO. Section III motivates and develops frequency prediction error $\Psi(\tau_{on},\tau_s)$. Section IV derives the relationship between Ψ-variance in the time domain and spectral density $S_y(f)$. Sections V and VI show time-domain and frequency-domain measurements, respectively, comparing an OCXO and a TCXO in limited live-time operation.

II. Dynamic ThêoH while Oscillator is On

We want to capture oscillator turn-on frequency stability and establish dynamic ADEV as a useful format [2]. ThêoH is preferred rather than ADEV since an oscillator is on for only a short time compared to its off-time. ThêoH is a hybrid (hence the "H") of ADEV for short-term averaging times plotted with a bias-removed version of Thêo1, called ThêoBR [3] for long-term. The Thêo portion characterizes to 75 % of a data run, whereas straight ADEV characterizes to only 20 % [1]. To illustrate, ThêoH plots are generated from measurements of oscillators powered on for a data run of only 3 s. Fig. 1 shows the measurement setup. Dynamic ThêoH is a waterfall graph where each 3 s run is parsed into ten sequential time segments, then ThêoH is computed for each segment and displayed as another waterfall plot, as shown in Fig. 2.

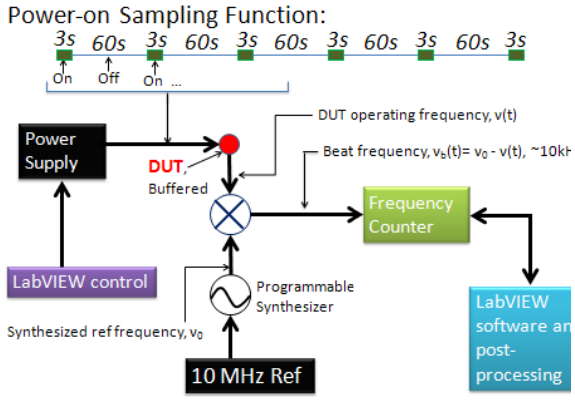


Figure 1. The device-under-test (DUT) is a temperature-compensated or oven-controlled oscillator with quartz or MEMs

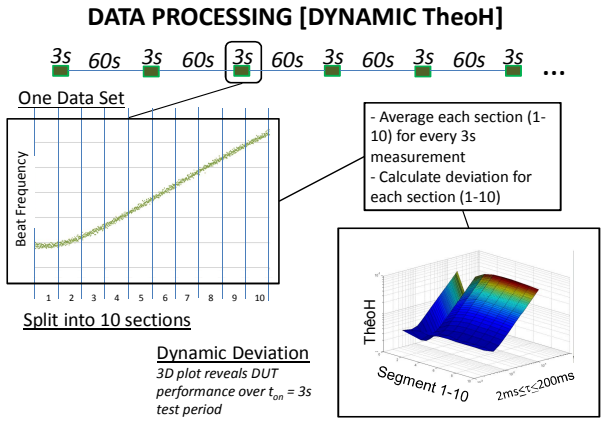


Figure 2. Dynamic ThêoH plot. Surface smoothness is a general measure that the test oscillator has attained steady-state operation after having been powered on.

Such plots provide a quick assessment of how long, and to what degree, it takes an oscillator to settle down to a consistent stability after being turned on.

We will use an example in which the emitter is repeatedly turned on for τ_{on} = 3 s every τ_s = 60 s. In addition to the frequency that is traced during start-up, discussed above, an important criteria is the start-up frequency reproducibility and its characterization, described next.

III. Frequency Prediction Error for Multiples of τ_S

We wish to estimate an oscillator's frequency at its next turn-on. While there are any number of different ways to make this estimate based on a history of actual measurements [4], we construct a two-sample frequency prediction to mimic the desirable properties of noise identification, convergence, convenience, and acceptance provided by the two-sample standard variance, better known as the Allan variance, and its square root, ADEV [5]. The two-sample, no dead-time Allan variance has widely accepted statistical properties, however, limited-live applications have substantial dead-time.

Fractional-frequency error $y_{on}(t)$ and its prediction at $y_{on}(t+\tau_s)$ is based on the reasonable assumption that (or one expects that) any given manufacturer wants $y_{on}(t+\tau_s)$ to be the same value as measured values of $y_{on}(t)$. If we write $y_{on}(t+\tau_s) = (1+\varepsilon)y_{on}(t)$, where ε is a random variable, we also expect average $\frac{1}{N}\sum_{n=1}^{N} y_{on}(t-n\tau_s)$ to be dependent on N, i.e. nonstationary, without a central limit, rendering this average of little or no practical use in the estimate of the frequency error designated as $\hat{y}_{on}(t+\tau_s)$. The efficient predictor in the presence of random walk noise is $\hat{y}_{on}(t+\tau_s) = y_{on}(t)$, the last measured value of y_{on}. The variance of this expectation can be written as a first difference

$$\Psi^2(\tau_{on},\tau_s) = \left\langle \left[\hat{y}_{on}(t+\tau_s) - y_{on}(t)\right]^2 \right\rangle, \quad (1)$$

where $\langle\cdot\rangle$ denotes an ensemble average. Like AVAR, (1) is the variance of an increment that converges in the limit.

DEFINITION: Samples of the fractional frequency-error function *y(t)* occur at a rate f_0 having an interval $\tau_0 = \frac{1}{f_0}$ (setup shown in Fig. 1). Given a sequence of fractional frequency errors $\{y_n : n=1,\ldots,M\}$ with a sampling period between adjacent measurements given by τ_0, we define the $m\tau_0$-average fractional-frequency deviate as

$$^{m}\bar{y}(t) \equiv \frac{1}{m}\sum_{j=0}^{m-1} y_{n-j},$$

where $y_n = y(t)$ with $n = t/\tau_0$ starting from a designated origin $t_0 = 0$. We also define psi-variance from the space of all possible two-sample increments as:

$$\Psi_y^2(\tau_{on},\tau_s) \equiv \left\langle \left[{}^{\tau_{on}}\bar{y}(t) - {}^{\tau_{on}}\bar{y}(t-\tau_s)\right]^2 \right\rangle \quad (2)$$

where $\langle\cdot\rangle$ denotes an ensemble average and ${}^{\tau_{on}}\bar{y}(t)$ is the mean frequency over duration $\tau_{on}=m\tau_0$. Fig. 1, top, shows the sampling function associated with $\Psi^2(\tau_{on}, \tau_s)$ acting on $\{y_n\}$, where τ_{on} is called the averaging or live interval and τ_s - τ_{on} is the oscillator's dead time. Note that $\Psi^2(\tau_{on}, \tau_s)$ becomes twice the two-point standard (Allan) variance $\sigma_y^2(\tau_s)$ if $\tau_{on} = \tau_s$.

IV. Relationship of $S_y(f)$ to Ψ–variance

For computing the usual power spectrum, we start with Parceval's theorem:

$$\Psi^2(\tau_{on},\tau_s) = 4\int_{\frac{1}{2\tau_s}}^{f_h} |H(f)|^2 S_y(f)\,df \quad (3)$$

where *H(f)* is the frequency-domain response of the time-domain sampling function of $\Psi^2(\tau_{on}, \tau_s)$ shown at the top of Fig. 1. $S_y(f)$ of the emitter is multiplied by the FT squared of the sampling function to obtain $|H(f)|^2$. We obtain:

$$\Psi^2(r,\tau_{on},\tau_s) = 4\int_{\frac{1}{2\tau_s}}^{f_h} \frac{S_y(f)\,df}{\pi^2\tau_{on}^{\,2} f^2}\left(\sin(\pi\tau_{on}f)\sin(\pi r\tau_s f)\right)^2 \quad (4)$$

where r = t/τ_s (starting from origin t_0) is a counting index r=1,2,3… representing the r^{th} data run of 3 s duration.

The f-domain response function $|H(f)|^2$ is shown in Fig. 3. This response is +20dB/decade like the Allan variance for low frequencies up to the peak at $f\tau_s$=½. There is insufficient roll off above this peak, so white and flicker of phase noise types will cause the level of $\Psi^2(\tau_{on},\tau_s)$ to depend on f_h in (4). This kind of dependence, though, is not a concern as of yet, since DUT random walk FM (and drift) are likely to dominate limited-live applications, as discussed in Section III, while white FM can occur from the measurement system at short- term.

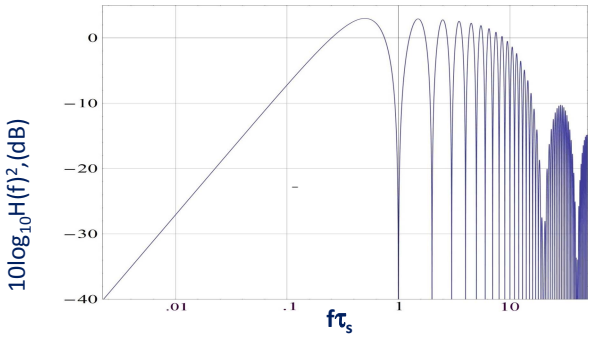


Figure 3. Frequency response $|H(f)|^2$ of $\Psi^2(\tau_{on},\tau_s)$ in (3), $\tau_{on} = (\tau_s)/20$, r=1.

Table 1 compares the transform to frequency spectrum $S_y(f)$ of $\frac{1}{2}\Psi^2(\tau_{on}, \tau_s)$ and $\sigma_y^2(\tau_s)$. We use the scaling factor "1/2" in $\frac{1}{2}\Psi^2(\tau_{on}, \tau_s)$ to normalize its result to equal $\sigma_y^2(\tau_s)$ if $\tau_{on}=\tau_s$, i.e., zero dead-time. Furthermore, the zero-dead time Allan and $\frac{1}{2}\Psi^2(\tau_{on}, \tau_s)$ respond identically to white FM noise having equal frequency-spectral coefficient h_o [6]. Flicker noise is given in terms of $\sigma_y^2(\tau_s)$ to simplify the formula. The Table evidences the bias on $\sigma_y^2(\tau_s)$ due to limited-live operation of the DUT.

We now observe the intrinsic level of random walk or drift that properly characterizes the DUT as τ approaches its maximum around 2 s, using ThêoH. Since fast-frequency measurements mask or are not sensitive to DUT-based PM-noise types that would appear as "super white" FM-noise in y(t) raw data, $\frac{1}{2}\Psi^2(\tau_{on}, \tau_s)$ is never biased by this noise when compared to AVAR. Since the bias never occurs, the unbiased white-FM-transform coefficient h_0 used for $S_y(f)$ does not depend on a high-cutoff, $f_c=1/(2\tau_{on})$, as indicated in Table 1. Random walk (and drift) are slightly biased (depends on $r = \tau_s/\tau_{on}$ [7]) and the positive τ-slope is the same for $\Psi^2(\tau_{on}, \tau_s)$ and $\sigma_y^2(\tau)$. It is important to note that flicker-FM noise using dead-time AVAR (here $\Psi^2(\tau_{on}=const., \tau_s=\tau)$) will appear as white-FM noise [8].

TABLE I. TABLE OF TRANSFORMS

Noise Type	Allan $\sigma_y^2(\tau)$	Psi Variance $\frac{1}{2}\Psi^2(\tau_{on},\tau_s)$
White FM	$\frac{h_0}{2\tau}$	$\frac{h_0}{2\tau}$
Flicker FM	$2h_{-1}\log 2$	$\frac{h_{-1}}{2\tau_{on}^2}\left(\begin{array}{l} 4\tau_s\tau_{on}\tanh^{-1}\frac{\tau_{on}}{\tau_s} \\ +2\tau_{on}^2\ln\frac{\tau_s}{\tau_{on}}+ \\ (\tau_s^2+\tau_{on}^2)\ln\left(1-\frac{\tau_{on}^2}{\tau_s^2}\right)\end{array}\right)$
Random Walk FM	$\frac{2\pi^2 h_{-2}}{3}\tau$	$\frac{\pi^2 h_{-2}}{3}(3\tau_s-\tau_{on})$

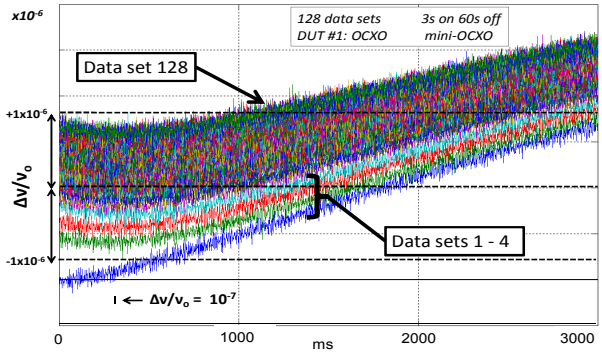


Figure 4. 128 sets of raw fractional frequency measurements. Set 1 is the first series of y(t), and the oscillator is first turned on from a "cold" start. Each trace is 3 s worth of data repeating every 60 s during which the oscillator is turned off. ν_o = 16.384MHz.

It remains to be seen if flicker-FM can be reliably detected by unraveling Ψ-variance to obtain AVAR. However, this distinction is usually unimportant to Doppler-relevant applications.

V. TIME DOMAIN MEASUREMENTS OF OSCILLATORS

We use a commercial miniature OCXO and TCXO at 16.384 and 26 MHz, respectively, as DUTs for an example. Fig. 4 shows 128 raw $y(\tau_{on})$ data runs of 2 ms sampled measurements on top of each other. At the very bottom, data set #1 starts the test oscillator. One can see that the first four sets capture a larger set-to-set overall variation than the remaining 124. In real applications, the oscillators are not cold-started but are in process, so the initialization sets such as 1 to 4 can generally be ignored. We process individual runs of Fig. 4 with dynamic ThêoH as described in Section II. This is shown in Fig. 5 along with averages of ThêoH. One can see that a consistent level of stability (drift+white FM) is reached after about 60 ms. Measurements are an equispaced sequence of fast-frequency errors, y(t), and are not time-errors, x(t). Thus, measurement noise is white FM and not typified by PM noise during runs of τ_{on}. Fig. 5 shows measurement white FM in short-term, i.e., $\sigma_y(\tau_0)\propto\tau^{-1/2}$.

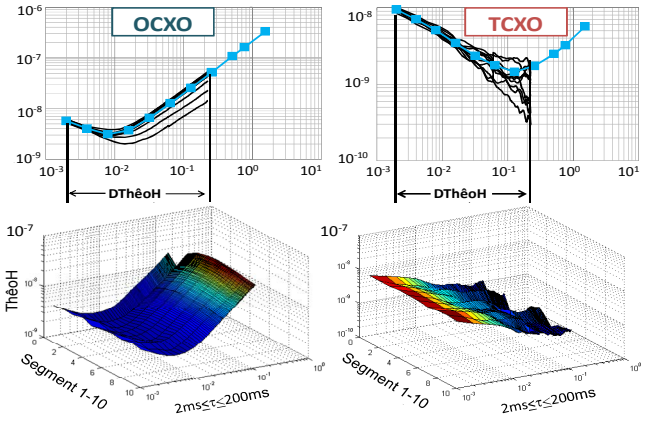


Figure 5. ThêoH deviation (top) and dynamic ThêoH (DThêoH) deviation (bottom) for the OCXO and TCXO. Note that the longest τ for DThêoH corresponds to 1/10 of the longest τ for ThêoH.

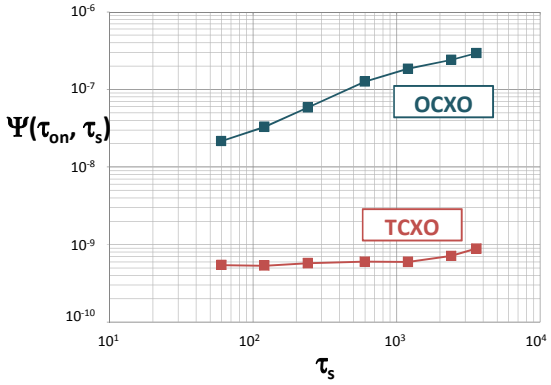


Figure 6. Ψ-deviation for the OCXO and TCXO. The minimum averaging time is τ_s = 60 s. τ_{on} = 3 s.

With the average of each data run, $^{\tau_{on}=3s}\overline{y}(t)$, we compute $\Psi^2(\tau_{on}, \tau_s)$ using all runs. Results are shown in Fig. 6, where we observe the level and rate of frequency reproducibility as a function of $n\tau_s$ for each time that the DUT is powered on. This level and rate may or may not limit other application-specific goals. Likewise, with a given level and rate, one may be forced to use application strategies or improve emitter reproducibility to achieve goals. Our finding is that there is no reliable method for estimating $\Psi(\tau_{on}, \tau_s)$ from $\sigma_y(\tau)$.

VI. Frequency Domain Measurements of Oscillators

Phase noise L(f) is important during limited-live Doppler tracking. L(f) is a convenient standard used to determine the error vs. range in offset-f, proportional to emitter velocity, as set by the emitter. L(f) is computed from the fast-frequency measurements obtained using Fig. 1. For a given τ_{on} data run sampled at t_0, the fractional-frequency spectrum $S_y(f)$ is obtained from the discrete FT of the series [6]:

$$Y(m\Delta f) = \frac{1}{N}\sum_{k=1}^{N} y(k\tau_0)e^{-2\pi i m\Delta f k\tau_0} \qquad (5)$$

where $\Delta f = 1/(N\tau_0)$. The one-sided spectral density of y(t) is computed by adding the squares of the real and imaginary components of Y and dividing by the RBW of the data run:

$$S_y(m\Delta f) = 2\frac{\left\{\mathrm{Re}\left[Y(m\Delta f)\right]\right\}^2 + \left\{\mathrm{Im}\left[Y(m\Delta f)\right]\right\}^2}{\Delta f} \qquad (6)$$

with BW = 1 Hz and RBW = Δf. Converting to L(f), we use [6]:

$$L(m\Delta f) = \frac{1}{2}S_\varphi(m\Delta f) = \frac{1}{2}\frac{\nu_0^2}{(m\Delta f)^2}S_y(m\Delta f) \qquad (7)$$

and obtain L(f) plotted on log-log scales in Fig 7.

In practice, the noise of each limited-live spectrum affects Doppler-track error. Averages of limited-live estimates of L(mΔf) are shown in Fig. 7 for TCXOs #1 and #2. A word of caution — L(f) derived from fast-frequency measurements in

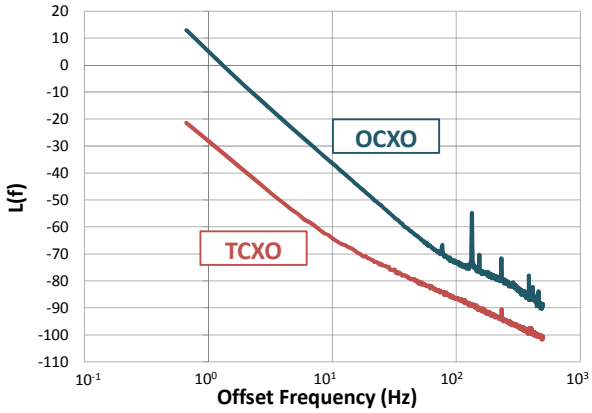


Figure 7. L(f) for the OCXO and TCXO from dead-time measurements.

Fig. 1 will not be sensitive to white and flicker PM noise, as mentioned earlier. This is not problematic to most limited-live characterizations, since the measurement high frequency cutoff (BW) is $f_c = 1/(2\tau_0)$ and τ_0 is of the order 10^{-3} in this case. Thus, L(f) is not computed beyond f of several hundred hertz, even in the best case.

For future studies, characterization of limited-live oscillators in this paper will be used while such oscillators are subject to temperature and vibration stresses.

ACKNOWLEDGMENT

The authors wish to thank Kyle Byrnes for setting up the preliminary phase of this experiment and writing initial computer code.

References

[1] D.A. Howe, "ThêoH: a hybrid, high-confidence statistic that improves on the Allan deviation," Metrologia, IOP, UK vol. 43, pp.S322-S331, 2006.

[2] L. Galleani and P. Tavella, "Tracking Nonstationarities In Clock Noises Using the Dynamic Allan Variance," FCS/PTTI, Vancouver, BC, pp.392-396, Aug. 2005.

[3] J.A. Taylor and D.A. Howe, "Fast TheoBR: A Method for Long Data Set Stability Analysis," IEEE Trans., vol.57, no. 9, pp.2091-2094, Sept. 2009.

[4] J. Levine, "Introduction to Time and Frequency Metrology," Review of Scientific Instruments, Vol. 70, pp 2567-2596, 1999.

[5] D.W. Allan, "Statistics of Atomic Frequency Standard," IEEE Proc., 54, No. 2, pp.221-231, 1966.

[6] D.A. Howe, D.W. Allan, and J.A. Barnes, "Properties of Signal Sources and Measurement Methods," Proc. 35th Annual Symp. on Freq. Control, A1-A48, 1981. Also see NIST Technical Note 1337, editors D.B. Sullivan, D.W. Allan, D.A. Howe, and F.L. Walls, March 1990, available at http://tf.boulder.nist.gov/general/pdf/868.pdf

[7] J.A. Barnes, "Tables of Bias Functions, B_1 and B_2, for Variances Based on Finite Samples of Processes with Power Law Spectral Densities," NBS Technical Note 375, Jan. 1969.

[8] P. Lesage, "Characterization of Frequency Stability: Bias Due to the Juxtaposition of Time-Interval Measurements," *IEEE Trans. Instrum. Meas.*, Vol. IM-32, No. 1, pp.204-207, March 1983.

Simplified Phase Noise Model for Negative Resistance Oscillators

Jeremy Everard and Min Xi
Department of Electronics
University of York,
Heslington, York,UK

***Abstract*— This paper describes a greatly simplified model for the prediction of phase noise in oscillators which use a negative resistance as the active element. It is based on a simple circuit consisting of the parallel addition of a noise current, a negative admittance/resistance and a parallel (Q limited) resonant circuit. The transfer function is calculated as a forward trans-resistance (V_{OUT}/I_{IN}) and then converted to power. The effect of limiting is incorporated by assuming that the phase noise element of the noise floor is KT/2 ie -177dBm/Hz at room temperature.. The answer is the same as more complex analyses but enables a simple clear insight into the operation of oscillators. The phase noise for a given power in the resonator in the resonator appears to be lower than in feedback oscillators. The reasons for this are explained.**

I. INTRODUCTION

The phase noise in oscillators often sets the ultimate limit in communications and RADAR systems. It is therefore useful to develop simple accurate linear theories which both highlight the underlying operating principles and is simple to use.

Oscillator phase noise for feedback oscillators was first described by Leeson [1] and the spectral density function was shown. This equation was manipulated and shown to have an optimum coupling coefficient between the resonator and the feedback amplifier for minimum phase noise as presented by Parker [2].

It is useful to develop simple linear theories from first principles and a simple accurate linear theory was developed by Everard at Philips Research Laboratories, UK in 1983 and published in [3] [4] [5]. At that time a similar theory for negative resistance oscillators was developed although some of the noise aspects were not completely clear.

II. FEEDBACK OSCILLATORS

The linear theory for feedback oscillators is based on the model shown in Figure 1 where the single input is modelled as two inputs which are added together. One of these is for noise injection and the other for the feedback path. A transfer function $V_{OUT(node2}/V_{IN2}$ can therefore be calculated, converted to power and a phase noise equation derived. The LCR circuit is used to model the resonator (and can be used for most types of resonators. Any coupling transformers can be incorporated by varying the values of L, C and R_{LOSS}.

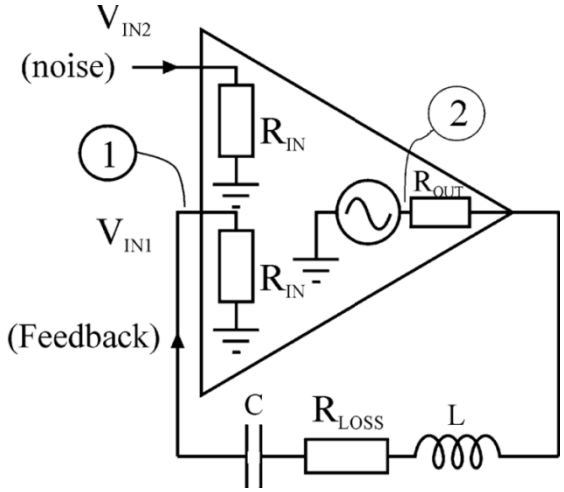


Figure 1 Simple Phase Noise Model for Feedback Oscillators

An equation for phase noise based on this model is shown in the following equation [5]:

$$L(f) = A \cdot \frac{FkT}{8\,(Q_0)^2\,(Q_L/Q_0)^2\,(1-Q_L/Q_0)^N\,P}\left(\frac{f_0}{f}\right)^2$$

Where:

1. N = 1 and A = 1 if P is defined as PRF and ROUT = zero
2. N = 1 and A = 2 if P is defined as PRF and ROUT = RIN
3. N = 2 and A = 1 if P is defined as PAVO and ROUT = RIN

If we take assumption 3, this equation becomes:

$$L(f) = 1 \cdot \frac{FkT}{8\,(Q_0)^2\,(Q_L/Q_0)^2\,(1-Q_L/Q_0)^2\,P_{AVO}}\left(\frac{f_0}{f}\right)^2$$

It is interesting to note the equation is very similar to the equation for power dissipated in a resonator (P_{RLOSS}) as a function of P_{AVO}

$$P_{RLOSS} = 2\frac{Q_L}{Q_0}\left(1 - \frac{Q_L}{Q_0}\right)P_{AVO}$$

The power in the resonator is (P_{RLOSS}) is maximum and hence phase noise is minimum when Q_L/Q_0 = ½. At this point, half the power is dissipated in the resonator, ¼ is transmitted and ¼ is reflected.

This therefore means that the phase noise is minimum when maximum power is dissipated in the resonator. This is also discussed in [6].

The phase noise equation for minimum noise is therefore:

$$L(f) = \frac{2FkT}{P_{AVO}}\left(\frac{f_0}{f}\right)^2$$

It should be noted that this equation is half the value of Leeson's equation when the optimum Q_L/Q_0 is included.. Leeson realised the factor of two shortly after he published his famous paper (private communication from Prof Leeson).

This will be compared later with the phase noise equation for negative resistance oscillators. This equation has been found to produce very accurate modelling of the phase noise of feedback oscillators (0 to 1dB) It also gives a clear insight into the operation of oscillators.

III. Negative Resistance Oscillators

A number of models have been produced for negative resistance oscillators [7-9]. A simple model for a negative resistance oscillator is shown in figure 2. This consists of the parallel addition of a noise current, a negative admittance/resistance and a parallel (Q limited) resonant circuit. The noise current is generated by R_{LOSS} and for this calculation the load resistance and noise figure are ignored. This is similar to the model shown in a paper by Hajimiri and Lee [9] although in their paper the calculation was performed in a different way and the emphasis was on the impulse sensitivity function.

The transfer function is calculated as a forward trans-resistance (Vout/Iin) and then converted to power. The effect of limiting is incorporated by assuming that the phase noise element of the noise floor is KT/2 ie -177dBm/Hz at room temperature

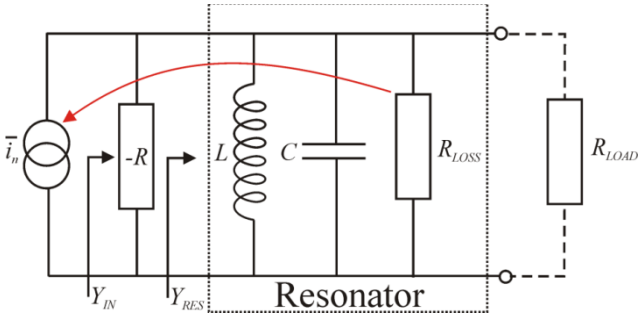


Figure 2 Simple Model for a Negative Resistance Oscillator

A. Calculation of the resonator Admittance

The admittance of the resonator is:

$$Y_{RES} = \frac{1}{j\omega L} + j\omega C + \frac{1}{R_{LOSS}}$$

If we now look at ω as the sum of a centre angular frequency and offset frequency ($\omega_0 + \Delta\omega$) where $\Delta\omega$ is the offset frequency:

$$Y_{RES} = \frac{-j}{(\omega_0 \pm \Delta\omega)L} + j(\omega_0 \pm \Delta\omega)C + \frac{1}{R_{LOSS}}$$

$$Y_{RES} = \frac{-j}{\omega_0\left(1 \pm \frac{\Delta\omega}{\omega_0}\right)L} + j\omega_0\left(1 \pm \frac{\Delta\omega}{\omega_0}\right)C + \frac{1}{R_{LOSS}}$$

At resonance: $\omega_0 L = 1/\omega_0 C$ Therefore

$$Y_{RES} = \frac{-j\omega_0 C}{\left(1 \pm \frac{\Delta f}{f_0}\right)} + j\omega_0\left(1 \pm \frac{\Delta f}{f_0}\right)C + \frac{1}{R_{LOSS}}$$

Where ω is the angular frequency.

Assuming that: $\Delta f << f_0$, then as:

$$Y_{RES} = j\omega_0 C\left(-\left(1 \mp \frac{\Delta f}{f_0}\right) + \left(1 \pm \frac{\Delta f}{f_0}\right)\right) + \frac{1}{R_{LOSS}}$$

$$Y_{RES} = 2j\omega_0 C\left(\frac{\Delta f}{f_0}\right) + \frac{1}{R_{LOSS}}$$

As $Q_0 = R_{LOSS}\omega_0 C \quad \therefore \omega_0 C = \frac{Q_0}{R_{LOSS}}$

$$Y_{RES} = \frac{1}{R_{LOSS}}\left[1 \pm 2jQ_0\left(\frac{\Delta f}{f_0}\right)\right]$$

This equation also describes the resonant response around the carrier for most other forms of parallel resonant circuits.

IV. Oscillation Conditions

If we now complete the oscillator by incorporating the negative resistance, the input admittance becomes

$$Y_{IN} = \frac{1}{R_{LOSS}}\left[1 \pm 2jQ_0\left(\frac{\Delta f}{f_0}\right)\right] - \frac{1}{R}$$

Under oscillating conditions the negative resistance cancels the positive resistance and therefore:

$$Y_{IN} = \left[\frac{\pm 2jQ_0}{R_{LOSS}}\left(\frac{\Delta f}{f_0}\right)\right]$$

We can now calculate the transfer function in terms of the forward trans-resistance and therefore calculate V_{OUT} in terms of i_n:

$$V_{OUT} = \frac{i_n}{Y_{IN}} = i_n \frac{R_{LOSS}}{2jQ_0}\left(\frac{f_0}{\Delta f}\right)$$

At resonance, Δf is zero and V_{OUT} becomes infinite whereas in reality it is actually just finite. However the

sideband noise is still significantly lower than the carrier for offset < 1Hz for most oscillators.

If we assume the noise is produced by the loss resistance in the amplifier then we can calculate the noise current as:

$$\left(\frac{i_n}{2}\right)^2 R_{Loss} = kTB$$

Therefore:

$$i_n = \sqrt{\frac{4kTB}{R_{LOSS}}}$$

The output voltage is now therefore

$$V_{OUT}(\Delta f) = \sqrt{\frac{4kTB}{R_{LOSS}}}\frac{R_{LOSS}}{2jQ_0}\left(\frac{f_0}{\Delta f}\right)$$

For noise, it is usually necessary to deal in terms of power the output power at an offset (not including zero) is therefore:

$$\frac{\left(V_{OUT}(\Delta f)\right)^2}{R_{LOSS}} = \frac{kTB}{{Q_0}^2}\left(\frac{f_0}{\Delta f}\right)^2$$

As this theory is a linear theory, the sideband noise is effectively amplified narrow band noise passed through a Q multiplication filter. Note the Q multiplication is approximately the ratio of the output voltage over the input voltage.

This is huge! It occurs because the voltage transfer function increases by $1/\Delta f$ as power is proportional to $1/\Delta f^2$.

The 3dB point in the resonator now becomes the 3dB above the noise floor!

To represent this as an ideal carrier plus sideband noise, the signal can be thought of as a carrier with a small perturbation rotating around it.

Note that there are two vectors rotating in opposite directions, one for the upper and one for the lower sideband. The sum of these vectors can be thought of as containing both amplitude modulation (AM) and phase modulation (PM). The component along the axis of the carrier vector being AM noise and the component orthogonal to the carrier vector being phase noise. PM can be thought of as a linear modulation as long as the phase deviation is considerably less than 0.1 rad.

Although linear, this theory can incorporate the non-linearities, i.e. limiting in the amplifier, by modifying the absolute value of the noise. If the output signal amplitude is limited with a 'hard' limiter, the AM component would disappear and the phase component would be half of the total value. This is because the input noise is effectively halved. This assumes that the limiting does not cause extra components due to mixing. Limiting also introduces a form of coherence between the upper and the lower sideband which has been defined as conformability by Robins [10]. The phase noise sideband output power is therefore

$$\frac{\left(V_{OUT}(\Delta f)\right)^2}{R_{LOSS}} = \frac{kTB}{2{Q_0}^2}\left(\frac{f_0}{\Delta f}\right)^2$$

To obtain $L(f)$ we now just need to divide by the output power of the oscillator (P) and therefore

$$L(f) = \frac{kTB}{2{Q_0}^2 P}\left(\frac{f_0}{\Delta f}\right)^2$$

This equation is very similar to that shown by Edson [7], Grebennikov [8] and Hajimiri and Lee [9]

V. Comparison of Feedback and Negative Resistance Oscillator Phase Noise Equations

It is interesting to compare this with the equation for phase noise for a feedback oscillator

$$L(f) = \frac{2kTB}{{Q_0}^2 P_{AVO}}\left(\frac{f_0}{\Delta f}\right)^2$$

This is 4 times larger. A factor of 2 can be explained by the fact that P in the negative resistance is the power dissipated in the resonator whereas PAVO is twice the value of the power dissipated in the resonator under optimum operating conditions. The other factor of two is probably due to the fact that the S/N ratio is set by the power at the input of the amplifier which is a ¼ in the feedback oscillator. Never the less it may still be important in some circumstances when the power in the resonator needs to be kept low that for the same power in the resonator the negative resistance amplifier offers a factor of two improvement in the phase noise.

VI. Evaluation of Negative Resistance Amplifiers

The easiest way to evaluate the performance of a negative resistance amplifier in terms of its noise figure and transposed flicker noise is to turn it into a two port amplifier to enable conventional measurement of noise figure and transposed flicker noise. A system is shown if figure 3 which uses a circulator to turn negative resistance into a conventional two power amplifier. The diplexer is added to try and ensure near 50 termination both inside and outside the band as illustrated in Figure 4, The diplexer has the additional requirement to incorporate the change of out of band impedance caused by the circulator. The S parameters for the circulator are measured and the filter response adjusted to take this into account. A plot of an example operating at 1 GHz is shown in Figure 5

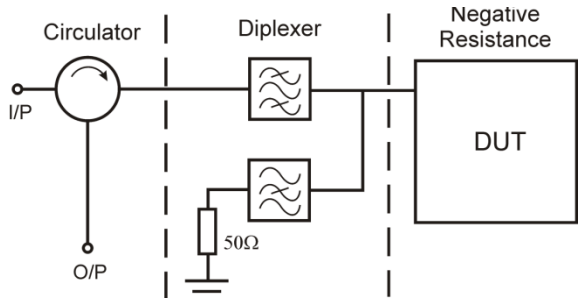


Figure 3 Technique to Evaluate Negative Resistance

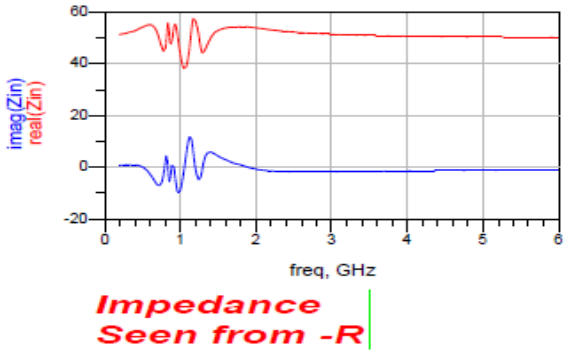


Figure 4 Variation in impedance see from –T (DUT)

An implementation for a typical circuit was implemented for this and tested in ADS harmonic balance software. The noise figure for the simulated negative resistance amplifier was measured in simulation and a phase noise plot is shown in Figure 5 and compared with theory. It is seen that the two plots are very similar.

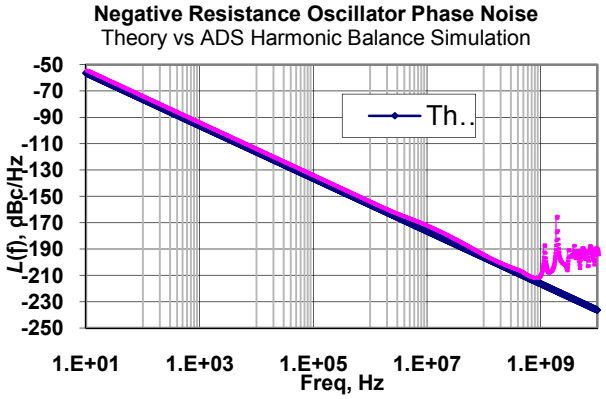


Figure 5 Comparison of theoretical and simulated Phase Noise

VII. Noise Floor Reduction

It is interesting to note that low noise floors appear to be available in the simulatation. A discussion of reduced noise floors is presented in an Agilent Patent by John R Burgoon in 1981.

VIII. Transposed Flicker Noise Reduction

A further justification for using a linear theory is if that transposed flicker noise corners can be significantly reduced. A number of authors have been working on this [11-15]. This group has recently shown that both series and parallel feedback can be used to reduce flicker noise significantly [16][17] Figure 6 and demonstrated some amplifiers with between 1 to 2 order magnitudes of flicker noise reduction. Typical phase noise plots are shown in figures 7 and 8. These phase noise plots were downconverted from 100MHz to enable them to be measured on this particular all digital phase noise measurement system.

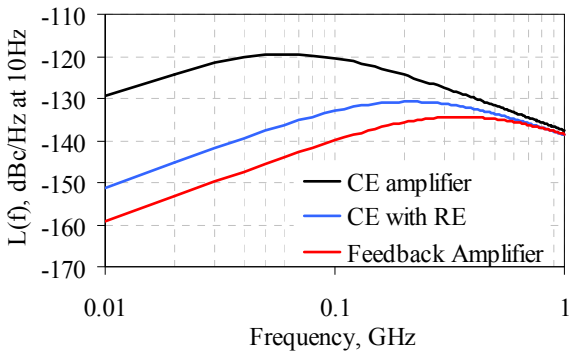


Figure 6 Transposed Flicker Noise Reduction for Series and Parallel Feedback Amplifiers

This group has recently shown that both series and parallel feedback can be used to reduce flicker noise and demonstrated some amplifiers with between 1 to 2 order magnitudes of flicker noise reduction. A typical phase noise plot is shown in figures 5 and 6.

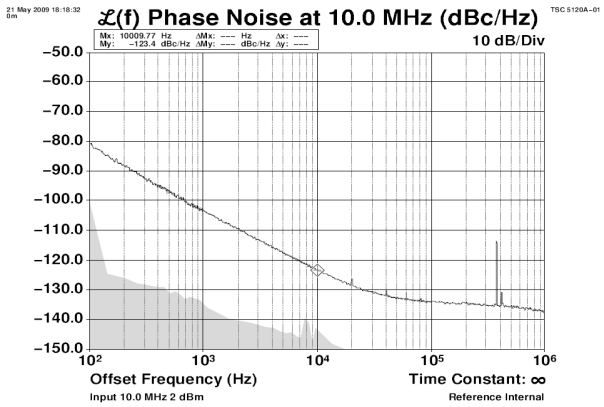


Figure 7 Phase Noise Plot for 100MHz Oscillator

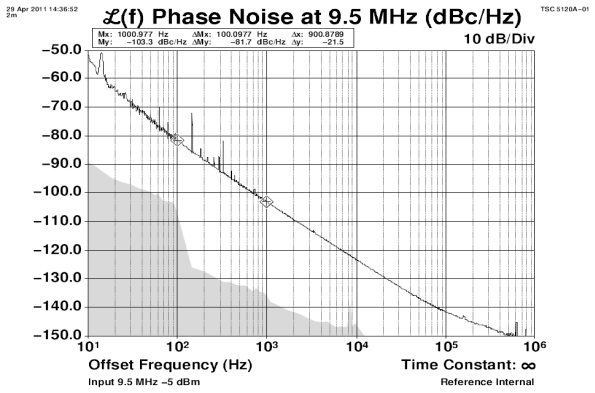
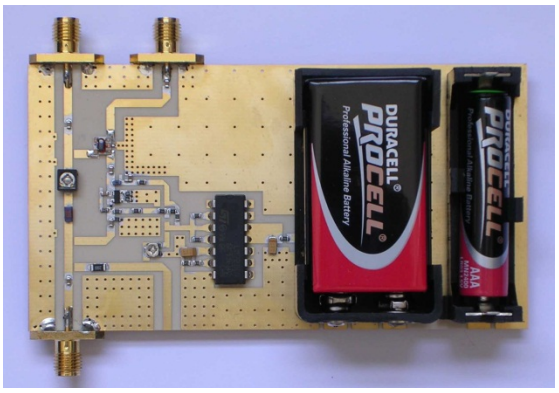


Figure 8 Phase Noise Plot for 100MHz Oscillator

These circuits have been used in a battery powered labkit for the one day courses on 'Low Phase Noise Oscillators: Theory and Laboratory run at the IEEE International

Microwave Symposium:- IMS 2009, 2010 and 2011. This labkit is shown in Figure 8 and enables five experiments to be performed.



IX. CONCLUSIONS

A simple linear theory for phase noise in negative resistance oscillators has been presented. A comparison between the phase noise equations for feedband and negative resiatantce amplifiers has been discussed.

X. ACKNOWLEDGEMENTS

J Everard would like to thank BAE Systems and the Royal Academy of Engineering for supporting him and this work. He would also like to acknowledge Philips Research Laboratories where very similar calculations were recorded in his work lab book in 1983 although the noise modelling for negative resistances was not presented clearly there. Min Xu would like to thank SEMTECH for their support during her studies towards a PhD.

XI. REFERENCES

[1] D.B. Leeson, A Simple Model of Feedback Oscillator Noise Spectrum", Proceedings of the IEEE, 54, pp. 329-330, Feb. 1966.

[2] T.E. Parker, "Current Developments in SAW Oscillator Stability", Proceedings of the 31st Annual Symposium on Frequency Control, Atlantic City, New Jersey, 1977, pp. 359–364.

[3] J.K.A. Everard "Low Noise Power Efficient Oscillators: Theory and Design" IEE Proceedings Part G, Vol. 133, No.4, August 1986, pp. 172-180.

[4] Jeremy Everard "Fundamentals of RF Circuit Design with Low Noise Oscillators" ISBN 0 47149793 2, Wiley - December 2000, Reprinted October 2002

[5] J.K.A. Everard "A Review of Low Noise Oscillator, Theory and Design" 1997 IEEE International Frequency Control Symposium, May 27th-30th 1997, Orlando, USA, pp.909-918.

[6] E. Vaury, V. Giordano, J.C. Nallatamby, J. Groslambert, M. Camiade, M. Pringent, O. Llopis, E. Gonzalez, M. Chaubet, J. Obregon. A New Method for the Design of Ulltra Low Noise Oscillators.. Joint EFTF/IFCS 1999

[7] WA Edson, "Noise in Oscillators", Proc. IRE,48, pp.1454-1466, August 1960

[8] A. Grebennikov, "RF and Microwave Transistor Oscillator Design", pp. 214-216, Wiley, 2007.

[9] T.H. Lee and A. Hajimir. Oscillator Phase Noise: A Tutorial. IEEE Journal of Solid State Circuits, Vol 35, No. 3 March 2000, pp.326- 336.

[10] W. P. Robins, Phase Noise in Signal Sources, Peter Peregrinus. Ltd., London, 1982

[11] F. L. Walls, E. S. Ferre-Pikal, and S. R. Jefferts, "Origin of 1/f. PM and AM noise in bipolar junction transistor amplifiers," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 44, no. 2, pp. 326–334, Mar. 1997.

[12] E. S. Ferre-Pikal, F. L. Walls, and C. W. Nelson, "Guidelines for designing BJT amplifiers with low 1/f AM and PM noise," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 44, no. 2, pp. 335–343, Mar. 1997.

[13] V. N. Kuleshov and T. I. Boldyreva, "1/f AM and PM noise in bipolar transistor amplifiers: Sources, ways of influence, techniques of reduction," in Proc. IEEE Int. Frequency Control Symp., May 1997, pp. 446–455.

[14] V. N. Kuleshov, "1/f models of bipolar junction transistor and their application to PM and AM noise calculation," in Proc. IEEE Int. Frequency Control Symp., May 1998, pp. 164–171. T.-D. Tomlin, K. Fynn, and A. Cantoni, "A model for phase noise generation in amplifiers," IEEE Trans. Ultrason. Ferroelectr. Freq.Control, vol. 48, no. 6, pp. 1547–1554, Nov. 2001.

[15] E. S. Ferre-Pikal and F. H. Savage, "Up-converted 1/f PM and AM noise in linear HBT ampliers," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 55, no. 8, p. 1698–1704, Aug. 2008.

[16] Konstantinos Theodoropoulos and Jeremy Everard, Residual Phase Noise Modelling of Amplifiers using Silicon Bipolar Transistors, Joint European Frequency and Time Forum and IEEE International Frequency Control Symposium, Besançon, France, April, 2009

[17] Konstantinos Theodoropoulos and Jeremy Everard, Residual Phase Noise Modeling of Amplifiers Using Silicon Bipolar Transistors, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control,, 57(3):562 - 573, March, 2010

An integrated phase noise measurement bench for on-chip characterization of resonators and VCOs

S. Godet[1,2,3], É. Tournier[1,2], O. Llopis[1,2],
[1]CNRS ; LAAS ; 7 avenue du colonel Roche, 31077 Toulouse, France
[2]Université de Toulouse ; UPS, INSA, INP, ISAE ; LAAS ; Toulouse, France
[3]*S. Godet is today with Axess Europe, Toulouse, France*

A. Cathelin
ST Microelectronics, Minatec, BP 257, 38016 Grenoble, France

***Abstract* : A phase noise measurement bench is integrated on a 3.6 mm^2 silicon chip. The bench includes a splitter with quadrature outputs, a phase detector, a low noise baseband amplifier and, if necessary, a synthesized source. Applications to the characterization of frequency sources and BAW resonators are discussed.**

***Index terms* : Phase noise, Built-in-self-test (BIST), integrated circuit, phase detection, SMR resonator**

I. INTRODUCTION

A phase noise measurement bench is generally a costly laboratory equipment, based on active references or long delay lines, including low noise amplifiers and phase detectors. The ability to include on a single silicon chip the elements of such a measurement bench allows the measurement and the monitoring of this critical parameter in small size systems, and probably also the correction of the sources phase noise using active loops. The device under test may either be a frequency source (ex : a VCO) or a two port system, such as an amplifier, a digital circuit… In the first case, the phase noise measurement bench must include a frequency reference element, either passive (resonator, delay line) or active (reference source). In the second case, it is mainly based on a low noise phase detection process.

A special case of two ports devices characterization is the measurement of piezoelectric resonators noise, which is very difficult to perform using conventional phase noise measurement techniques on 50 Ω load. Indeed, these devices need different impedances conditions to take benefit of their series or parallel resonances. While realized on a chip, the input and output impedances can be controlled and the resonator is characterized in the same conditions it experiences when it is included in a circuit.

Such an approach can be part of a Built-In-Self Test system [1-4], and used for reliability applications. As an example, the ability to measure the phase noise degradation in an embedded system allows to switch from the noise degraded device (oscillator, amplifier…) to another device, thus maintaining the system performance at its best.

Various elements of a phase noise bridge have thus been designed using an integrated circuit foundry (ST Microelectronics, BiCMOS7RF 0.25 µm). The system has been firstly designed to measure the phase noise of integrated BAW resonators using an SMR technology (CEA-LETI). However, the replacement of one of the resonators by a short line results in an enhancement of the detection of the frequency fluctuations of the source, and thus in the realization of a narrow bandwidth frequency discriminator. This discriminator is able to measure the phase noise spectrum of any type of LC resonator based VCOs.

II. PHASE NOISE MEASUREMENT SYSTEM

Figure 1 represents the typical circuit topology of such a phase noise measurement bench. The main device is a phase detector, generally realized with a mixer with the two signals (LO and RF) maintained close to the phase quadrature (±90°). The phase detector is followed by a baseband amplifier, which should feature an equivalent input low frequency noise lower than the output noise of the phase detector.

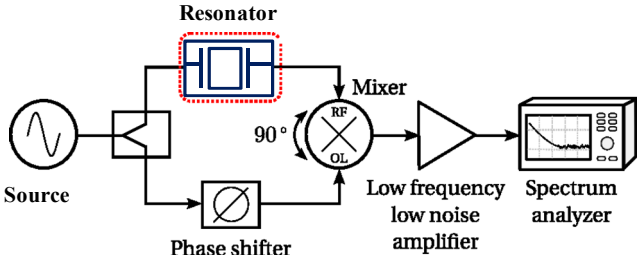


Figure 1 : Typical topology of a phase noise measurement bench, in the configuration for source measurement (frequency discriminator). The residual phase noise measurement system is the same, with the resonator substituted by the device under test

If the purpose of the measurement bench is to characterize the frequency source, the resonator is used as the frequency reference element, and is part of the measurement bench. It allows, thanks to its large phase to frequency slope (dϕ/df), the conversion of the frequency fluctuations of the source into phase fluctuations, which are then detected by the phase detector.

If the purpose of the measurement bench is to measure the residual phase noise of two ports devices, such an amplifier, the device under test (DUT) replaces the resonator, but a frequency source should be available in the system. This oscillator must be tunable and, above all, must feature low AM and FM noise. The AM noise is the most critical specification in this case [5], because the FM noise is simply removed by equilibrating the delay in the two arms of the phase detector. In the case of a resonator measurement, the problem of the parasitic FM noise detection of the source is solved by measuring two identical resonators, one in each arm of the phase detector, thus canceling the delay caused by one resonator. The resulting phase noise is the addition of the phase noise of the two resonators, i.e. a phase noise increased of 3 dB compared to a single resonator, if the resonators are identical. Finally, in the case of the measurement of frequency conversion devices (frequency dividers, multipliers…), once again two identical devices must be used, in order to reach the same frequency on the RF and LO ports of the mixer.

One of the devices depicted in Figure 1 is very difficult to integrate : the phase shifter. It is realized on a macroscopic phase noise measurement system with a stretched line, which allows phase rotations of generally more than 180° at all the frequencies of interest. Such a wide control of the phase is difficult to get on integrated circuits. However, in the integrated world, the phase is generally well known and does not shift too much on the small lines used in a silicon chip. We thus propose a solution to this problem which avoids the realization of this phase control.

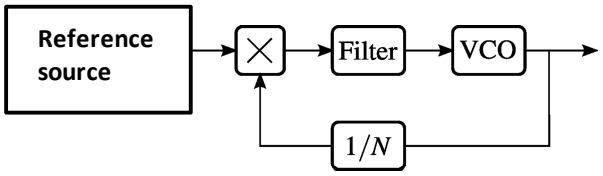
Finally, the output signal is analyzed on a low frequency spectrum analyzer. We have not integrated on a chip such a complex system, although it would be possible to do it. In our case, the final analysis is still performed out of the chip. However, it would be easy to integrate a power measurement on a given frequency bandwidth, which would allow to get an estimate of the phase noise at a given frequency offset carefully chosen for the application foreseen.

We will now describe all the elements of this system and the way they have been effectively integrated on silicon, starting with the input stages and ending with the phase detector stage.

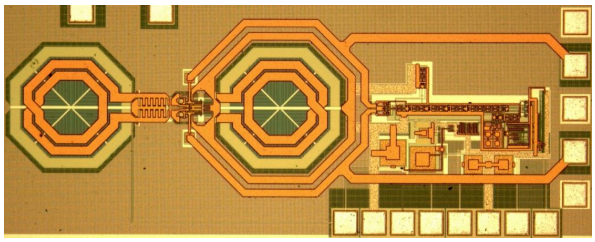
A. Synthesized frequency source

An integrated synthesizer has been realized, and is depicted in Figure 2. It is based on a LC differential VCO, which features a phase noise of -80 dBc/Hz at 10 kHz offset from a 7.5 GHz carrier, and on a PLL realized with a digital prescaler. The PLL allows the locking of the VCO on a low frequency reference source (30 MHz) and the generation of stable signals with a 30 MHz step on the whole VCO locking range (6.8 GHz to 8 GHz). Only the reference source (quartz oscillator) is maintained outside the chip.

a)



b)



Figures 2a and 2b : Synthesized frequency source (6.8 GHz to 8 GHz) based on a digital prescaler PLL ; circuit topology (a) and photograph of the chip (b)

B. Input stage : splitter and quadrature output

The problem of the phase quadrature at the mixer input can be solved very easily in an integrated circuit approach with a power divider delivering signals already in quadrature. If the phase shift in the DUT is small, or if it can be compensated using two identical devices, the phase quadrature is maintained down to the mixer stage, and the phase detection is performed at its maximum sensitivity.

To realize this function, a digital frequency divider by two has been used. This circuit, depicted in Figure 3, is based on D flip-flop cells and generates four output signals which are all in phase quadrature (0°, 90°, 180° and 270°). However, such a digital circuit located at a critical stage of the phase bridge can be the main noise source of the system. Our goal has thus been to optimize the phase noise of this circuit [6], in order to get a very low residual phase noise divider. This has been performed using the phase noise simulation capabilities of Cadence™, and particularly the ability of this software to isolate the main noise source of the circuit. An important optimization work has been focused on the D flip-flop topology, and more precisely on the current source of this differential element. More details on this process can be found in reference [6].

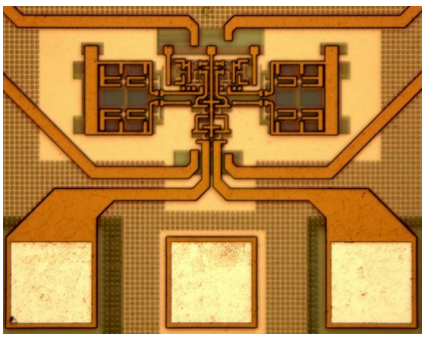


Figure 3 : Input power divider and frequency divider by two, featuring four outputs in phase quadrature

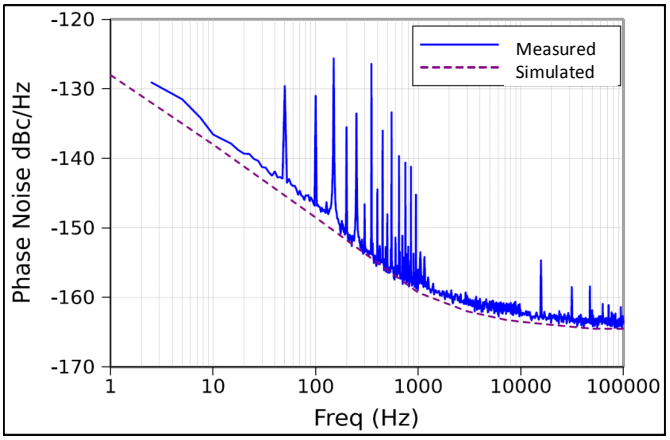


Figure 4 : Measured and simulated phase noise of the input frequency divider by two

Figure 3 represents the realized circuit, and Figure 4 its residual phase noise simulated using Cadence™ and measured on a conventional (macroscopic) phase noise measurement set-up. The agreement between simulated and measured data is excellent, and the phase noise floor effectively reached is as low as -164 dBc/Hz at 100 kHz offset from an input frequency of 3.5 GHz.

C. *Phase detector*

Phase detectors are classically realized with four balanced diodes mixers. Such phase detectors are very reliable and can work on a large frequency range. However, the main drawback of this approach lies in the need for a sufficient LO power to drive the mixer (typically 7 dBm) and on the conversion losses of passive mixers. It is therefore interesting to investigate on active mixers for phase detection. The residual phase noise of some bipolar transistor, when loaded and biased properly, is as low as the residual phase noise of schottky diodes silicon mixers, and these devices can lead to a substantial improvement of the phase detection coefficient.

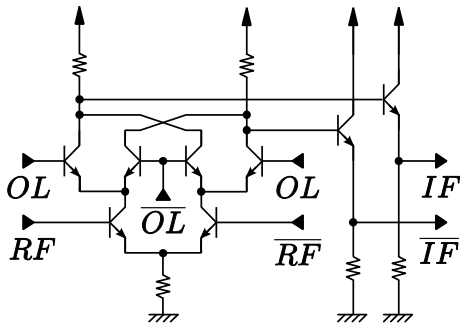


Figure 5 : Active phase detector, with resistive loads

Various active phase detectors have thus been designed, based on Gilbert cell mixer and using bipolar devices. Two versions of this circuit have been realized, one based on resistive loads (Figure 5) and featuring a phase detection coefficient of 350 mV/rad, and another one based on active loads and featuring a phase detection coefficient of 970 mV/rad. Up to now, only the first circuit has been measured and used in the integrated phase noise bench, although the other circuit features a better simulated phase noise performance. When this second mixer will have been fully characterized, it will be the purpose of a dedicated publication. However, the resistive loads mixer is already particularly interesting because it features a phase detection coefficient higher than passive mixers and works on much lower level signals. Moreover, its phase noise is low, as it can be deduced from the overall system phase noise measurement presented in section III (Figure 9).

D. *Low noise amplifier*

The low noise amplifier must feature an equivalent input noise level lower than the output noise of the phase detector. Using the resistive loads mixer, the phase detection coefficient is (a little higher but) close to the phase detection coefficient of a passive mixer, and the amplifier performance should be as close as possible as the one of the best operational amplifiers. Such metrological amplifiers feature an equivalent voltage noise floor in the range of 1 nV/√Hz, and a 1/f noise corner frequency lower than 1 kHz. Our goal has thus been to integrate an amplifier with similar performances.

Our first attempt with a conventional differential topology first stage has led us to noise performances in the range of 20 nV/√Hz or more at the amplifier input. The strategy to improve the amplifier performance has been based on the duplication of the bipolar transistors involved in the first differential stage and, also, in the gain stage. Indeed, setting in parallel many transistors reduces the equivalent noise (at a cost of an input impedance reduction, which is not mandatory here) because the noise sources of each device are added incoherently meanwhile the current (the information) is added coherently. The result of this study has been the realization of a state of the art low noise amplifier, featuring a noise floor of 1.05 nV/√Hz and a noise corner frequency in the kHz range. This is particularly interesting, taking into account the small size (640 µm X 227 µm) and, above all, the low power consumption of this circuit : 2.5 V, 18 mA (45 mW).

Figure 6 represents the circuit realized, and Figure 7 the measured low frequency noise of this amplifier. The noise measurement is in perfect agreement with the simulated noise, and demonstrates that the design goals have been fulfilled. More details on the noise optimization process for this circuit can be found in reference [7].

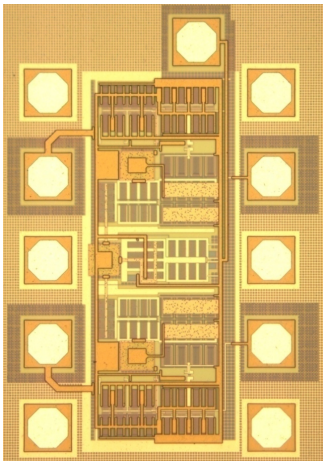


Figure 6 : Photograph of the low noise output amplifier

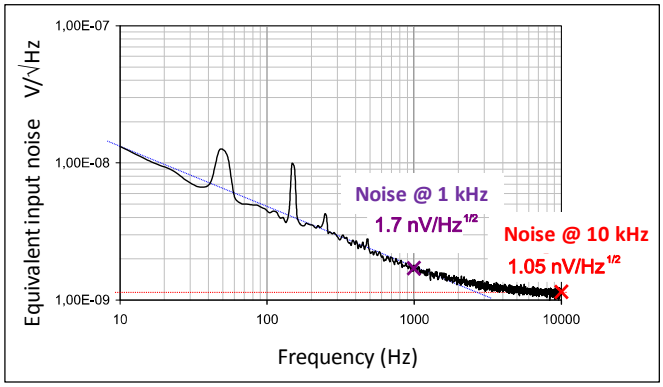


Figure 7 : Measured equivalent input voltage noise of the amplifier

E. Overview of the whole integrated circuit

These elements of the circuit are then put together on a single silicon chip. The final circuit is represented in Figure 8, and its size is in the range of 1.67 mm X 2.17 mm. The central part is dedicated to the DUT (case of residual phase noise measurement) or the resonators (BAW) and lines, which are reported on the circuit using a flip-chip technology. As the input stage provides four outputs in quadrature, we use these four outputs in order not to lose power, and the mixer is also designed with four inputs. The signal at the output of the low noise amplifier is then analysed on a FFT spectrum analyser. The system noise floor is measured using four identical short lines in place of the resonators.

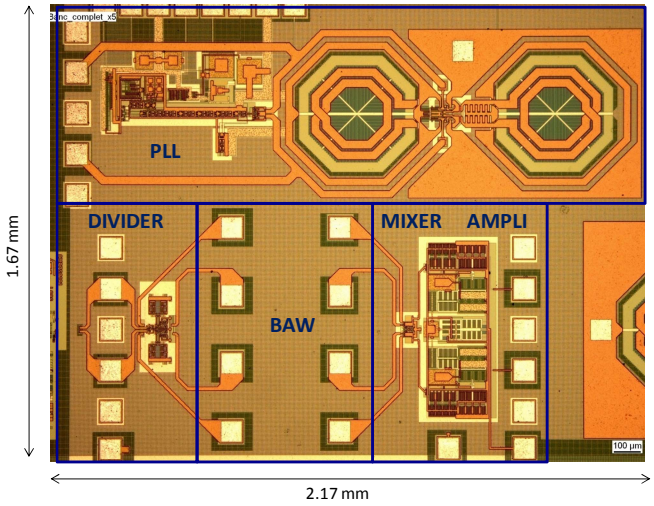


Figure 8 : Overview of the integrated phase noise bench

III. Application to the Characterization of Integrated BAW Resonators Phase Noise

Unlike many passive microwave resonators, BAW resonators may induce a 1/f noise component in the circuit. As an example, we have measured some years ago on a conventional phase noise bridge a 1/f phase noise component in FBAR resonators [8], and the measured phase noise level in these devices was sufficient to be the main noise contribution to an oscillator realized with one of these resonators. However, the same measurements performed in the same conditions on some new SMR resonators have given results much too close from the measurement noise floor to conclude on the noise contribution of these devices. This is due not only to the absolute level of phase noise (or more precisely to the fluctuation of the resonator centre frequency) but also to the fact that a proper impedance must be used to get an observable resonance with these devices.

We have thus used our integrated phase noise bench to measure the residual phase noise of those SMR resonators, the resonators being realized at CEA-LETI in Grenoble (France). More details on these resonators and their applications can be found in reference [9]. Four identical resonators have thus been mounted using flip-chip technology on the silicon circuit. Unfortunately, some contacts problems due to differences in metals pads between the two devices have prevented the measurement of this integrated circuit. We have thus used the separated elements to assemble a partially integrated phase noise test bench, and we have obtained the data depicted in Figure 9. The measured noise is the one of the four devices (incoherent addition of the noise of each device), which is effectively higher than the noise floor. The phase noise of a single resonator can be deduced from this plot by subtracting 6 dB to this curve.

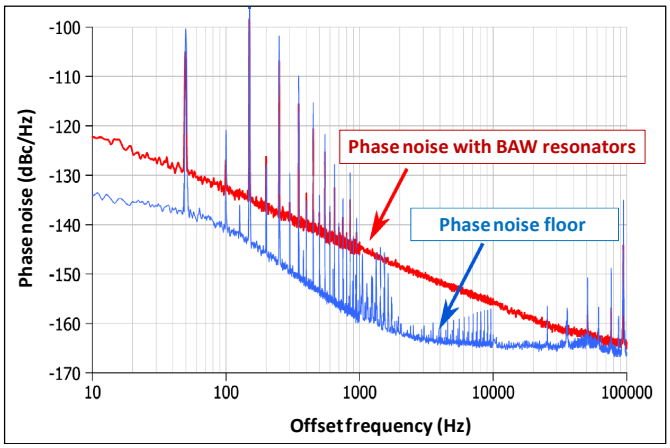


Figure 9 : Measured phase noise of the four SMR resonators, and phase noise floor of the measurement bench

Apart from the resonators residual phase noise, this measurement demonstrates a relatively good phase noise floor for our system. The phase noise floor is close to the measured phase noise of the input stage (frequency divider by two), with a bump on the spectrum added around 100 Hz offset, which is probably due to a contribution from the phase detector noise.

IV. Conclusion

A phase noise measurement bench integrated on a silicon chip has been presented. This bench may be used either for residual phase noise measurements, or for oscillator phase noise measurement in the frequency discriminator configuration, with a BAW resonator as frequency reference. Some of the elements of this bench are at the state of the art for integrated circuits functions in terms of noise performance (low noise amplifier, frequency divider). An application to the

measurement of the noise properties of some BAW-SMR resonators is also presented.

ACKNOWLEDGMENTS

We would like to thank A. Reinhardt and J.B. David from CEA-LETI for integrated resonator realization. This work has been performed in the frame of the contract "MOBILIS" which was part of the European FP6-IST program.

REFERENCES

[1] B.R. Veillette, G.W. Roberts, "On-chip measurement of the jitter transfert function of charge-pump phase locked loops", IEEE Journal of Solid State Circuits, vol 33, N° 3, March 1998, pp. 483-491.

[2] W. Khalil, B. Bakkaloglu, S. Kiaei, "A self calibrated on chip phase noise measurement circuit with -75 dBc single tone sensitivity at 100 kHz offset", IEEE Journal of Solid State Circuits, vol 42, N° 12, Dec. 2007, pp. 2758-2765.

[3] A. Valdes-Garcia, W. Khalil, B. Bakkaloglu, J. Silva-Martinez, E. Sanchez-Sinencio, "Built-in Self Test of RF Transceiver SoCs : from Signal Chain to RF Synthesizers", 2007 IEEE RFIC Symposium, Honolulu, June 2007, pp. 335–338.

[4] A. Cathelin, S. Godet, O. Llopis, E. Tournier, S. Thuriès, "Circuit électronique comprenant un dispositif de mesure du bruit de phase d'un dispositif oscillant et/ou résonant", Patent STMicroelectronics SA and CNRS, B0/-110879FR, 05-GR1-404, 2008.

[5] G. Cibiel, M. Regis, E. Tournier, O. Llopis "AM noise impact on low level phase noise measurements", IEEE trans. on Ultrasonics Ferroelect. and Freq. Control, vol. 49, n° 6, June 2002, pp. 784-788.

[6] S. Godet , E. Tournier , O. Llopis , A. Cathelin , J. Juyon, "A low phase noise and wide-bandwith BiCMOS SiGe:C 0.25um digital frequency divider for an on-chip phase-noise measurement circuit", 9th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SIRF 2009), San Diego, Jan. 2009, 4 p.

[7] S. Godet , E. Tournier , O. Llopis , A. Cathelin , J. Juyon, "A baseband ultra-low noise SiGe: C BiCMOS 0.25 um amplifier and its application for an on-chip phase-noise measurement circuit", 9th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SIRF 2009), San Diego (USA), Jan. 2009, 4 p.

[8] S. Gribaldo, C. Chay, E. Tournier, O. Llopis, "Experimental study of phase noise in FBAR resonators", IEEE Transactions on UFFC, vol 53, n° 11, nov. 2006, pp. 1982-1987.

[9] A. Reinhardt, G. Parat, E. Defaÿ, M. Aïd, "Acoustic Technologies for Advanced RF Architectures" 2010 NEWCAS Conference, pp. 161-164.

Drive Level Dependence of Doubly Rotated Langasite Resonators with Different Configurations

Haifeng Zhang[a], Yuan Xie [a], Joseph A. Turner[b], and John A. Kosinski[c, d]
[a]Department of Engineering Technology, University of North Texas, Denton, TX, 76207
[b]Department of Engineering Mechanics, University of Nebraska-Lincoln
W317.4 Nebraska Hall, Lincoln, Nebraska 68588
[c]U.S. Army RDECOM CERDEC I2WD, ATTN: RDER-IWT, Fort Monmouth, NJ 07703-5211
[d]Department of Chemistry, Medical Technology, and Physics,
Monmouth University, West Long Branch, NJ 07764-1898

Abstract -**The miniaturization of crystal resonators and filters towards the MEMS (Micro-Electro-Mechanical Systems) and nano-structured scales demands improvement of nonlinear piez oelectricity theory and a better understanding of th e nonlinear b ehavior of new crystal materials. The nonlin earities affect the quality factor and acousti c behavior of MEMS and nano-structured resonators and filters. Among these nonlinear effects, drive level dependence (DLD), which describes the instability of the resonator frequency due to voltage level and/or power density, is an urgent problem for miniaturized resonators. Langasite, a new promising piezoelectric material, is of recent interest for a variety of applications because of its good temperature behavior, good piezoelectric coupling, low acoustic loss and high Q factor. It has been recently used to make high temperature MEMS. In this article, experimental measurements of drive level dependence of langasite resonators with different configurations (plano-plano, single bevel, and double bevel) are reporte d. The results show that the resonator configuration affects the DL D of the l angasite resonator. The drive sensitivity of resonators with different configurations is determined.**

I. Introduction

The miniaturization of crystal resonators and filters towards the MEMS and nano-structured scales demands better understanding and modeling of the nonlinear behavior of crystal materials. The nonlinearities affect the quality factor and acoustic behavior of the MEMS (Micro-Electro-Mechanical Systems) and nano-structured resonators and filters. These effects include force-frequency effect, electroelastic effect, mode coupling and activity dip, acceleration sensitivity, intermodulation and drive level dependence. Among these nonlinear effects, drive level dependence (DLD), which describes the instability of the resonator frequency due to voltage level and/or power density, is an urgent problem for miniaturized resonators. This behavior is of current interest due to the increasing importance of low phase noise oscillators, because there is a direct relationship between resonator drive level and oscillator phase noise [1]. Except for the influence on the frequency, drive level dependence could also worsen the activity dip problem [2], [3].

Langasite, a new promising piezoelectric material, combines many of the advantages of quartz, barium titanate and lithium niobate -- having high electromechanical coupling and good frequency-temperature characteristics. It is a potential material to substitute for current applications dominated by quartz. One of the most important advantages of langasite is that it will not undergo phase transitions up to its melting temperature of 1473 deg C, which makes it very promising to make high temperature MEMS sensors. Common piezoelectric devices made by quartz and lithium niobate are limited to an operation temperature of 520 deg C or 400-500 deg C; MEMS made by langasite will be able to operate at least 1050 deg C. Recently, several attempts have been made to use langasite as the material to make high temperature MEMS sensors [4] and nano devices [5], [6], [7]. Investigation of nonlinearities of langasite is critical for the real application and optimal design of these devices, and research on DLD is especially important for these miniaturized devices. The measurement of drive level dependence for Y-cut langatate and langanite resonators has been performed by Kim [8]; however, measurements for doubly rotated langasite resonators have not been performed yet. In this article, we show that the different configurations of the resonators also affect the drive level sensitivity of doubly rotated langasite resonators. Experimental measurements of drive level dependence of langasite resonators with different configurations (plano-plano, single bevel, and double bevel) are reported. The drive sensitivity of resonators with different configurations is determined.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

II. Sample preparation

There are three langasite resonator samples under test. Their diameter, thickness and cut angle are the same. All other parameters are different. The detailed design parameters for these three resonators are listed in Table I.

Table I: Resonator design parameters.

Resonator configuration	Plano-plano	Plano convex with double bevel	Plano convex with single bevel
Material orientation	YXlw $\Theta=45^0/\Phi=65^0$	YXlw $\Theta=45^0/\Phi=65^0$	YXlw $\Theta=45^0/\Phi=65^0$
Plate diameter (mm)	13.0	13.0	13.0
Thickness (mm)	0.65	0.65	0.65
Electrode diameter (mm)	5.50	5.61	5.61
Radius of curvature (diopter)		4	4
Abrasive size to grid the bevels (micron)		1	1
Fundamental resonant frequency (Hz)	2,577,150	2,752,600	2,409,750

III. Experiment setup

The experiment setup is shown schematically in Fig. 1. An Agilent 33220A function generator is used to excite the crystal; the function generator is connected to and controlled by a computer with Labview. The signal from the function generator is fed into the crystal resonator through a PI network shown in Fig. 2. The resistance values in the PI network are shown in Table II. The response signal from the crystal resonator is collected by a Tektronix TDS 3054 C oscilloscope, which is subsequently connected to and controlled by the same computer with Labview.

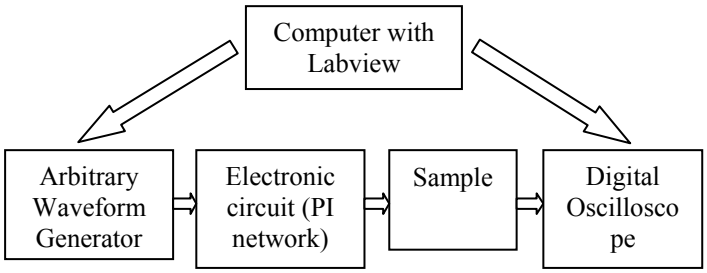


Fig. 1. The experiment setup.

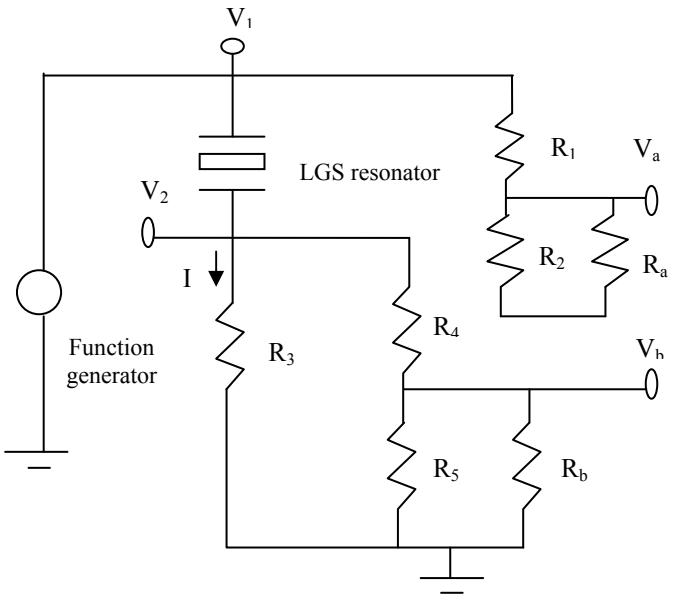


Fig. 2. The PI network circuit diagram.

Table II. The list of resistance values in the PI network.

R_1	R_2	R_3	R_4
1210 OHMS	51 OHMS	51 OHMS	1210 OHMS
R_5	R_a	R_b	I
51 OHMS	50 OHMS	50 OHMS	1.14Vb

IV. Measurement results

The relation between the resonator current and frequency is measured and shown in Figs. 3, 4, and 5 as a function of the magnitude (Volt) of the sinusoidal excitation signal from the function generator. As expected, the resonance curves shift towards a lower frequency as the voltage increases, which is a "softening" behavior. This is contrary to the DLD of SC cut quartz resonators in [9], as the resonance curves for SC cut quartz resonators bend towards higher frequency, which is a "hardening" behavior. The relative drive sensitivity D defined in [9] is extracted in Figs. 6, 7 and 8. The results are listed in Tables III and IV; Table III compares the D for the LGS resonators and SC cut quartz resonators with different

curvatures. It is apparent that the drive sensitivity of the tested LGS resonators exceeds SC cut quartz resonators. Table IV compares the D for LGS resonators and LGN, LGT resonators. It seems that the drive sensitivity of LGS resonators in fundamental mode is smaller than LGN and LGT resonators. Higher overtones of the tested LGS resonators could not be measured due to undesired mode interference.

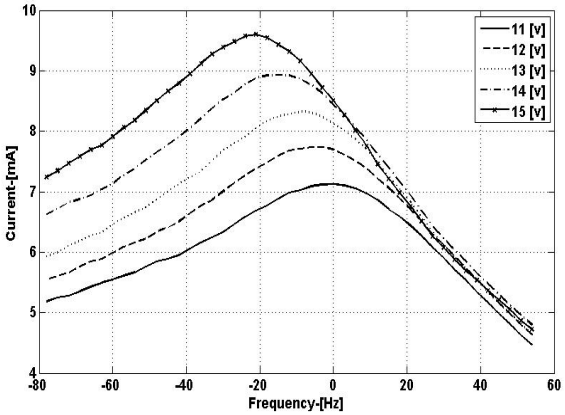


Fig. 3. Drive level dependence of LGS resonator with plano-plano configuration.

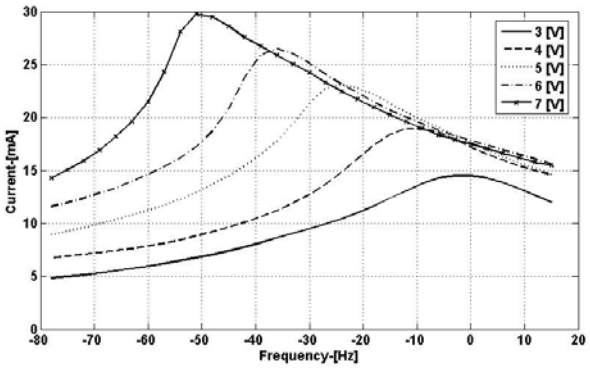


Fig. 4. Drive level dependence of LGS resonator with bevel-bevel configuration.

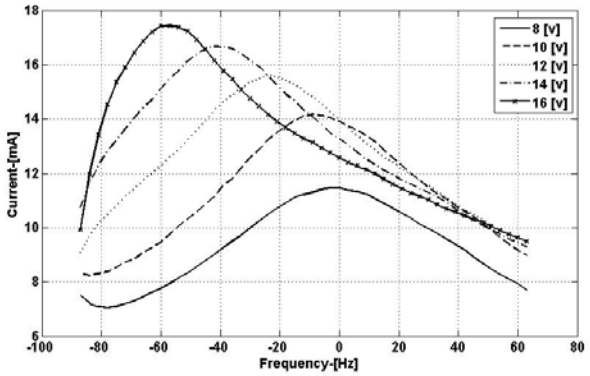


Fig. 5. Drive level dependence of LGS resonator with plano-bevel configuration.

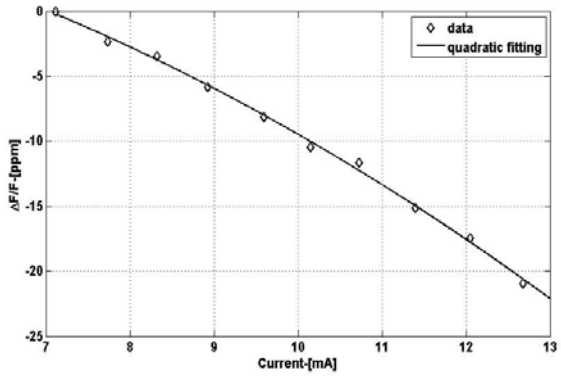


Fig. 6. Drive sensitivity of LGS resonator with plano-plano configuration.

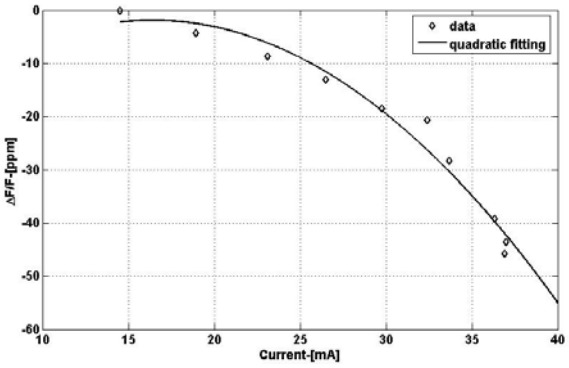


Fig. 7. Drive sensitivity of LGS resonator with bevel-bevel configuration.

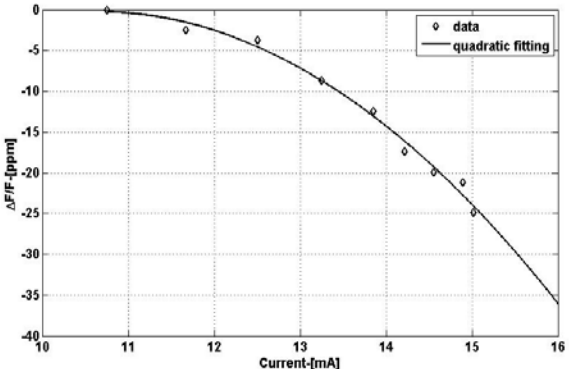


Fig. 8. Drive sensitivity of LGS resonator with plano-bevel configuration.

Table III: Comparison with SC-cut Quartz Resonators

Drive Sensitivity D [ppm/mA^2]		**Drive Sensitivity D [ppm/mA^2] for a SC cut r esonators at 5.115 MHz fundamental mode**	
LGS double bevel	-0.0954	Diopter 1.00	0.0076
LGS plano-plano	-0.1741	Diopter 1.25	0.0116
LGS single bevel	-1.2460	Diopter 1.5	0.0127

Table IV: Comparison with Langatate and Langanite Resonators

Drive level Sensitivity D [ppm/mA^2]	
LGS double bevel	-0.0954
LGS plano-plano	-0.1714
LGS single bevel	-1.2460
LGN	-3.34 (3rd OT)
LGT	-1.63 (3rd OT)

V. Conclusion

The drive level dependence of doubly rotated langasite resonators show softening behavior, which is contrary to the hardening behavior a SC cut quartz resonator exhibits. The change of resonator configuration has altered the drive level sensitivity. The resonator with single bevel configuration has the largest drive sensitivity coefficient; the resonator with double bevel configuration has the least drive sensitivity. Compared with SC-cut quartz resonators, the langasite resonator has generally larger drive level dependence. Compared with langatate and langanite resonators, langasite resonators show smaller drive level dependence.

VI. Acknowledgments

This research was supported by the Army Research Office under Grant No. W911NF-10-1-0293.

VII. Reference

[1] M. S. Patel, Y. K. Yong, M. Tanaka, and T. Imai, "Drive level dependency in quartz resonators", *Proc. 2005 Ann. Symp. On Freq. Cont.,* pp. 793-801, 2005.

[2] B. Parzen, *Design of Crystal and O ther Harmonic Oscillators,* John Wiley and Sons, 1983.

[3] A. Wood and A. Seed, Activity dips in at-cut crystals, *Proc. 21st Ann. Symp. On Freq. Cont.,* pp. 420-435, 1967.

[4] E. Ansorge, S. Schimpf, S. Hirsch, J. Sauerwald, H. Fritze, and B. Schmidt, "Evaluation of langasite as a material for high temperature microsystem", *Sensors and Actuators,* pp. 393-396, 2006.

[5] H. Fritze, H. L. Tuller, H. Seh, and G. Borchardt, "High temperature nanobalance sensor based on langasite", *Sensors and Actuators,* pp. 103-107, 2001.

[6] H. Fritze, H. Seh, H. L. Tuller, and G. Borchardt, "Operation limits of langasite high temperature nanobalances", *Journal of the European Ceramic Society*, pp. 1473-1477, 2001.

[7] H. Seh, H. L. Tuller, and H. Fritze, "Langasite for high-temperature acoustic wave gas sensors", *Sensors and Actuators,* pp. 169-174, 2003

[8] Y. K. Kim, Amplitude-Frequency effects of Y-cut langanite and langatate, *Proc. 2003 Ann. Symp. On Freq. Cont.,* pp. 631-636, 2003.

[9] R. L. Filler, "The amplitude-frequency effect in SC-cut resonators", *Proc. 39th Annu. Symp. Freq. Contr.,* pp. 311-316, 1985.

Effect of DC Electric Field on the Dispersive Characteristics of Acoustic Waves in Piezoelectric Layered Structure

S.I. Burkov, O.P. Zolotova
Condensed Matter Physics Department
Siberian Federal University
Krasnoyarsk, Russia
E-mail: sergbsi@gmail.com

B.P. Sorokin
Technological Institute for Superhard and Novel Carbon Materials
Troitsk, Moscow region, Russia
E-mail: bpsorokin2@rambler.ru

***Abstract*—Boundary conditions for a layered piezoelectric structure under the influence of a uniform dc electric field have been obtained. Dispersive dependences and anisotropy of phase velocities, electromechanical coupling coefficients, power flow angle, and controlling coefficients of dc electric field influence as a functions of h×f product for the Rayleigh and Love modes in bismuth germanium oxide/fused silica and langasite/fused silica layered structures have been calculated.**

I. INTRODUCTION

Development of acoustoelectronics (AE) is now concerned with the design of complicated structures such as layered piezoelectric devices [1, 2]. For example, by this way there is opened an attractive possibility to include the non-piezoelectic substrate with good acoustical properties into device meanwhile the acoustic wave can be excited using the piezoelectric film/plate. However, to extend the class of devices it is desirable to obtain the ability to control their parameters during the wave propagation process. On the other side the choice of optimal film/crystalline layer thickness h is important to obtain a given frequency characteristics of device [3]. In present paper for a number of layered piezoelectric structures the influence of dc electric field on the dispersive parameters of Rayleigh and Love (SH) waves which are the functions of h×f product has been examined (f denotes a wave frequency). As a model structures the crystalline layers of strong piezoelectric crystal bismuth germanium oxide ($Bi_{12}GeO_{20}$, BGO) or langasite ($La_3Ga_5SiO_{14}$, LGS) on fused silica have been investigated. Langasite is a most promising crystal in a practical sense because it has a good combination of properties, such as thermal stability, low attenuation of high frequency elastic waves and greater electromechanical coupling coefficients (EMCC) compared with quartz [4].

II. ELASTIC WAVE PROPAGATION IN A LAYERED PIEZOELECTRIC MEDIUM UNDER THE INFLUENCE OF HOMOGENEOUS DC ELECTRIC FIELD)

Let us choose the working orthogonal coordinate system in which the axis X_3 is directed along the outward normal to the surface layer occupying the space of $0 \le X_3 \le h$, and the X_1 axis coincides with the wave propagation direction. For small-amplitude plane waves propagation under the influence of dc electric field on the piezoelectric crystal the equations of motion and electrostatics and the state equations of the piezoelectric medium should have the form:

$$\begin{aligned} &\rho_0 \ddot{\tilde{U}}_I = \tilde{\tau}_{IK,K}, \ \tilde{D}_{M,M} = 0, \\ &\tilde{\tau}_{IK} = C^*_{IKPQ}\tilde{\eta}_{PQ} - e^*_{NIK}\tilde{E}_N, \\ &\tilde{D}_N = e^*_{NIK}\tilde{\eta}_{IK} + \varepsilon^*_{MN}\tilde{E}_M. \end{aligned} \qquad (1)$$

In equation (1) the following notation has been used: ρ_0 is the crystal density in the undeformed (initial) state, and $\tilde{U}_I$ is the vector of dynamic elastic displacements. Here and below summation over twice repeated indices is meant. Effective elastic, piezoelectric and dielectric constants which are the linear functions of dc electric field are defined by:

$$\begin{aligned} &C^*_{ABKL} = C^E_{ABKL} + \left(C^E_{ABKLQR}d_{JQR} - e_{JABKL}\right)M_J E, \\ &e^*_{NAB} = e_{NAB} + \left(e_{NABKL}d_{JKL} + H_{NJAB}\right)M_J E, \\ &\varepsilon^*_{NM} = \varepsilon^\eta_{NM} + \left(H_{NMAB}d_{PAB} + \varepsilon^\eta_{NMP}\right)M_P E. \end{aligned} \qquad (2)$$

Here E is the value of the dc electric field, M_J is the unit electric field vector, and C^E_{ABKLQR}, e_{NABKL} and H_{NMAB} are the nonlinear elastic, piezoelectric and electrostrictive constants (material tensors).

Acoustic wave propagation in a piezoelectric plate under the action of E must satisfy the relevant boundary conditions. Thus, the normal components of stress tensor must be equal

to zero on the free surface. A continuity of the tangential components of electric field to the interface between piezoelectric layer vacuum should be provided by the condition of continuity of electric potential. Also the equality of normal components of stress tensor τ_{IK} and the continuity of the electric potential φ at the boundary "crystalline layer – substrate" at $X_3 = 0$ should be ensured. As a consequence all these conditions can be written as:

$$\begin{aligned}
&\tau_{3A}^{(1)} = 0\big|_{x_3=h}, \qquad D_3^{(1)} = D^{(\text{вак})}\big|_{x_3=h};\\
&\tau_{3A}^{(1)} = \tau_{3A}^{(2)}\big|_{x_3=0}, \qquad D_3^{(1)} = D_3^{(2)}\big|_{x_3=0};\\
&\varphi^{(1)} = \varphi^{(2)}\big|_{x_3=0}, \qquad U_A^{(1)} = U_A^{(2)}\big|_{x_3=0}.
\end{aligned} \tag{3}$$

As a result of substitution the solutions in the form of homogeneous plane waves into the boundary conditions (3) one can obtain a system of equations for calculating the acoustic waves parameters of a layered structure in a generalized form when a layer and substrate are the piezoelectric crystals:

$$\begin{aligned}
&\sum_{m=1}^{4}\Big\{b_m\Big[\Big(C_{I3KL}^{*(1)} + 2d_{AKF}^{(1)}C_{3IFL}^{(1)E}M_A E\Big)k_L^{(m)}\alpha_K^{(m)} + e_{3PI}^{*(1)}k_p^{(m)}\alpha_4^{(m)}\Big] -\\
&-\sum_{n=1}^{8}a_n\Big[\Big(C_{I3KL}^{*(2)} + 2d_{AKF}^{(2)}C_{3IFL}^{(2)E}M_A E\Big)k_L^{(n)}\alpha_K^{(n)} + e_{3PI}^{*(2)}k_p^{(n)}\alpha_4^{(n)}\Big]\Big\} = 0,\\
&\sum_{m=1}^{4}\Big\{b_m\Big[\Big(e_{3KL}^{*(1)} + 2d_{JKP}^{(1)}e_{3PL}^{(1)}M_A E\Big)k_L^{(m)}\alpha_K^{(m)} + \varepsilon_{3K}^{*(1)}k_K^{(m)}\alpha_4^{(m)}\Big] -\\
&-\sum_{n=1}^{8}a_n\Big[\Big(e_{3KL}^{*(2)} + 2d_{JKP}^{(2)}e_{3PL}^{(2)}M_A E\Big)k_L^{(n)}\alpha_K^{(n)} + \varepsilon_{3K}^{*(2)}k_K^{(n)}\alpha_4^{(n)}\Big]\Big\} = 0,\\
&\sum_{m=1}^{4}\Big[U_I^{(1)(m)}b_m\Big] - \sum_{n=1}^{8}\Big[U_I^{(2)(n)}a_n\Big] = 0,\\
&\sum_{n=1}^{8}\Big[-a_n\Big(C_{3KPI}^{*(2)} + 2d_{APF}^{(2)}C_{3KFI}^{(2)E}M_A E\Big)k_I^{(n)}\alpha_p^{(n)} -\\
&-a_4 e_{P3K}^{*(2)}k_p^{(n)}\alpha_4^{(n)}\Big]\exp\big(ik_3^{(n)}h\big) = 0,\\
&\sum_{n=1}^{8}\Big[-a_n\Big(e_{3KL}^{*(2)} + 2d_{JKP}^{(2)}e_{3PL}^{(2)}M_J E\Big)k_L^{(n)}\alpha_K^{(n)} -\\
&-a_4\Big(\varepsilon_{3K}^{*(2)}k_K^{(n)} - i\varepsilon_0\Big)\alpha_4^{(n)}\Big]\exp\big(ik_3^{(n)}h\big) = 0.
\end{aligned} \tag{4}$$

Numerical superscripts (1) and (2) denote the substrate and layer terms, respectively; $\alpha_K^{(m)}$ and b_m are amplitudes and weight coefficients of m-th partial wave (m = 1, ..., 4) in the piezoelectric substrate, respectively; $\alpha_K^{(n)}$ and a_n are the same ones of n-th partial wave (n = 1, ..., 8) in the piezoelectric layer.

Note that the equations (4) for the boundary conditions have been derived at the assumptions about the influence of homogeneous dc electric field on the crystal layer without edge effects. In the obtained equations it has been taken into account all changes in the configuration of an anisotropic continuum, in particular, changes of the sample form (geometric nonlinearity) due to direct piezoelectric effect, and changes of the material constants (2) (physical nonlinearity) under the influence of strong dc electric field.

For the purpose of this paper there was selected a model structure which was consisted of a piezoelectric crystalline layer mounted on an isotropic substrate (fused silica). Thus the equations (4) should be specified as

$$\begin{aligned}
&\sum_{m=1}^{3}b_m C_{I3KL}^{(1)}k_L^{(m)}\alpha_K^{(m)} -\\
&-\sum_{n=1}^{8}a_n\Big[\Big(C_{I3KL}^{*(2)} + 2d_{AKF}^{(2)}C_{3IFL}^{(2)E}M_A E\Big)k_L^{(n)}\alpha_K^{(n)} + e_{3PI}^{*(2)}k_p^{(n)}\alpha_4^{(n)}\Big] = 0,\\
&b_4\varepsilon_{3K}^{(1)}k -\\
&-\sum_{n=1}^{8}a_n\Big[\Big(e_{3KL}^{*(2)} + 2d_{JKP}^{(2)}e_{3PL}^{(2)}M_A E\Big)k_L^{(n)}\alpha_K^{(n)} + \varepsilon_{3K}^{*(2)}k_K^{(n)}\alpha_4^{(n)}\Big] = 0,\\
&\sum_{m=1}^{4}U_I^{(1)(m)}b_m - \sum_{n=1}^{8}U_I^{(2)(n)}a_n = 0,\\
&\sum_{n=1}^{8}\Big\{-a_n\Big[\Big(C_{3KPI}^{*(2)} + 2d_{APF}^{(2)}C_{3KFI}^{(2)E}M_A E\Big)k_I^{(n)}\alpha_p^{(n)} -\\
&-a_4 e_{P3K}^{*(2)}k_p^{(n)}\alpha_4^{(n)}\Big]\Big\}\exp\big(ik_3^{(n)}h\big) = 0,\\
&\sum_{n=1}^{8}\Big[-a_n\Big(e_{3KL}^{*(2)} + 2d_{JKP}^{(2)}e_{3PL}^{(2)}M_J E\Big)k_L^{(n)}\alpha_K^{(n)} -\\
&-a_4\Big(\varepsilon_{3K}^{*(2)}k_K^{(n)} - i\varepsilon_0\Big)\alpha_4^{(n)}\Big]\exp\big(ik_3^{(n)}h\big) = 0.
\end{aligned} \tag{5}$$

III. Dispersive Characteristics of Elastic Waves in Layered Piezoelectric Structures

To define in the linear approximation a measure of dc electric field influence on the phase velocities of Rayleigh and Love waves it is convenient to introduce the controlling coefficients as:

$$\alpha_v = \frac{1}{v(0)}\left(\frac{\Delta v}{\Delta E}\right)_{\Delta E\to 0}. \tag{6}$$

On the basis of equations (1) - (5) in a layered piezoelectric structures for a few modes of Rayleigh and Love waves the calculation of such parameters as the phase velocities, EMCCs and α_v coefficients has been carried out. The material coefficients of linear and nonlinear electromechanical properties of BGO and LGS were taken from [4, 5].

Figure 1 shows the dispersion curves of phase velocities, EMCCs and α_v coefficients of Rayleigh and Love waves depending on the h×f value when waves are propagated along the [100] direction of (001) BGO/fused silica. In this case pure piezoactive Love wave and non-piezoactive Rayleigh wave are propagated. It is clear that the interval of phase velocities of Love waves lies between the values of the shear waves velocities of fused silica and BGO. Rayleigh wave phase velocity at h×f increasing tends to the Rayleigh wave phase velocity of BGO (1625.92 m/s). The maximum value of $K^2 = 2.9$ % at h×f = 1000 m/s takes place for zero Love mode. When dc electric field is applied along the E||[010], i.e. orthogonally to the sagittal plane, the modes of elastic waves remain the pure ones. The α_v coefficient of Love wave at h×f increasing tends to $\alpha_v = -3\cdot10^{-11}$ m/V for the slow shear bulk acoustic wave (BAW) of BGO, and α_v coefficient of all the Rayleigh type waves at h×f increasing tends to $\alpha_v = -6.9\cdot10^{-12}$ m/V value as the same one for the fast shear BAW of BGO.

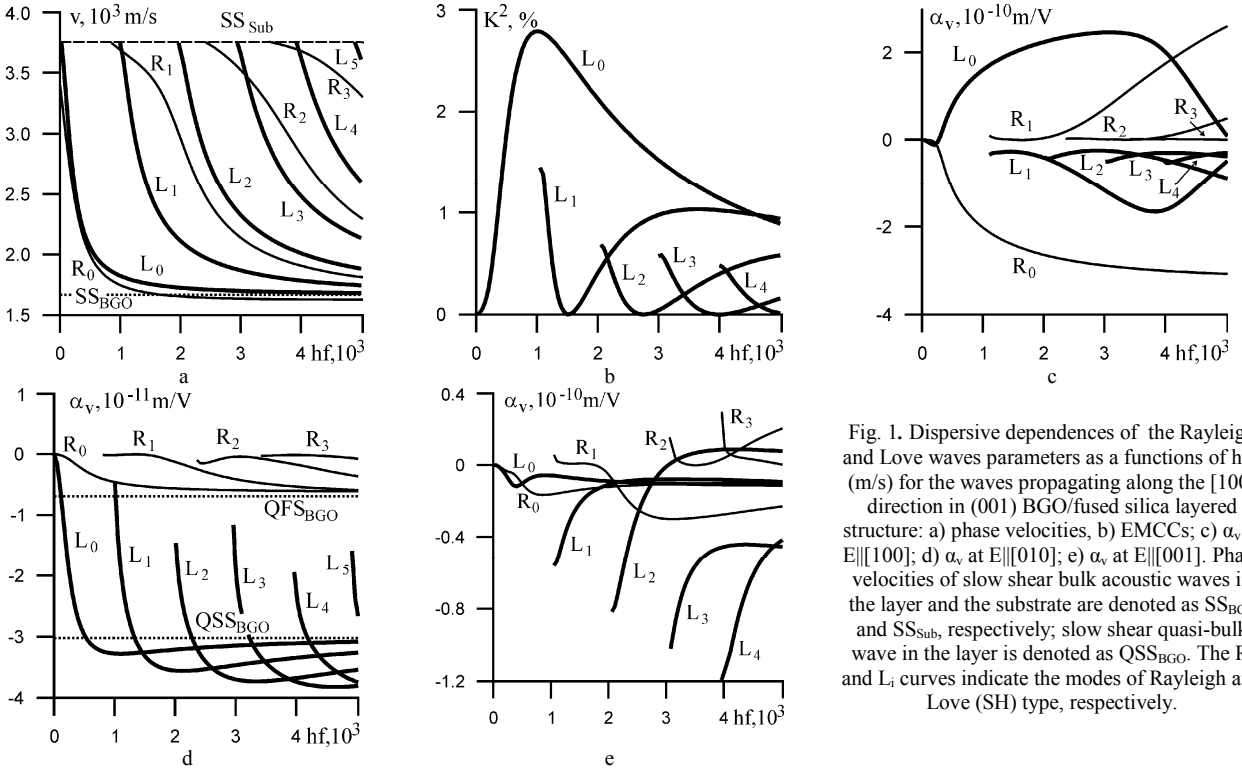


Fig. 1. Dispersive dependences of the Rayleigh and Love waves parameters as a functions of h×f (m/s) for the waves propagating along the [100] direction in (001) BGO/fused silica layered structure: a) phase velocities, b) EMCCs; c) α_v at E||[100]; d) α_v at E||[010]; e) α_v at E||[001]. Phase velocities of slow shear bulk acoustic waves in the layer and the substrate are denoted as SS_{BGO} and SS_{Sub}, respectively; slow shear quasi-bulk wave in the layer is denoted as QSS_{BGO}. The R_i and L_i curves indicate the modes of Rayleigh and Love (SH) type, respectively.

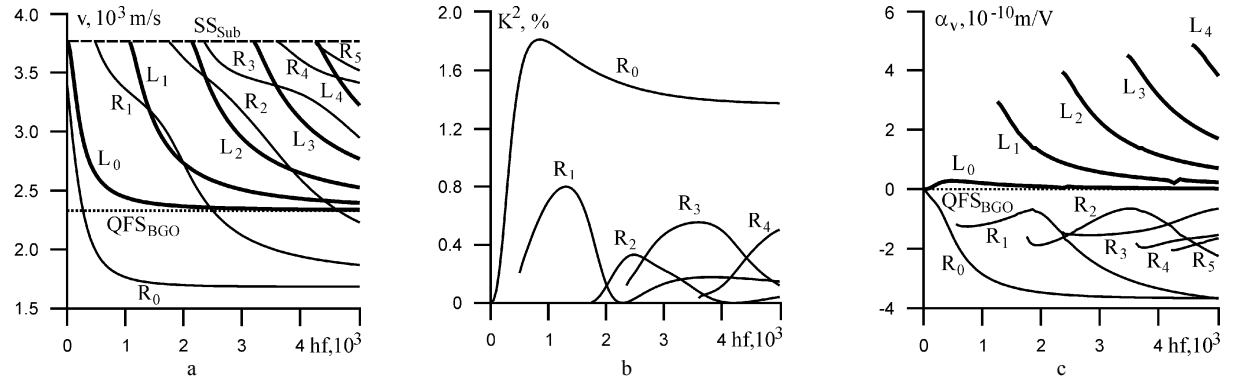


Fig. 2. Dispersive dependences of the Rayleigh and Love waves parameters as a functions of h×f (m/s) for the waves propagating along the [110] direction in (001) BGO/fused silica layered structure: a) phase velocities, b) EMCCs; c) α_v at E||[001].

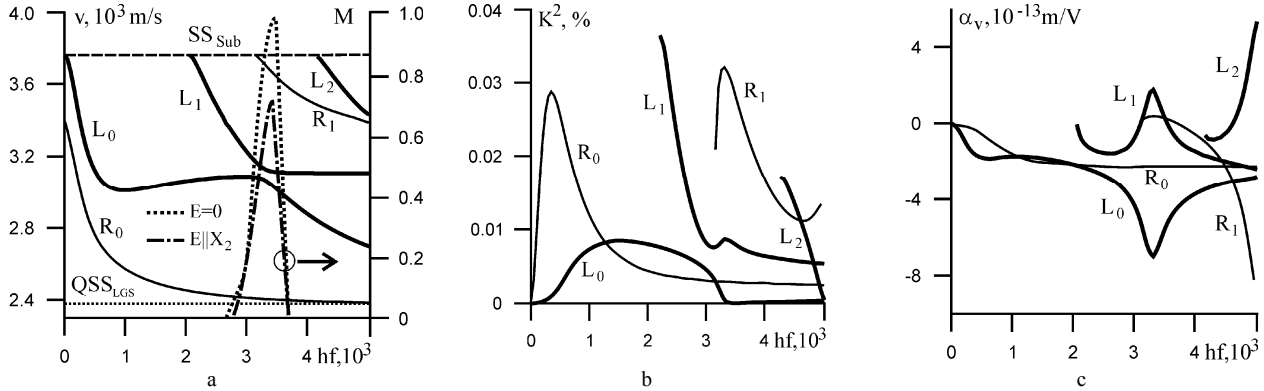
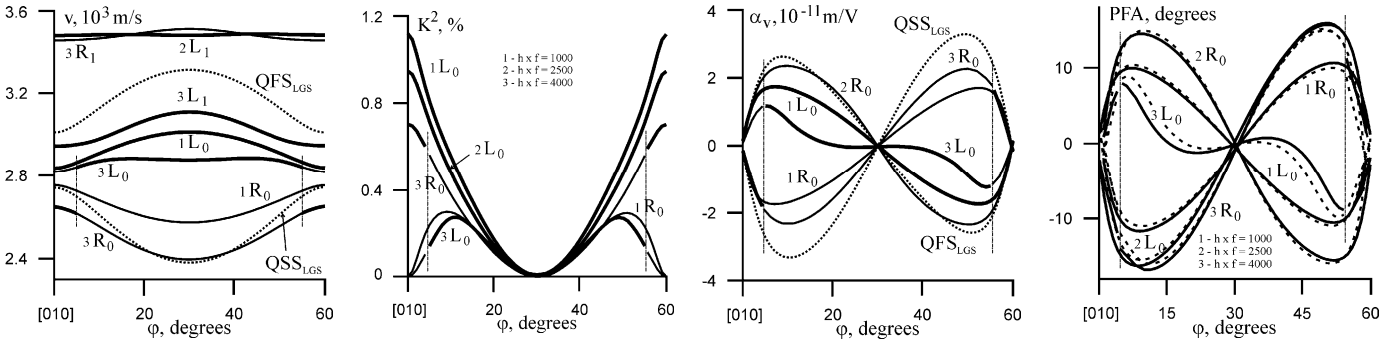


Fig. 3. Dispersive dependences of the Rayleigh and Love waves parameters as a functions of h×f (m/s) for the wave propagating along the [100] direction in (001) LGS/fused silica layered structure: a) phase velocities; b) EMCCs; c) α_v at E||[001].



a b c d

Fig. 4. Anisotropy of Rayleigh and Love wave parameters for the wave propagating in the (001) LGS/fused silica layered structure for a few values of h×f (m/s): a) phase velocities; b) EMCCs; c) α_v at E||X_3; d) PFA at E = 0 (solid lines) and E||X_1 (dashed lines). Curves denoted as QSS_{LGS} and QFS_{LGS} are associated with quasi-bulk slow and fast shear waves in a layer, respectively.

When dc electric field is applied along the E||[100] or E||[001] elastic waves modes will turn into a generalized form. Note that if dc electric field coincides with the wave propagation direction there should be observed a maximal E influence on the phase velocities especially for R_0 and L_0 modes (fig. 1c). Thus the maximal α_v value for the zero Love mode of $2.5 \cdot 10^{-10}$ m/V at h×f = 3050 m/s is achieved at E||[100].

Figure 2 shows the dispersion curves of phase velocities, EMCCs and α_v coefficients of Rayleigh and Love waves depending on the h×f value when waves are propagated along the [110] direction of (001) BGO/fused silica. In this case a Rayleigh wave possesses a piezoelectrical activity. The range of variation of the Love waves phase velocities lies between the phase velocities of shear wave in fused quartz and fast quasi-shear wave in bismuth germanium oxide. The maximal value of K^2 = 1.8 % at h×f = 850 m/s exists for R_0 zero mode of Rayleigh wave. In the most practically significant version of dc electric field application E||[001] the hybridization between the modes is negligibly small [6]. The α_v coefficient of Love wave under h×f increasing tends to $\alpha_v = -5.6 \cdot 10^{-13}$ m/V for fast shear acoustic wave in the BGO layer.

In the [100] propagation direction of (001) LGS/fused silica layered structure there are only Rayleigh and Love waves of generalized type (fig. 3). The small EMCCs indicate on the low piezoelecrical activity of these waves (fig. 3b). An unique feature is the effect of hybridization between zero and first modes of SH waves (generalized Love waves). Either way the applied electric field reduces the effect of hybridization leading to a "splitting" of phase velocity values in this area and reducing the hybridization coefficient M [6]. Maximal values of $\alpha_v = -3.9 \cdot 10^{-12}$ m/V and $\alpha_v = 2.4 \cdot 10^{-12}$ m/V occur at the h×f = 3300 m/s close the center of hybridization area for the zero and first SH modes, respectively. The α_v coefficient increasing for the second mode with h×f increasing is also caused by a similar hybridization effect arising between the L_1 and L_2 modes.

Figure 4 shows the anisotropy characteristics of elastic waves propagating in (001) LGS/fused silica structure. Vertical dashed lines mark the points of hybridization that in this case occur between the R_0 and L_0 modes when h×f = 4500 m/s. The counting of φ angle is produced from the initial [010] direction along which the pure modes of Rayleigh and Love waves are propagated. In this direction there is a maximum of K^2 = 1.1 % at h×f = 1000 m/s for the L_0 mode. Power flow angle (PFA) has a maximal value of 16,4° for L_0 mode at h×f = 2500 m/s and of -16,8° for R_0 mode at h×f = 4000 m/s. Application of dc electric field along the X_3 direction which is three-fold axis leads to the symmetry decreasing up to point symmetry group 3 and as a result the value of PFA is increased insignificantly (up to 0,2°), but E application along X_1 axis leads to PFA increasing up to 2°.

IV. Conclusion

Using the obtained theoretical results one can analyze in detail the dispersive behavior and anisotropy of acoustic modes in a piezoelectric layered structure under the influence of homogeneous dc electric field if constants of linear and nonlinear electro-mechanical properties of the crystal are known. It was shown that for different variants of an electric field application interaction can occur between of different elastic modes. BGO layered structure has a better EMCC than that for LGS layered structure. The data obtained may be of interest in terms of study of AE devices geometry controlling by the dc electric field.

References

[1] S.J. Ippolito, A. Ponzoni, K. Kalantar-Zadeh, W. Wlodarski, E. Comini, G. Faglia, and G. Sberveglieri, "Layered WO_3/ZnO/36° $LiTaO_3$ SAW gas sensor sensitive towards ethanol vapour and humidity," Sensors and Actuators B, vol. 117, pp. 442–450, 2006.

[2] G. Wingqvist, J. Bjurstrom, and I. Katardjiev, "Shear mode AlN thin film electroacoustic resonator for biosensor applications," Trans. IEEE Ultrasonics Symposium, pp. 50-53, 2005.

[3] Y. Wang, K. Hashimoto, T. Omori, and M. Yamaguchi, "Change in piezoelectric boundary acoustic wave characteristics with overlay and metal grating materials," Trans. on Ultrason., Ferroel. and Freq. Control, vol. 57, pp. 16-22, January 2010.

[4] K.S. Aleksandrov, B.P. Sorokin, and S.I. Burkov, Effective Piezoelectric Crystals for Acoustoelectronics, Piezotechnics and Sensors, vol. 2. Novosibirsk: SB RAS Publishing House, 2008, 429 p.

[5] K.S. Aleksandrov, B.P. Sorokin, and S.I. Burkov, Effective Piezoelectric Crystals for Acoustoelectronics, Piezotechnics and Sensors, vol. 1. Novosibirsk: SB RAS Publishing House, 2007, 501 p.

[6] .E. Kuznetsova, B.D. Zaitsev, A.A. Teplykh, and I.A. Borodina, "Hybridization of acoustic waves in piezoelectric plates," Acoustical Physics, vol. 53, pp. 64-69, January 2007.

Determination of the mode shapes of energy trapping resonators using X-ray topography.

B. Capelle, J. Détaint & Y. Epelboin.

IMPMC Universités Pierre et Marie Curie -P6, Paris Diderot-P7 et CNRS.
4, Place Jussieu. 75005 Paris. France.

Bernard.Capelle@impmc.upmc.fr

***Abstract*— This paper considers the determination of the mode shapes using x-ray topography and the application of this method to resonant piezoelectric devices made of various materials (Quartz, LGS, Lithium tantalate). The basis of the method is recalled briefly and its application to thickness mode resonators or filters is particularly considered. A comparison of the observed mode shapes with those computed using the "essentially thickness mode" theory of H.F.Tiersten indicates that this theory give very good results concerning the modes shapes and the properties of the thickness mode devices even when materials presenting large coupling coefficients are used.**

INTRODUCTION.

X-ray topography is a powerful mean to observe the vibration modes of the piezoelectric devices. Many years ago, Spencer et al. [1] have given an interesting overview of the application of the Lang topography [2] to determine the mode shapes of quartz resonators and monolithic filters. Several authors have proposed approximate expressions of the diffracted intensity in presence of vibration modes that can have some validity in particular conditions. However, the image formation occurring due to x-ray diffraction by deformed crystals is a very complex phenomenon governed by the dynamical theory of x-ray diffraction and the Takagi-Taupin partial derivatives equations. A complete description of this theory can be found in the book of A. Authier [3]. Numerical simulations of diffraction images obtained with vibrating crystals, made in solving these equations, have shown that X-ray topography can provides accurate images of the vibration modes of the piezoelectric devices if favorable diffraction conditions are chosen [4]. In other cases, the relation the between the diffracted intensity and the displacement may be quite complex. In Figure 1 (from [4]), we compare a computed topograph image obtained in solving the Takagi-Taupin equations and the corresponding experimental image (AT-cut quartz 19MHz, fundamental mode thickness 80µm, (-2,1.0) diffraction at λ#0.70Å). These diffraction conditions are particularly favorable to obtain a good observation of the mode shape (near linear increase of the diffracted intensity with the displacement in time integrated conditions). In such conditions, X-ray topography gives more accuracy and more information in a single experiment than the other techniques. The contrasts recorded in the topographs are dependant on the product of the displacement vector by the diffraction vector, on the values of partial derivatives of the displacement and on several parameters characterizing the diffraction process. X-ray topography allows observing the different components of the displacement, often separately.

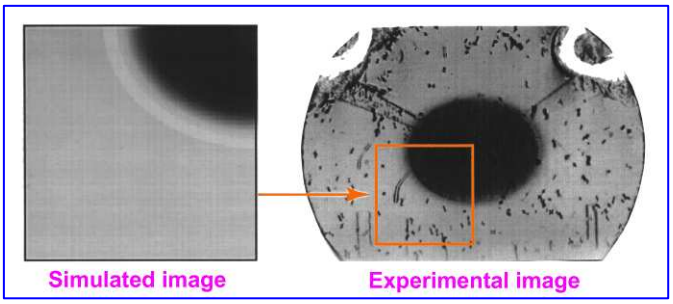


Figure 1. Comparison of computed and experimental image.

They can also be determined simultaneously if white beam topography is used. The X-ray films or the nuclear plates have a large dynamic (up to four decades) and X-ray counting can be used if necessary. The shapes so recorded of the thickness modes of energy trapping devices are most often extremely similar to those predicted by the theory proposed by H.F. Tiersten et al. [5-6]. The topography experiments reported here were made using synchrotron radiation at the LURE (Orsay - France) and at ESRF (Grenoble - France) using beam lines ID19 and BM05.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

I. OBSERVED AND COMPUTED MODES IN AT-CUT QUARTZ.

A. Plane resonators with conventional electrodes:

Here after we compare the mode shapes observed for 12 MHz plane AT-cut quartz resonators (12.7mm diameter, about 134.5µm thick) having various electrode geometries, but the same electrode area. The topographs of figure 2 and 3 were obtained using (-2,1.0) reflection with an X-ray wave length of 0.70Å. As already noticed these diffraction conditions are very favorable to obtain a nearly linear relationship between the displacement and the increase of the diffracted intensity (up to the saturation of the film, and to a certain level of excitation). At very large exciting currents the diffraction process change some how (kinematical diffraction) and other diffraction techniques have to be used to determine the mode shapes [7].

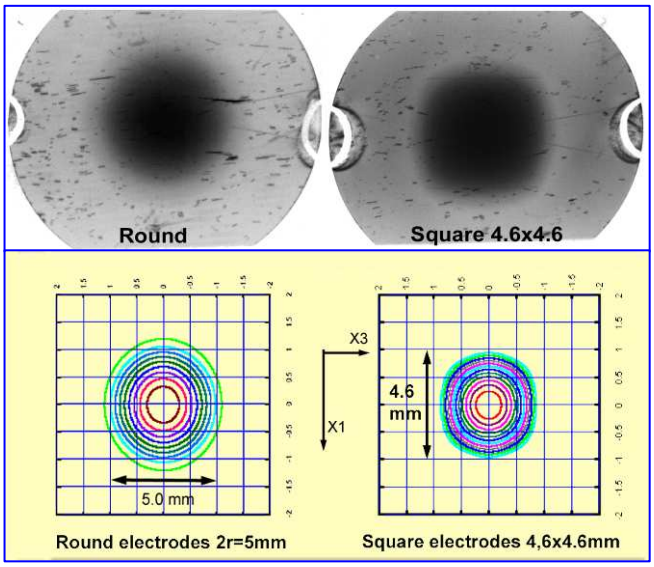


Figure 2. Observed and computed mode shapes of 12.MHz resonators (a)

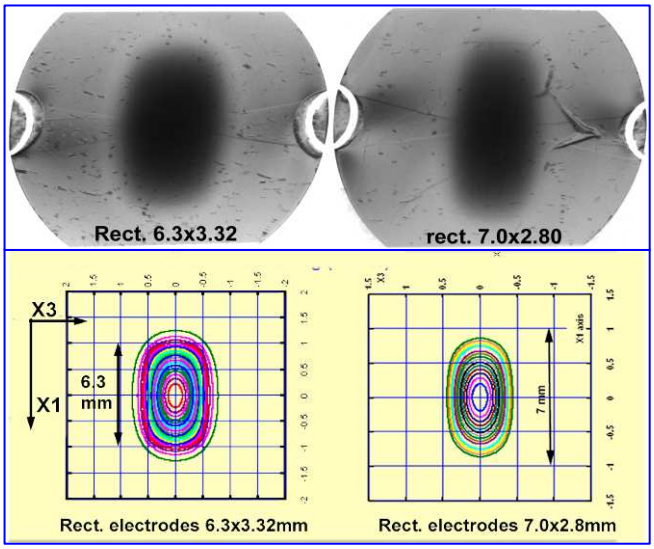


Figure 3. Observed and computed mode shapes of 12.MHz resonators (b)

Energy trapping of the modes in the electrode region occurs due to the electrical effect of the electrodes and to their mass (mass loading 0.96%). We can observe that the trapping effect is efficient for such values of the electrode dimensions and of the mass loading. The computed modes shapes (here normalized to 1 at the center of the electrodes) are very similar to the observed ones. A more detailed observation of the properties of these resonators reveals some difference concerning the resonance frequencies of the anharmonic modes and the values of the elements of the equivalent schemes. The equivalent inductance is some how more favorable (lower) for the round electrodes. This advantage results of a smaller efficiency of the corners of the square or rectangular electrodes.

Anharmonic modes: The dimensions of the electrodes chosen allow the existence of two anharmonic modes (except for the 7x2.8 electrodes that give only one anharmonic).

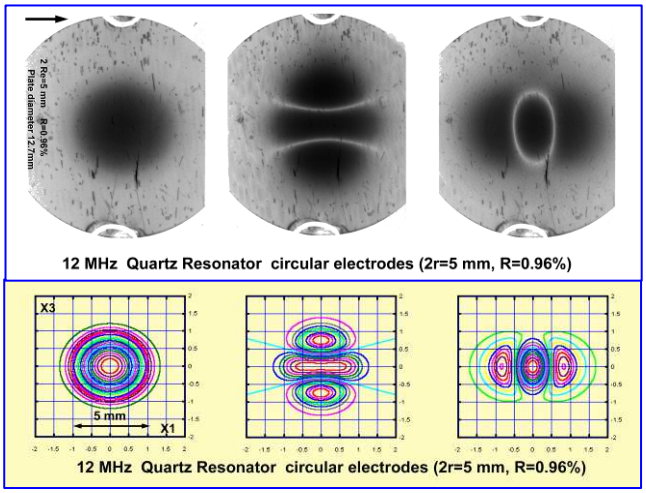


Figure 4. Mode shapes of the anharmonics (round electrodes).

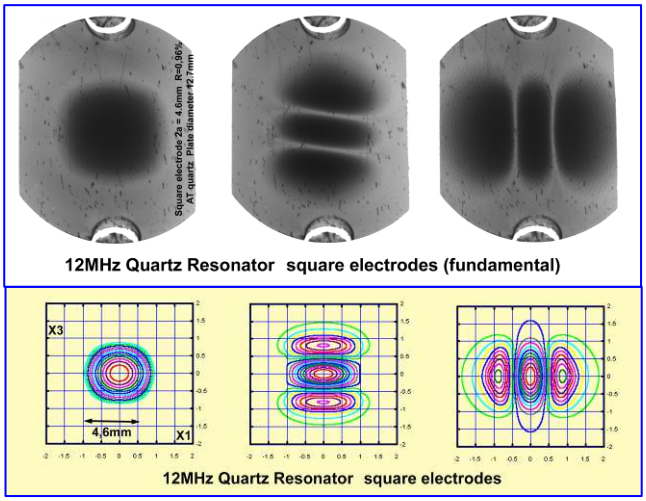


Figure 5. Mode shapes of the anharmonics (Square electrodes).

The modes corresponding to the round and the square electrodes are represented in figure 4 and 5 (same diffraction conditions as for figure 2 and 3). For some of these modes

slight differences appear between the experiment and the calculation, this is mostly the case for the 2nd mode of the resonator with square electrodes. The difference is probably due to a small defect of parallelism in the external region of the plate. These anharmonic modes appear to be more sensitive to imperfections and to the electrodes tabs than the first one, due to their larger extension. The theory elaborated by H.F. Tiersten indicates the existence, in most cases, of an intrinsic lateral anisotropy of the modes [8]. For the fundamental mode of the AT-cut of quartz this anisotropy is noticeable while it is weak for the 3rd and 5th overtone modes. It results of the ratio of the M_1 to the P_1 coefficients of the Tiersten partial derivative equations that governs the lateral dependence of the main displacement u_1 (M_1= 0.1098D+12, P_1=0.686D+11 N/m²). This anisotropy can be observed in the case of the round electrodes where the mode is slightly elongated in the x direction.

B. Electrodes taking into account the lateral anisotropy.

The use of elliptical electrodes with a particular axis ratio derived from the theory $([M_1/P_1]^{1/2}=1.265)$ leads to devices with more interesting properties, particularly for filter applications (less unwanted response and lower inductances at equal electrode surface, etc…). In this case the mode shape has the same eccentricity as the electrodes and all the anharmonic modes with 2p nodal line converge toward the same modes that have either p elliptical nodal lines or 2p radial nodal line (or a mix of that). The modes with radial nodal lines are not electrically excited (centric-symmetry in a particular coordinate system), except when some asymmetry exists (this is the case for the resonator represented in figure 6). This electrode geometry leads also to the maximum frequency separation between the main mode and the anharmonics. It is thus easier to suppress them by choice of the dimensions of the electrodes and/or of the mass loading. These properties make that well designed resonators using this electrode shape are more interesting to build filters

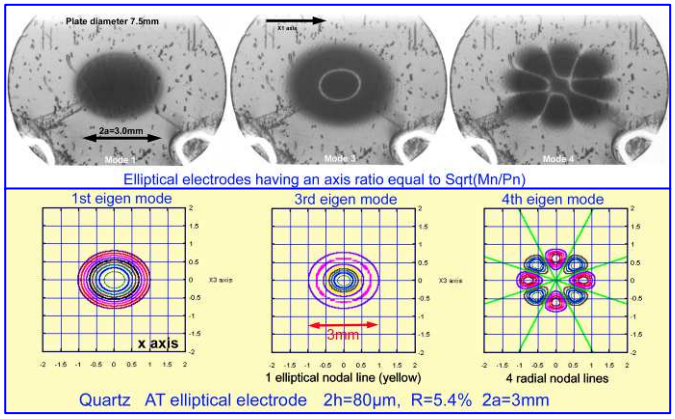


Figure 6. Modes shapes of a resonator using elliptical electrodes.

Figure 6 represents several modes of a resonator using a large elliptical electrode with an important mass loading made on a plate presenting a slight defect of symmetry (the electrode tabs are not directed along one axis of the ellipse); this allows the existence of some theoretically unexcited modes (here the mode with 4 straight nodal lines). However, the other observed modes shapes are very similar to the computed ones (same sample as for figure 1 and same diffraction conditions).

Another way to illustrate the lateral anisotropy of the fundamental modes of AT-cut quartz is given in figure 7 where we compare the mode shape produced by elliptical electrodes with the same axis ratio as considered above, either elongated in the correct direction (x1=x axis) or elongated in the x3=z' direction. It can be observed that in the first case the mode is elliptical with the same eccentricity as the electrodes while in the second case, it is nearly circular (same diffraction conditions as for figures 2 to 4).

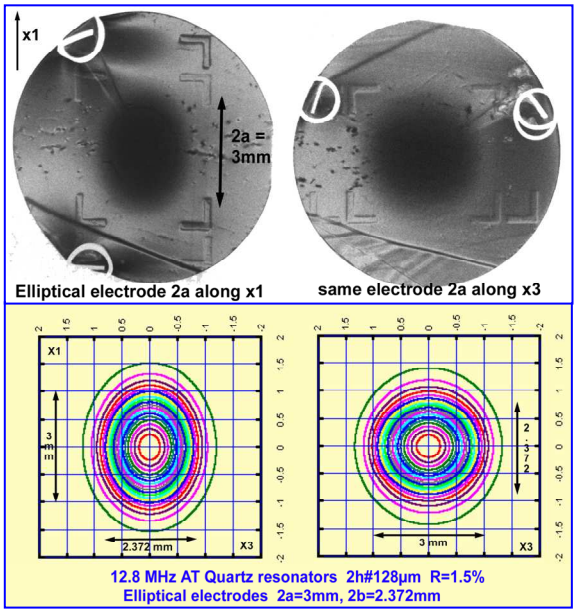


Figure 7. Elliptical electrodes elongated in the x or Z' directions.

II. LANGASITE RESONATORS

The theory proposed by Tiersten et al. makes use of some approximations considering a moderate piezoelectric effect. Most of these approximations concern in fact the cut-off frequencies of the different regions of the device that appear in the partial derivative equations. However these quantities are fundamentally one dimensional ones and they can be computed exactly for any piezoelectric material. Thus it is possible to expect good results of the application of the theory to the materials having medium or even large coupling coefficients. Some times ago [9], we have made calculations of the modes shapes of the overtone modes of plane and plano-convexe Y-cut langasite resonators. For several overtones, the results were not in very good agreement with the observations. At this time, we have noticed, for this material and also for gallium phosphate, that the computed mode shapes, and somehow the computed inductances were sensitive to the constant set used for the calculation. Since that time, these materials have been perfected and several new sets of material constants have been measured. New calculations were made using some of these more recent sets of material constants [10-11-12] for the same resonators using Y-cut LGS (thickness=0.432mm, electrode diameter D=3.5mm, plate

diameter #10mm). As can be observed in figure 8 and 9, a much better agreement is obtained, even for the overtones modes of large ranks using the constants of reference [11] (Experimental conditions (-2,1.0) diffraction, λ#0.70Å, plate thickness 0.432mm (more absorbing crystal)). The sensitivity of several important parameters appearing in the energy trapping theory to the values of the elastic constants noticed for LGS and gallium phosphate results much probably of the fact that these quantities imply many products of differences between constants or products of constants

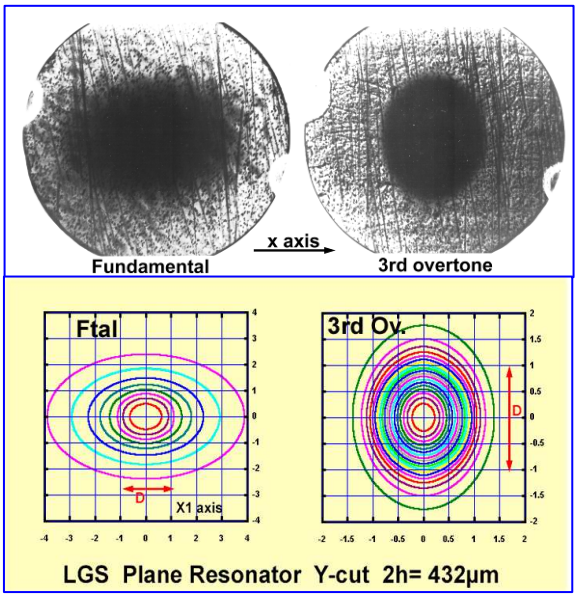


Figure 8. Fundamental and 3rd overtone modes of a LGS plane resonator

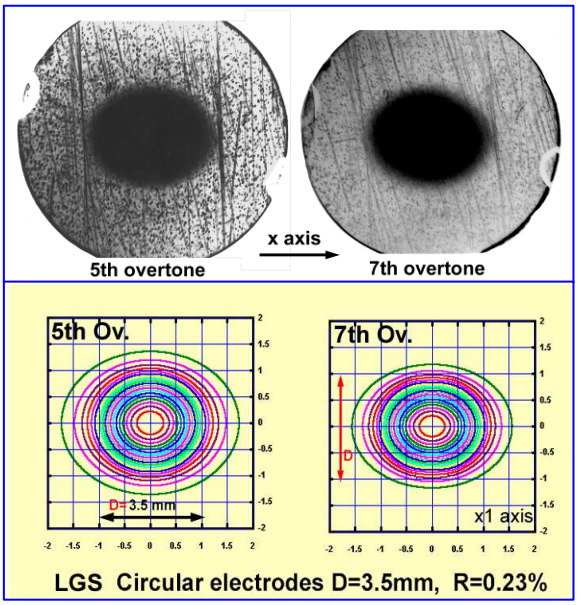


Figure 9. 5th and 7th overtone modes of a Y-cut LGS plane resonator.

. Another observation concerns the quite large anisotropy of the overtone modes of Y-cut LGS resonators.

Similarly, a better agreement between the observed and the computed modes can also be obtained for gallium phosphate resonators using the recently determined constants [13-14].

III. LITHIUM TANTALATE RESONATORS.

We have observed that the theory of essentially thickness modes of H.F. Tiersten leads to realistic results for many orientations of this material (X-cut, Y rotated cuts, doubly rotated cuts, etc…). The mode shapes computed with different sets of constants [115-16-17] have only minor differences. Hereafter, we compare the results given by this theory to the mode shapes observed using x-ray topography for x-cut resonators. Figure 10 represents the electrical response of a 27.7 MHz fundamental X-cut resonator having small electrodes and a small mass loading (aluminum electrodes) that give no trapped anharmonics for the fast shear mode (filter type resonators).

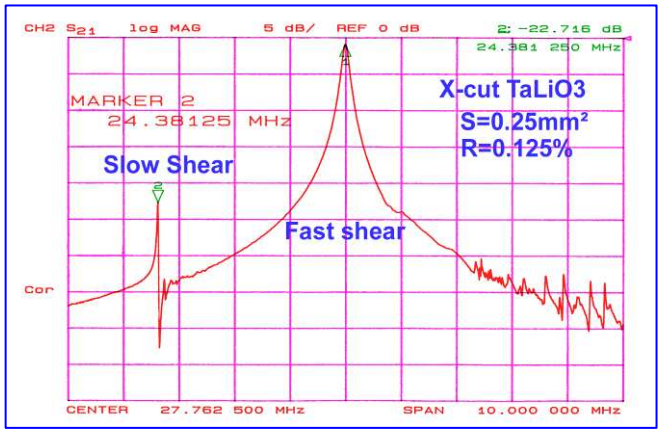


Figure 10. Response of an X-cut resonator (with compensation of Co)

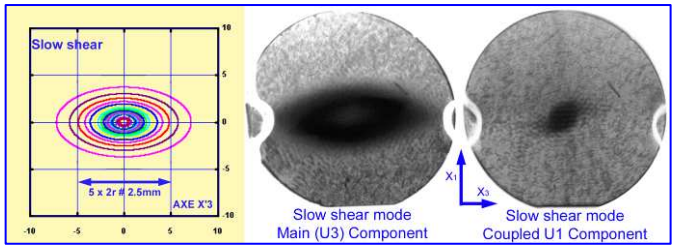


Figure 11. Slow shear mode of an X cut resonator

The slow shear mode is sufficiently close to the fast shear, so that theses mode are coupled and that some excitation of the u1 component (of the fast shear) occurs at the resonance frequency of the slow shear. This can be observed experimentally by X-ray topography (figure 11) or in considering the response curve obtained with compensation of the static capacitance (the slow shear mode is in the "slope" of the response of the fast shear). This is a forced mode effect included in the theory. It results mostly of the very large coupling coefficient of the fast shear mode (very broad response curve). This effect can be taken into account theoretically in considering the forced mode (linear combination of all the Eigen-modes verifying the boundary conditions) at this frequency, or more simply in connecting in parallel the motional arms of the equivalent schemes of both

modes. This coupling appears in the topographs taken at the resonance frequency of the slow shear mode (figure 11). The observed and the computed slow shear modes have a large and very similar anisotropy. It should also be noticed that the slow shear is very weakly trapped by small electrodes presenting a small mass loading and that its experimental mode shape is probably slightly modified, in the most external region, by the mounting clips of the plate and the conducting cement (plate diameter 5.5mm). The coupled fast shear mode excited at the resonance of the slow shear seems correctly confined in these conditions but it is also coupled to a weak plate mode (flexure). The mode shapes of the anharmonics of the fast shear were observed in a similar resonator made with larger square electrodes (S#0.66mm², 2h#70µm, plate diameter 5.5mm). Due to the large coupling coefficient (about 45%) two of them have a resonance frequency lower than the "nominal" anti-resonance frequency. The computed modes shapes are in reasonable agreement with the computed ones if the modest Q factor of mode 2 is considered. It should be noticed that for all the orientations studied (except for the Z-cut), an important lateral anisotropy of the trapped mode modes was observed in lithium tantalate. In some cases, concerning for example the doubly rotated cuts giving a zero temperature coefficient for the overtone modes, no trapping may exist in one direction for a particular overtone rank.

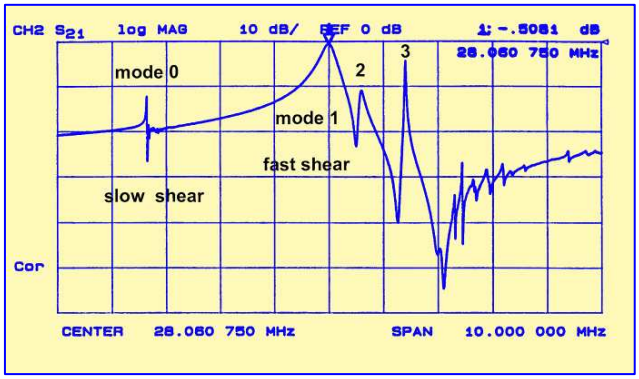


Figure 12. Response curve of an X-cut resonator made with large electrodes

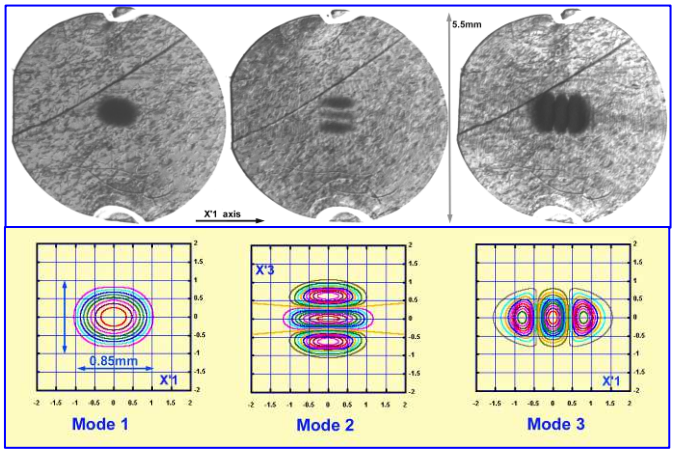


Figure 13. Mode shape of the fundamental and of its two first anharmonics.

IV. UNWANTED STATIONARY AND PROGRESSIVE PLATE MODES

The theory of the essentially thickness modes takes into account the possibility of (small) components other than the main displacement but neither that of strong resonant plate modes nor that of progressives modes. The resonant plate modes are most often very weakly coupled to the thickness modes existing in plates having a large ratio of the lateral dimensions to the thickness. They are more frequent in devices using rectangular plates and/or rectangular electrodes and in beveled resonators with a moderate ratio of the lateral dimensions to the thickness. They are very difficult to avoid in miniaturized devices where this ratio is small, at least, for one lateral dimension. Several approximate or exact theories are able to consider the coupling of the plate modes with the thickness modes. However there is a difficulty to describe theoretically the fixation conditions of the plate in a manner which represent accurately those used in real devices. Most often the actual plate modes (resonant or progressive) are extremely dependant to the fixation conditions. Nevertheless, these theories are extensively used for the design of miniaturized resonators.

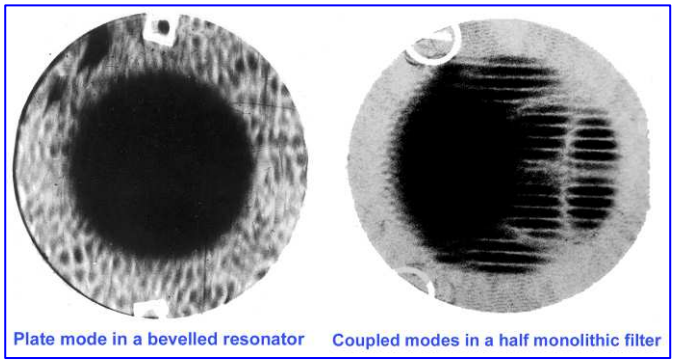


Figure 14. Plate modes coupled to the fundamental thickness shear.

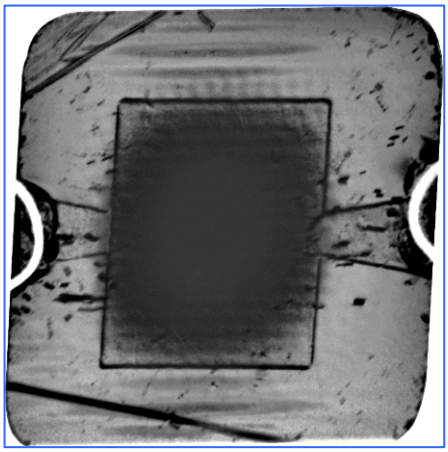


Figure 15. Flexure coupled to thickness shear in a plane rectangular AT-cut quartz resonator

Some examples of coupled plate modes are given in figure 14. One was observed in a beveled resonator. The other was observed in a monolithic filter (only one pair of electrodes is excited here). Our observations have shown that the coupled

modes are more frequent in plates having a rectangular shape and/or using rectangular electrodes (figure 15). In quartz resonators with one small lateral dimension, it is usual to choose the plate mode coupled to the thickness mode so to keep an acceptable Q factor and a good temperature coefficient.

A. *Progressive coupled modes*

We have observed progressive modes coupled to the thickness shear in all the fundamental mode plane resonators studied by time resolved (synchronous) topography. They carry a noticeable energy in devices made with high coupling materials and, in all cases, including quartz; they explain the modest Q.F products observed for fundamental mode devices. We consider that they can be detected in resonators having a Q factor lower than a few hundred thousands. The observation of these progressive plate modes by stroboscopic X-ray topograph indicates clearly that they carry energy towards the fixations of the device. The energy loss per period and the stored energy can be estimated roughly using x-ray topographs (or more adequately using video recordings of the diffraction images obtained in stroboscopic conditions). They seem to correspond approximately to the measured Q factor. For the different kinds of fundamental mode resonators this quantity is typically of the order of a very few thousands for X-cut lithium tantalate, of several ten thousands for medium coupling materials such as LGS and GaPO4, and up to a very few hundred thousands for large fundamental mode quartz resonators. Most often, the progressive modes have at least two components (u1 and u3) and present a quite large standing wave ratio which facilitates their detection. *It is extremely probable, that in fact, the progressive modes are the real responsible of the activity dips (by a modification of the progressive wave due to a thermal effect) and of several other annoying phenomena existing in resonators.* These modes are very complex and presently no satisfactory theory (or numerical solution) seems to exist. An example of the progressive modes observed in an X-cut TaLiO3 resonator by time resolved topography is presented in figure 16. It should be noticed that in these images a non-linear relation exist between the diffracted intensity and the strains (giving rise to the phenomena of partial contrast inversion).

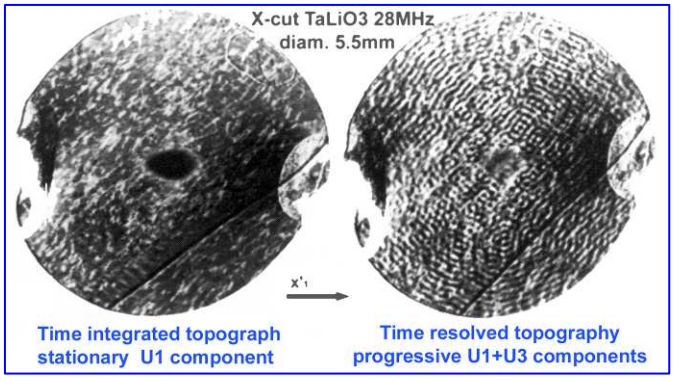


Figure 16. Progressive mode coupled to the fast shear in X-cut TaLiO3.

V. CONCLUSION

X-Ray topography constitutes an accurate method to determine the mode shapes of piezoelectric devices. The theory of the essentially thickness mode established by H.F. Tiersten predicts most often very accurately the mode shape of the resonators that have sufficiently large ratios of the lateral dimensions to the thickness. Several effects resulting of the finite dimensions of the plate are taken into account by this theory, but other ones are not relevant of its hypothesis. This is the case of the coupling of the thickness mode with resonant or progressive plate modes of significant amplitudes which are difficult to avoid in miniaturized devices. The progressive plate modes are clearly responsible of the limited Q.F product, at least, of the fundamental mode of plane thickness mode resonators. They are much probably implied in many other annoying phenomena such as the activity dips.

ACKNOWLEDGMENT

The authors acknowledge the allocation of beam time by ESRF. They are also indebted to A. Jeanne- Michaud for his essential contribution to this study.

REFERENCES

[1] W.J. Spencer in Physical Acoustics vol..5. W.P.Masson, R.N Thurston Editors. Academic Press NY p.111-161 (1969).

[2] A.R. Lang. Acta Cryst. v.12 p.249 (1959).

[3] A. Authier. Dynamical theory of X-ray diffraction. Oxford University Press Oxford. 2002.

[4] Y. Epelboin, J. Detaint, B. Capelle, J. Appl. Cryst. v. 31, p. 574-582 (1998).

[5] H.F. Tiersten J. Acoust. Soc. Am. v..57, p.667-681 (1975).

[6] D.S. Stevens, H.F. Tiersten., J. Acoust. Soc. Am. v. 79, p.1811-1826. (1986).

[7] B. Capelle. J. Detaint and Y. Epelboin. J. Appl. Cryst. v. .34, p. 625-629 (2001).

[8] J. Detaint, A. Zarka, B. Capelle, Y. Epelboin Proc. 1998 IEEE Int. Frequency Control Symposium p.882-890 (1998)

[9] .J. Detaint, A. Zarka, B. Capelle, D. Palmier, E. Philippot. Proc. 1997 IEEE Int. Frequency Control Symposium p. 566-578 (1997).

[10] D. Malocha, M.P. Da Cunha, E. Adler, S. Frederck, M. Chou. Proc. 2000 IEEE Int Frequency Control Symposium p.200-205 (2000)

[11] R. C. Smythe and G. E. Hague. Proc. 2000 IEEE IEIA & Int. Frequency Control Symposium p.191 - 194. (2000).

[12] J. A. Kosinski, R. A. Pastore, E. Bigler, M.P. da Cunha, D. C. Malocha, J. Detaint. Proc. 2001 IEEE Int. Frequency Control Symposium. p 278-286 (2001).

[13] H. Thanner, W. Wallnofer, C. Reiter, P. W. Krempl and P. M. Worsch, Annales de Chimie, Sc. Materiaux, v.26, no.1, pp. 91-94 (2001)

[14] P. Armand, M. Beaurain, B. Rufflé, B. Menaert and P. Papet. Inorg. Chem. v.48, p. 4988–4996 (2009).

[15] R.T. Smith, F.S. Welsh. J. Aplied Phys. V.42 n°6 p.2219-2230 (1971)

[16] G. Kovacs, M. Anhorn, H.E. Engan, G. Visintini, C.G.W. Ruppel. Proc. 1990 IEEE Ultrasonics Symposium p.435-438 (1990).

[17] R.M. Taziev, A.S. Kozlov, I.L. Vasiliev, I.B. Yakovkin Proc.1997 IEEE Int. Frequency Control Symposium. p.867-872 (1997)

Compact Parametric Model of Capacitive BAW Resonators

Giorgio Casinovi, Ashwin K. Samarao and Farrokh Ayazi
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

***Abstract*—This paper presents a new equivalent-circuit model of SiBARs, derived in a mathematically rigorous way from the physics equations governing the behavior of the device. The model is parametric, that is, the model component values can be computed directly from the dimensions of the resonator and the properties of the material. The model also accounts fully, accurately and automatically for aspects of the device behavior that arise from the interaction of multiple physics domain: the shift in the resonance frequency with the polarization voltage ("spring softening"), the effect of the polarization voltage and gap size on the insertion loss, and the reduction in the resonator loaded *Q* as the polarization voltage is increased. In contrast, those effects are not automatically accounted for by other models, such as mass-spring combinations. Finally, the model described is load- and source-independent, and thus it reproduces accurately the behavior of the device itself, regardless of the type and characteristics of the surrounding circuitry.**

I. INTRODUCTION

Much research activity in recent years has been directed at the development of bulk acoustic resonators that are compatible with standard integrated circuit technologies. In this respect, capacitive silicon resonators offer a particularly attractive option, since they can be made entirely of materials that are used routinely in IC fabrication processes, resulting in significant advantages in terms of ease of integration and cost savings. For this reason, their use is becoming increasingly common in a wide variety of RF circuits, such as reference oscillators and filters [1].

Disk resonators were among the first examples of micromechanical bulk resonators, but more recently width-extensional-mode resonators based on an alternative, rectangular-bar geometry were demonstrated [2]. In this paper this particular type of capacitive resonators will be referred to as silicon bulk acoustic resonators, or SiBARs. The basic structure of a SiBAR is schematically shown in Fig. 1: the resonating bar element is placed between two electrodes, supported by two thin tethers. A DC polarization voltage applied between the resonator and the electrodes generates an electrostatic field in the capacitive gaps. When an AC voltage is applied to the drive electrode, the electrostatic force applied to the corresponding face of the resonator creates an elastic

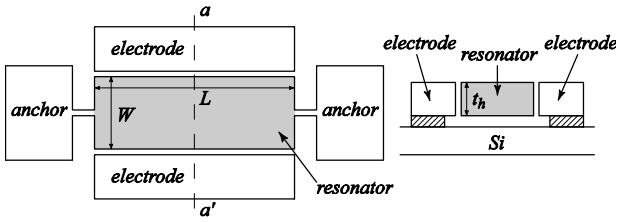


Figure 1. Schematic structure of a SiBAR: top view (left) and cross-sectional view (right)

wave that propagates through the bar. Small changes in the size of the capacitive gap on the other side of the device induce a voltage on the sense electrode, the amplitude of which peaks near the mechanical resonant frequencies of the bar.

While the behavior of bulk resonators is well understood in broad, qualitative terms, sufficiently accurate analytical models of those devices that can be used in place of full-blown finite-element simulation are still lacking. This is a significant drawback in the design of mixed-technology systems that rely on micromechanical resonators. In the simulation of such systems it is common practice to use simple RLC models for the resonators [3][4]. However these models are usually derived using empirical or semi-empirical techniques, often relying on qualitative analogies with similarly behaving systems, such as mass-spring combinations. Consequently, the accuracy of such models cannot be demonstrated or even estimated theoretically: at best, it can only be verified numerically or empirically, and on a case by case basis. In fact, an empirical *RLC* model could not track changes in *Q* and the insertion loss of a silicon bulk acoustic resonator (SiBAR) with the polarization voltage without the introduction of an artificial parasitic resistance [5].

This paper presents an equivalent circuit model of a SiBAR that is derived in a mathematically rigorous way from the governing equations of the device. The model is parametric, in that the model component values can be computed directly from the dimensions of the resonator and the properties of the material. The only exceptions are the resistances in the model, whose values depend on the energy

This work was supported in part by a NSF EAGER grant.

losses in the resonator, which are difficult to model analytically.

The model also accounts fully, accurately and automatically for aspects of the device behavior that arise from the interaction of multiple physics domain: the shift in the resonance frequency with the polarization voltage ("spring softening"), the effect of the polarization voltage and gap size on the insertion loss, and the reduction in the resonator loaded Q as the polarization voltage is increased. In contrast, those effects are not automatically accounted for by other models, such as mass-spring combinations. Finally, the model described in this paper is load- and source-independent, and thus it reproduces accurately the behavior of the device itself, regardless of the type and characteristics of the surrounding circuitry. The model has been validated numerically against finite-element simulations and experimentally against measurements taken on a SiBAR fabricated using the HARPSS process [2].

II. Resonator Model

As in all capacitive MEMS resonators, the operation of a SiBAR is based on the principle of electromechanical transduction. Let W, L and t_h denote respectively the width, length and thickness of the resonator, as shown in Fig. 1. If a DC polarization voltage V_p is applied across the drive and sense capacitive gaps, the resulting electric fields create pressure on the faces of the resonator and a consequent elastic deformation of the structure. If a small-signal voltage v_d is superimposed to the DC voltage at the drive electrode, the change in the electric field in the capacitive gap and the corresponding change in the pressure applied to the resonator face generate an elastic wave that propagates across the width of the resonator.

Therefore a compact SiBAR model can be derived from the equations of elastic wave propagation, under the approximate assumption that the resonant modes of interest are created by the propagation of plane waves in the direction of the width of the resonator. Under this assumption, an analytical solution of the elasticity equations can be obtained. By means of appropriate circuit synthesis methods, this solution is used to derive an equivalent circuit, that is, a circuit the behavior of which is described by the same equations that characterize the operation of the SiBAR.

The resulting equivalent circuit model of the SiBAR is shown in Fig. 2. The values of the circuit elements are given by the following expressions

$$\begin{aligned} C_1 &= C_g/8 & L_1 &= 8/(\omega_0^2 C_g) \\ C_2 &= 2C_g/\pi^2 & L_2 &= \pi^2/(2\omega_0^2 C_g) \\ C_0 &= \varepsilon^2 A^3 c_{33}/WQ_p^2 & R_1 R_2 &= (2\pi/\omega_0 C_g)^2 \end{aligned} \tag{1}$$

In the expressions above, Q_p is the value of the DC charge that is present on the drive and sense gap capacitances. Specifically $Q_p = C_g V_p = (\varepsilon A/g)V_p$, where g is the size of the capacitive gap, ε the dielectric permittivity, $A = Lt_h$ and C_g the value of the gap capacitance. Furthermore $\omega_0 = 2\pi f_0$, where f_0 is the mechanical resonance frequency of the resonator, and c_{33} is the stiffness coefficient of the resonator material in the direction of propagation of the elastic waves.

The values of R_1 and R_2 depend on the energy losses in the resonator, that is, on its mechanical Q. Note that (1) imposes a constraint on the product $R_1 R_2$. Consequently the value of only one of those resistors can be chosen independently. For example, the measured ratio between the input voltage and the short-circuit output current of the SiBAR at resonance can be used for this purpose. An approximate but sufficiently accurate expression for this ratio can be obtained by observing that near the resonant frequency the current through the series $R_2 L_2 C_2$ circuit will dominate the current through the parallel $R_1 L_1 C_1$ circuit: it is therefore reasonable to assume that $i_3 \cong i_4$. After some algebraic manipulations, the following expression is obtained

$$i_s = \frac{j\omega C_g/L_2 C_0}{-\omega^2 + j\omega(R_2/L_2) + (1 - 2C_2/C_0)/(L_2 C_2)} v_d \tag{2}$$

The expression above shows that the electrical resonant frequency ω_1 — i.e. the frequency at which i_s achieves its maximum value — is lower than the mechanical resonant frequency ω_0 and is given by

$$\omega_1 = \omega_0 \sqrt{1 - 2C_2/C_0} \tag{3}$$

In particular, since the value of C_0 depends on the DC bias voltage V_p, so does the value of ω_1. This is a well-known effect, which has been consistently observed in experimental setups [3][5]. Setting $\omega = \omega_1$ in (2) yields the desired

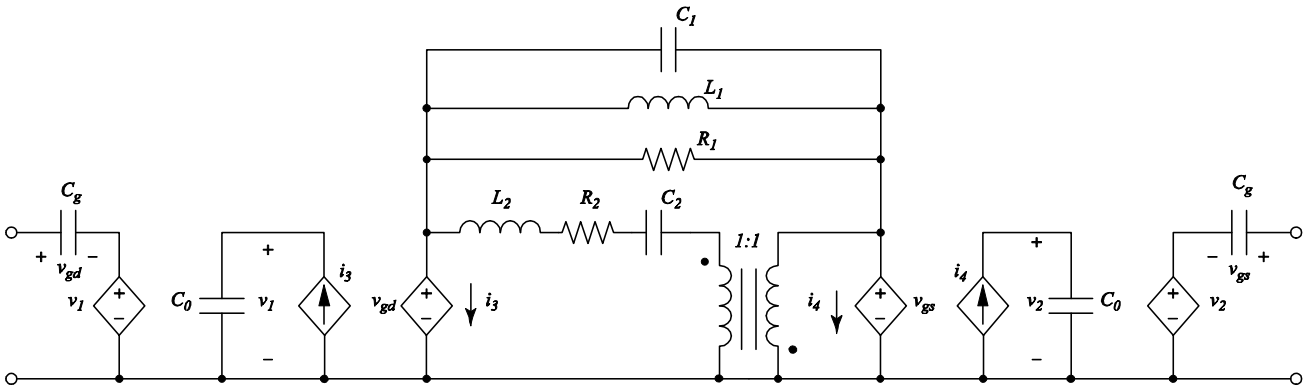


Figure 2. SibAR equivalent circuit model

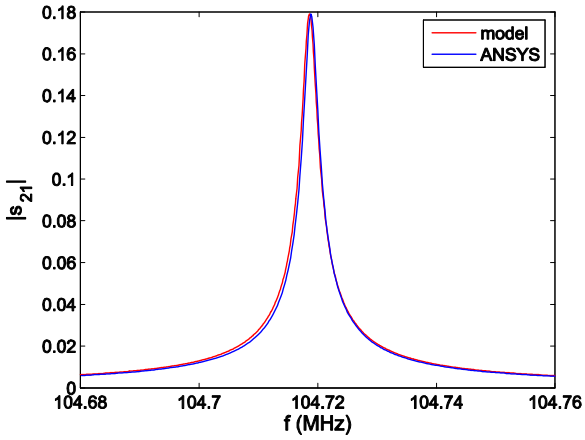


Figure 3. SiBAR s_{21} parameter extracted from simulations

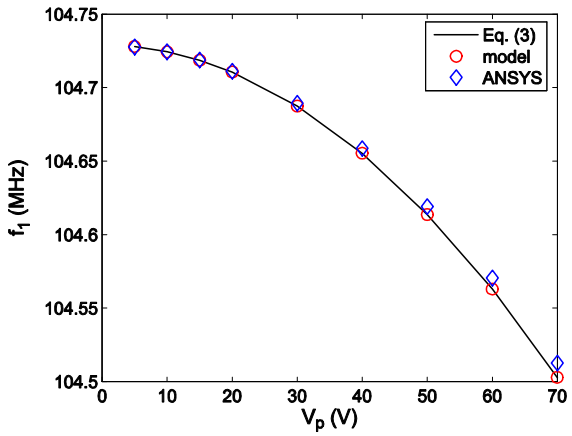


Figure 4. SiBAR electrical resonance frequency vs. polarization voltage

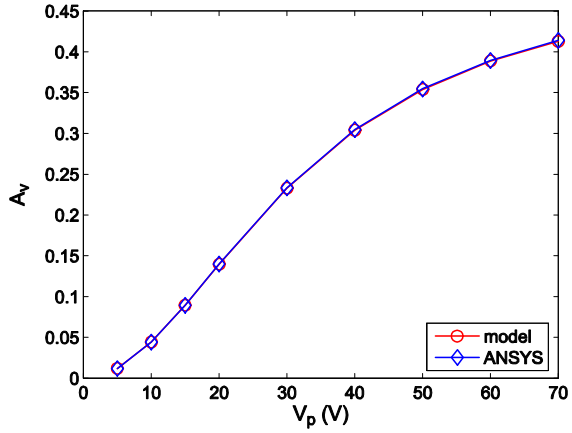


Figure 5. SiBAR voltage gain at resonance vs. polarization voltage

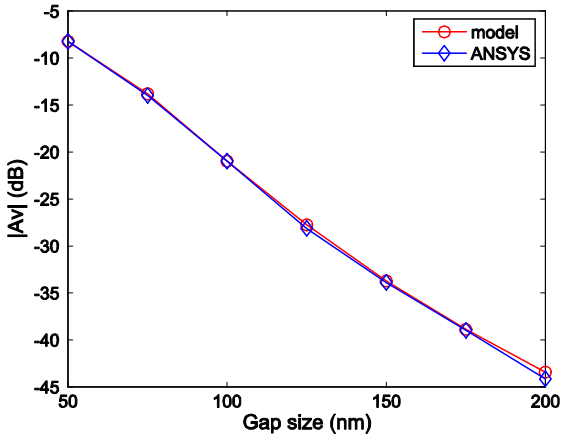


Figure 6. SiBAR voltage gain at resonance vs. capacitive gap size

relationship between the value of R_2 and the v_d/i_s ratio at resonance, namely

$$R_2 = \frac{C_g}{C_0}\left(\frac{v_d}{i_s}\right)_{\omega=\omega_1} \tag{4}$$

Note that once the value of R_2 has been computed for a particular geometry, it can be extrapolated to other geometries if the rate of energy loss in the resonator does not change.

Further improvements to the model can be made by solving the elasticity equations more accurately, taking the finite thickness of the resonator into account [6]. Because of space limitations the calculations are omitted here, but the end result is a modified expression for C_0, while the rest of the equivalent circuit model remains unchanged.

III. Numerical Simulations

The equivalent circuit model described in the previous section was compared to finite-element simulations of a SiBAR performed in ANSYS. A detailed description of the ANSYS model is given in [6]. The device used in the comparisons was a resonator of dimensions L = 400 µm, W = 40 µm, t_h = 20 µm, with a DC bias voltage V_p = 15 V. Fig. 3 compares the s_{21} parameter obtained from an ANSYS simulation of the device with 50 Ω source and load resistances with the s_{21} parameter of the model under the same conditions.

A second set of simulations was then performed on the same device with various DC polarization voltages, ranging from 5 V to 70 V, while keeping the values of the source and load resistances fixed at 50 Ω. Fig. 4 shows the electrical resonant frequency $f_1 = \omega_1/2\pi$ of the device at each value of V_p, obtained both from ANSYS and model simulations. The same figure shows also the graph of the relationship between f_1 and V_p given by (3). It can be seen that the simulation results track very closely the relationship in (3), which in turn matches actual measurements taken on SiBARs, as will be shown in Sec. IV. Fig. 5 shows the value of the voltage gain $A_v = v_{out}/v_{in}$ at resonance as a function of V_p, obtained both from ANSYS and model simulations.

A final set of simulations was performed to evaluate the effect of the size of the capacitive gaps on the voltage gain of the device. It should be noted that for these simulations the values of R_1 and R_2 were changed assuming an inversely proportional relationship to C_g. This assumption stems from the fact that R_1 and R_2 are determined by the mechanical Q of the resonator, which is independent of the gap size. Based on the expressions in (1), keeping the Q of resonant circuits

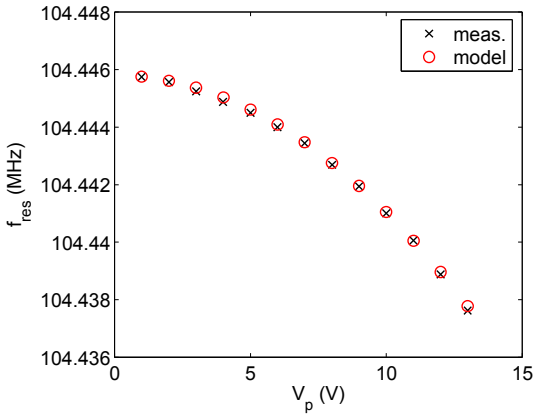


Figure 7. SiBAR resonance frequency predicted by the model compared to measured values

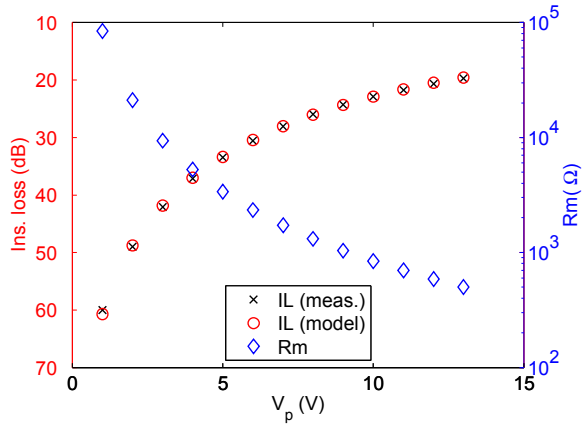


Figure 8. Left axis: SiBAR insertion loss predicted by the model compared to measured values. Right axis: SiBAR motional resistance

$R_1L_1C_1$ and $R_2L_2C_2$ constant with the gap size requires the resistance values to be inversely proportional to C_g. The results of this last set of simulations obtained from the model and from ANSYS are shown in Fig. 6. It can be seen that in all cases the graphs in the figures show excellent agreement between the two sets of data, a result that clearly supports the validity of the model.

IV. Experimental Model Validation

In addition to finite-element simulations, the SiBAR model was validated experimentally against measurements taken on a device fabricated using the HARPSS process [2]. The nominal dimensions of the device were L = 400 µm, W = 40 µm, t_h = 20 µm. The extracted values of the width and gap size were W = 40.028 µm and g = 99 nm, respectively. Fig. 7 compares the measured resonance frequency of the device with that predicted by the model, as a function of the polarization voltage V_p. Fig. 8 shows a similar comparison for the insertion loss of the device, also as a function of V_p. For reference purposes, the corresponding values of the motional resistance R_m are also plotted. Finally, the values of the quality factor Q of the device, both measured and predicted by the model, are plotted versus V_p in Fig. 9. Although the agreement is not as good as in the two previous figures, it can be seen that the

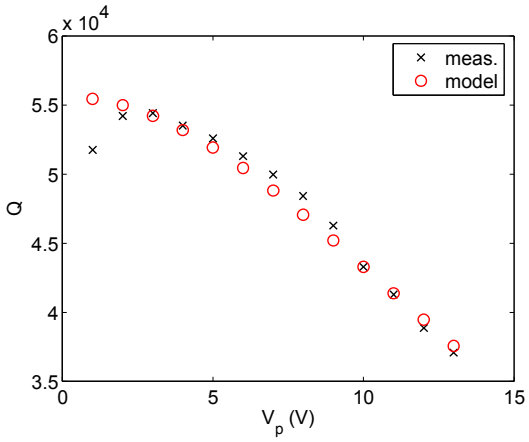


Figure 9. SiBAR quality factor Q predicted by the model compared to measured values

model tracks with good accuracy the drop in Q with the increase in V_p without introducing an artificial additional resistance, which is instead needed in simpler, empirically derived resonator models [5].

V. Conclusion

This paper has introduced a parametric equivalent circuit model for SiBARs that is derived directly from the equations governing the electromechanical behavior of the device. Consequently the values of most of the components in the model can be computed analytically, based on the dimensions of the resonator and the properties of the material. The only model parameters that cannot be determined in this way are the resistance values, which are tied to energy losses in the device. Simulation results show that the performance parameters extracted from the model are in excellent agreement with those obtained both from finite-element analysis of the device and from actual measurements. Thus the model can be profitably used in the design of complex mixed-technology circuits that include SiBARs among their components, such as VHF oscillators.

References

[1] C. T.-C. Nguyen, "MEMS technology for timing and frequency control," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 54, no. 2, pp. 251–270, Feb. 2007.

[2] S. Pourkamali, G. K. Ho, and F. Ayazi, "Low-impedance VHF and UHF capacitive silicon bulk acoustic wave resonators — Part I: Concept and fabrication," *IEEE Trans. Electron Devices*, vol. 54, no. 8, pp. 2017–2023, Aug. 2007.

[3] J. Wang, Z. Ren, and C. T.-C. Nguyen, "1.156-GHz self-aligned vibrating micromechanical disk resonator," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 51, no. 12, pp. 1607–1628, Dec. 2004.

[4] J. R. Clark, W.-T. Hsu, M. A. Abdelmoneum, and C. T.-C. Nguyen, "High-Q UHF micromechanical radial-contour mode disk resonators," *J. Microelectromech. Syst.*, vol. 14, no. 6, pp. 1298–1310, Dec. 2005.

[5] S. Pourkamali, G. K. Ho, and F. Ayazi, "Low-impedance VHF and UHF capacitive silicon bulk acoustic wave resonators — Part II: Measurement and characterization," *IEEE Trans. Electron Devices*, vol. 54, no. 8, pp. 2024–2030, Aug. 2007.

[6] G. Casinovi, X. Gao, and F. Ayazi, "Lamb waves and resonant modes in rectangular-bar silicon resonators," *J. Microelectromech. Syst.*, vol. 19, no. 4, pp. 827–839, Aug. 2010.

Analytical Solutions of Nonlinear Vibrations of Thickness-shear and Flexural Modes of Quartz Plates

Ji Wang[1], *Rongxing Wu*[1], *Jianke Du*[1], *Wei Yan*[2] *and Dejin Huang*[1]
[1]Piezoelectric Device Lab, School of Mech Eng & Mechanics, Ningbo Univ, 818 Fenghua Rd, Ningbo, Zhejiang, 315211, China
[2]School of Arch, Civil Eng & Env, Ningbo Univ, 818 Fenghua Rd, Ningbo, Zhejiang, 315211, China
E-mail: wangji@nbu.edu.cn

***Abstract*—As the systematic study of nonlinear vibrations of quartz crystal resonators, we have established the nonlinear Mindlin plate equations which include the strongly coupled thickness-shear and flexural modes. Based on the simplification of these complicated nonlinear equations, equations of thickness-shear vibrations of quartz crystal plates have been solved by the combination of the Galerkin method and homotopy analysis method (HAM). The amplitude-frequency relation we have obtained shows that the nonlinear frequency of thickness-shear vibrations depends on the vibration amplitude and thickness and length of plate, which is quite different from the linear cases. All these efforts will result in applicable solution techniques for further studies of nonlinear effects of quartz plates under bias fields for the precise analysis and design of quartz crystal resonators.**

I. INTRODUCTION

Common types of piezoelectric resonators such as the quartz crystal resonators always operate at a desired resonant frequency with broad applications for frequency control and detection in electronic and mechanical products and systems [1-2]. The linear piezoelectric plate theories for these applications have been successfully studied by Mindlin, Tiersten, and Lee et al and yielded many satisfactory results [3-5]. However, with demand for higher frequency and miniaturization of piezoelectric structures, many nonlinear phenomena such as the derive level dependency (DLD) and activity dip have emerged and required analysis [6]. These intrinsic nonlinearities are responsible for the coupling between the vibration modes and external or internal bias field effects which ultimately cause frequency instability and affect the operation precision of piezoelectric devices. Therefore, there is a strong need to study nonlinear properties and behaviors of vibrations of quartz crystal resonators for many different purposes. Actually, some preliminary and challenging work has been done by many researchers. Yang has established equations for the extensional and flexural vibrations of electroelastic plates under strong electric field [7]. Yang and his collaborators also studied many nonlinear couplings between thickness-shear and other vibration modes [8-9]. In his doctoral dissertation, Patel has employed the finite element method based on equations of nonlinear three-dimensional piezoelectricity to analyze well-known DLD phenomenon in quartz crystal resonators [10].

In our earlier studies, we have already established the nonlinear Mindlin plate equations which include the thickness-shear and flexural vibration modes that are strongly coupled [11-12]. We found the nonlinear Mindlin plate equations are indeed complicated, and cannot be directly solved by known methods we are familiar with from linear equations. On the other hand, even with powerful computers, both linear and nonlinear Mindlin plate equations are still difficult to solve with numerical methods due to difficulties in solving extremely large scale linear systems at high frequency. The finite difference method also has been tried without any valuable results [13]. For this reason, we focused on solving the nonlinear equations of flexural and thickness-shear vibrations of quartz crystal plate by the Galerkin method and homotopy analysis method (HAM). Both the analytical and numerical results have shown some inherent characteristics of nonlinear vibration and given some valuable directions for further research.

II. THE NONLINEAR MINDLIN PLATE EQUATIONS

Through the consideration of kinematic and material nonlinearities, nonlinear Mindlin plate equations have been established by following the procedure of Mindlin with the signature expansion of displacements in the thickness coordinate [11-12]. The nonlinear equations of flexural and thickness-shear vibrations are [12]

$$
\begin{aligned}
&2bc_{1212}\left(u_{2,11}^{(0)}+u_{1,1}^{(1)}\right)+2bc_{121112}u_{2,11}^{(0)}u_{2,1}^{(0)}\left(u_{2,1}^{(0)}+u_{1}^{(1)}\right)\\
&+bc_{121112}u_{2,1}^{(0)}u_{2,1}^{(0)}\left(u_{2,11}^{(0)}+u_{1,1}^{(1)}\right)+2bc_{1111}u_{2,11}^{(0)}u_{2,1}^{(0)}u_{2,1}^{(0)}\\
&+2bc_{122212}u_{1,1}^{(1)}u_{1}^{(1)}\left(u_{2,1}^{(0)}+u_{1}^{(1)}\right)+bc_{122212}u_{1}^{(1)}u_{1}^{(1)}\left(u_{2,11}^{(0)}+u_{1,1}^{(1)}\right)\\
&+2bc_{1122}u_{1,1}^{(1)}u_{1}^{(1)}u_{2,1}^{(0)}+bc_{111111}\left(u_{2,1}^{(0)}\right)^{3}u_{2,11}^{(0)}u_{2,1}^{(0)}\\
&+bc_{111122}u_{2,1}^{(0)}u_{2,11}^{(0)}u_{1}^{(1)}u_{1}^{(1)}u_{2,1}^{(0)}+bc_{111122}u_{2,1}^{(0)}u_{2,1}^{(0)}u_{1,1}^{(1)}u_{1}^{(1)}u_{2,1}^{(0)}\\
&+bc_{112222}\left(u_{1}^{(1)}\right)^{3}u_{1,1}^{(1)}u_{2,1}^{(0)}+2bc_{111212}\left(u_{1}^{(1)}u_{1,1}^{(1)}\right)\left(u_{1,1}^{(1)}u_{1,1}^{(1)}+u_{1}^{(1)}u_{1,11}^{(1)}\right)u_{2,1}^{(0)}\\
&+bc_{1111}u_{2,1}^{(0)}u_{2,1}^{(0)}u_{2,11}^{(0)}+bc_{1122}u_{1}^{(1)}u_{1}^{(1)}u_{2,11}^{(0)}+\frac{1}{4}bc_{111111}\left(u_{2,1}^{(0)}u_{2,1}^{(0)}\right)^{2}u_{2,11}^{(0)}
\end{aligned}
$$

This research is supported by grants from the National Natural Science Foundation of China (Nos 10932004 & 10772087) and the Doctoral Program Fund of Ministry of Education of China (No 20093305110003).

$$
\begin{aligned}
&+\frac{b}{2}c_{111122}u_{2,1}^{(0)}u_{2,1}^{(0)}u_1^{(1)}u_1^{(1)}u_{2,11}^{(0)}+\frac{1}{4}bc_{112222}\left(u_1^{(1)}u_1^{(1)}\right)^2u_{2,11}^{(0)}\\
&+bc_{111212}\left(u_1^{(1)}u_{1,1}^{(1)}\right)^2u_{2,11}^{(0)}=2b\rho\ddot{u}_2^{(0)},\\
&\frac{2b^3}{3}c_{1111}u_{1,11}^{(1)}+\frac{2b^3}{3}c_{111111}u_{1,1}^{(1)}u_{2,11}^{(0)}u_{2,1}^{(0)}+\frac{b^3}{3}c_{111111}u_{1,11}^{(1)}u_{2,1}^{(0)}u_{2,1}^{(0)}\\
&+\frac{2b^3}{3}c_{111122}u_{1,1}^{(1)}u_1^{(1)}u_{1,1}^{(1)}+\frac{b^3}{3}c_{111122}u_{1,11}^{(1)}u_1^{(1)}u_1^{(1)}\\
&+\frac{2b^3}{3}c_{121112}u_1^{(1)}u_{1,11}^{(1)}\left(u_{2,1}^{(0)}+u_1^{(1)}\right)+\frac{2b^3}{3}c_{121112}u_1^{(1)}u_{1,1}^{(1)}\left(u_{2,11}^{(0)}+u_{1,1}^{(1)}\right)\\
&+\frac{2b^3}{3}c_{111212}\left[u_{1,1}^{(1)}u_{1,1}^{(1)}\left(u_{2,1}^{(0)}+u_1^{(1)}\right)+u_1^{(1)}u_{1,11}^{(1)}\left(u_{2,1}^{(0)}+u_1^{(1)}\right)\right]\\
&+\frac{2b^3}{3}c_{111212}u_1^{(1)}u_{1,1}^{(1)}\left(u_{2,11}^{(0)}+u_{1,1}^{(1)}\right)+\frac{b^3}{3}c_{121112}u_1^{(1)}u_{2,1}^{(0)}u_{2,1}^{(0)}u_1^{(1)}u_{1,11}^{(1)}\\
&+\frac{b^3}{3}c_{121112}u_1^{(1)}\left[2u_{2,11}^{(0)}u_{2,1}^{(0)}u_1^{(1)}u_{1,1}^{(1)}+u_{2,1}^{(0)}u_{2,1}^{(0)}u_{1,1}^{(1)}u_{1,1}^{(1)}\right]\\
&+b^3c_{122212}u_1^{(1)}u_1^{(1)}u_{1,1}^{(1)}u_{1,1}^{(1)}u_1^{(1)}+\frac{b^3}{3}c_{122212}u_1^{(1)}u_1^{(1)}u_1^{(1)}u_{1,11}^{(1)}u_1^{(1)}\\
&-2bc_{1212}\left(u_{2,1}^{(0)}+u_1^{(1)}\right)-bc_{121112}u_{2,1}^{(0)}u_{2,1}^{(0)}\left(u_{2,1}^{(0)}+u_1^{(1)}\right)\\
&-bc_{122212}u_1^{(1)}u_1^{(1)}\left(u_{2,1}^{(0)}+u_1^{(1)}\right)-bc_{2211}u_{2,1}^{(0)}u_{2,1}^{(0)}u_1^{(1)}\\
&-bc_{2222}u_1^{(1)}u_1^{(1)}u_1^{(1)}-\frac{b}{4}c_{221111}\left(u_{2,1}^{(0)}u_{2,1}^{(0)}\right)^2u_1^{(1)}\\
&-\frac{b}{2}c_{221122}u_{2,1}^{(0)}u_{2,1}^{(0)}u_1^{(1)}u_1^{(1)}u_1^{(1)}-\frac{b}{4}c_{222222}\left(u_1^{(1)}u_1^{(1)}\right)u_1^{(1)}\\
&-bc_{221212}u_1^{(1)}\left(u_1^{(1)}u_{1,1}^{(1)}\right)^2=\frac{2b^3}{3}\rho\ddot{u}_1^{(1)}\ ,
\end{aligned}
\tag{1}
$$

where $u_2^{(0)}$ and $u_1^{(1)}$ are the displacements of flexural and thickness-shear vibrations modes, c_{ijkl} and c_{ijklmn} are the second and third elastic constants of material, b and ρ are the half thickness of plate and density of material, respectively. Not surprisingly, we found that the nonlinear Mindlin plate equations with the consideration of kinematic and material nonlinearities are indeed complicated. Therefore, we first carry out the analysis of linear Mindlin plate equations for comparison purpose. When the nonlinear terms in (1) are neglected, we have

$$
\begin{aligned}
&2bc_{66}\left[u_{2,11}^{(0)}+u_{1,1}^{(1)}\right]=2b\rho\ddot{u}_2^{(0)},\\
&\frac{2b^3}{3}c_{11}u_{1,11}^{(1)}-2bc_{66}\left[u_{2,1}^{(0)}+u_1^{(1)}\right]=\frac{2b^3}{3}\rho\ddot{u}_1^{(1)}.
\end{aligned}
\tag{2}
$$

By assuming a long thickness-shear wave mode, we have corresponding displacements as [3-4]

$$
u_2^{(0)}=A\sin(\xi x_1)\cos(\omega t),\quad u_1^{(1)}=B\cos(\eta x_1)\cos(\omega t),
\tag{3}
$$

where $A, B, \xi, \eta,$ and ω are amplitudes, wavenumbers, and angular frequency, respectively.

For the case when the two ends of the plate are free, the corresponding boundary conditions we have are [3, 14]

$$
u_{2,1}^{(0)}(x_1=\pm a)=0,\quad u_1^{(1)}(x_1=\pm a)=0.
\tag{4}
$$

Substituting (3) into (4), we have

$$
\xi=\frac{n\pi}{2a}, n=1,3,5,7\cdots,\quad \eta=\frac{m\pi}{2a}, m=1,3,5,7\cdots.
\tag{5}
$$

By inserting (5) back into (3) and (2), and consider thickness-shear approximation [3-4], we have

$$
\omega^2\cong\frac{3c_{66}}{b^2\rho}+\frac{c_{11}}{\rho}\left(\frac{m\pi}{2a}\right)^2.
\tag{6}
$$

This is the known long thickness-shear vibration frequency of a finite quartz crystal plate [1]. Particularly, if $m=1$, it is the fundamental long thickness-shear frequency.

III. Galerkin Approximation

We are mostly interested in the effect of the flexural mode and nonlinear terms on thickness-shear vibration frequency and mode shapes. The Galerkin method is frequently employed to solve nonlinear wave propagation problems in finite regions [15-16]. By assuming the spatial variation of displacements in the form of linear mode shape, we can use the orthogonal properties of the linear mode shapes to obtain equation to determine temporal behavior.

Based on (3) and (5), for the lowest-order effect of the flexural mode on the fundamental thickness-shear mode $m=1, n=1$, we have

$$
u_2^{(0)}=\overline{A}u(t)\sin\left(\frac{\pi x_1}{2a}\right),\quad u_1^{(1)}=\overline{B}u(t)\cos\left(\frac{\pi x_1}{2a}\right),
\tag{7}
$$

where $\overline{A}$ and $\overline{B}$ are new nonlinear amplitudes. We move the inertia term from right side of the second equation of (1) to left side and define this expression as the weighting error. We also require this weight error to vanish in one period, or over the length, as

$$
\int_{-a}^{a}\varepsilon\cdot\cos\left(\frac{\pi x_1}{2a}\right)dx_1=0.
\tag{8}
$$

Substituting (7) into (8) and integrating through the length of the plate, we have the equation

$$
\ddot{u}+\widetilde{C}_1u+\widetilde{C}_2u^3+\widetilde{C}_3u^5=0,
\tag{9}
$$

where

$$
\begin{aligned}
\widetilde{C}_1&=\omega_F^2\cong\frac{3c_{66}}{b^2\rho}+\frac{c_{11}\pi^2}{4a^2\rho},\\
\widetilde{C}_2&=\frac{\overline{B}^2c_{22}}{b^2\rho}+\frac{\overline{B}^2c_{222}}{4b^2\rho}+\frac{\overline{B}^2c_{266}}{b^2\rho}+\frac{\overline{A}\overline{B}c_{266}\pi}{2ab^2\rho}+\frac{\overline{B}^2c_{166}\pi^2}{12a^2\rho}\\
&\quad+\frac{\overline{A}^2c_{12}\pi^2}{4a^2b^2\rho}+\frac{\overline{A}^2c_{166}\pi^2}{4a^2b^2\rho}+\frac{\overline{A}\overline{B}c_{166}\pi^3}{24a^3\rho}+\frac{\overline{A}^3c_{166}\pi^3}{8\overline{B}a^3b^2\rho},\\
\widetilde{C}_3&=\frac{\overline{A}^2\overline{B}^2c_{122}\pi^2}{10a^2b^2\rho}+\frac{\overline{B}^4c_{266}\pi^2}{20a^2b^2\rho}+\frac{\overline{B}^4c_{266}\pi^2}{60a^2\rho}+\frac{\overline{A}^2\overline{B}^2c_{166}\pi^4}{240a^4\rho}+\frac{\overline{A}^4c_{112}\pi^2}{80a^4b^2\rho},
\end{aligned}
\tag{10}
$$

where ω_F is fundamental and referential frequency of long thickness-shear vibration of quartz crystal plate.

IV. Nonlinear Solutions

The standard homotopy analysis method proposed by Liao has been successfully applied to solve many types of nonlinear equations in science and engineering [17-19]. We want to

employ this method to solve nonlinear equation of thickness-shear vibrations of quartz crystal plate as shown in (9). First, we introduce two new variables, $\tau = \omega t$ and $y = u$, to (9) and assume corresponding initial condition as

$$\begin{aligned} &\omega^2 y'' + \widetilde{C}_1 y + \widetilde{C}_2 y^3 + \widetilde{C}_3 y^5 = 0, \\ &y(0) = \widetilde{B}, y'(0) = 0, \end{aligned} \tag{11}$$

where ω is the nonlinear frequency which will be determined later, and $\widetilde{B}$ is the initial condition. It should also be pointed out that primes denote derivatives with respect to the new variable τ in (11).

Now we want to solve (11) by homotopy analysis method. Based on the boundary conditions (11), it is reasonable to assume the solution expression as [17]

$$y = \sum_{m=0}^{+\infty} c_m \cos[(2m+1)\tau], \tag{12}$$

where c_m is a constant coefficient. Under the rule of solution expression in (12), we choose

$$y_0 = \widetilde{B}\cos(\tau), \tag{13}$$

as the initial guess of $y(\tau)$. First of all, we define two Taylor series as

$$\begin{aligned} &\phi(\tau,q) = y_0(\tau) + \sum_{m=1}^{+\infty} y_m q^m, \ \omega(q) = \omega_0 + \sum_{m=1}^{+\infty} \omega_m q^m, \\ &y_m(\tau) = \frac{1}{m!}\frac{\partial^m \phi(\tau,q)}{\partial q^m}\bigg|_{q=0}, \ \omega_m(q) = \frac{1}{m!}\frac{\partial^m \omega(q)}{\partial q^m}\bigg|_{q=0}, \end{aligned} \tag{14}$$

where $q \in [0,1]$ denotes an embedding parameter. Furthermore, we define auxiliary linear and nonlinear operators as

$$\begin{aligned} &L[\phi(\tau,q)] = \omega_0^2\left[\frac{\partial^2 \phi(\tau,q)}{\partial \tau^2} + \phi(\tau,q)\right], \\ &N[\phi(\tau,q),\omega(q)] = \omega^2 \frac{\partial^2 \phi(\tau,q)}{\partial \tau^2} + \widetilde{C}_1 \phi(\tau,q) + \widetilde{C}_2 \phi^3(\tau,q) + \widetilde{C}_3 \phi^5(\tau,q), \end{aligned} \tag{15}$$

where ω_0 is the initial guess of nonlinear frequency.

We construct the zeroth-order deformation equation

$$(1-q)L[\phi(\tau,q) - y_0(\tau)] = qhH(\tau)N[\phi(\tau,q),\omega(q)], \tag{16}$$

where $h \neq 0$ is an auxiliary parameter, and $H(\tau) \neq 0$ is an auxiliary function of τ. From (16) we know that if q increases from 0 to 1, $\phi(\tau,0)$ and ω_0 varies form initial guess to the exact solutions. If auxiliary function $H(\tau)$ and auxiliary parameter h are so properly chosen that $\phi(\tau,q)$ and $\omega(q)$ are convergent at $q = 1$.

Differentiating the zeroth-order deformation equation (16) m times with respect to q, then setting $q = 0$, and finally dividing them by $m!$, we obtain the higher-order deformation equation

$$\begin{aligned} &L[y_m(\tau) - \chi_m y_{m-1}(\tau)] = hqH(\tau)R_m(y_{m-1},\omega_{m-1}), \\ &\chi_m = \begin{cases} 0, & m \le 1, \\ 1, & m > 1. \end{cases} \end{aligned} \tag{17}$$

with the boundary conditions as

$$y_m(0) = 0, y_m'(0) = 0, \tag{18}$$

where

$$\begin{aligned} R_m(y_{m-1},\omega_{m-1}) &= \frac{1}{(m-1)!}\frac{\mathrm{d}^{m-1} N[\phi(\tau,q),\omega(q)]}{\mathrm{d}q^{m-1}}\bigg|_{q=0} \\ &= \sum_{j=0}^{m-1} y_j''(\tau) \sum_{r=0}^{m-1-j} \omega_r \omega_{m-1-j-r} + \widetilde{C}_1 y_{m-1}(\tau) \\ &\quad + \widetilde{C}_2 \sum_{j=0}^{m-1} y_j(\tau) \sum_{r=0}^{m-1-j} y_r(\tau) y_{m-1-j-r}(\tau) \\ &\quad + \widetilde{C}_3 \sum_{j=0}^{m-1} y_j(\tau) \sum_{i=0}^{m-1-j} y_i(\tau) \sum_{k=0}^{m-1-i-j} y_k(\tau) \sum_{r=0}^{m-1-i-k-j} y_r(\tau) y_{m-1-i-k-j-r}(\tau). \end{aligned} \tag{19}$$

Based on solution expression (12) and inherent properties of nonlinear equation (12), (17) can be expressed as

$$hqH(\tau)R_m(y_{m-1},\omega_{m-1}) = \sum_{n=0}^{\mu_m} b_{m,n}(\omega_{m-1})\cos[(2n+1)\tau], \tag{20}$$

where $b_{m,n}$ is a constant, μ_m is a positive integer determined by m and $H(\tau)$. In order to avoid appearance of secular terms such as $\tau\cos(\tau)$ in solution of (17), we must force coefficient of $\cos(\tau)$ to vanish as

$$b_{m,0}(\omega_{m-1}) = 0, \tag{21}$$

which provides us with another algebraic equation for ω_{m-1}.

For simplicity, we assume $H(\tau) = 1$, hence the solution of (17) can be expressed in the form

$$\begin{aligned} y_m(\tau) &= \chi_m y_{m-1}(\tau) \\ &\quad - \frac{1}{\omega_0^2}\sum_{n=0}^{\mu_m} \frac{b_{m,n}}{4n(n+1)}\cos[(2n+1)\tau] + D_1\sin(\tau) + D_2\cos(\tau), \end{aligned} \tag{22}$$

where D_1 must be forced to zero for satisfying the rule of solution expression [17], and D_2 is determined by the initial conditions. By Mathematica programs, we can obtain the higher order expression for y_m and ω_{m-1}. For example, the zeroth- to first-order solutions of nonlinear frequency are gave as

$$\begin{aligned} \omega_0 &= \frac{\sqrt{8\widetilde{C}_1 + 6\widetilde{B}^2\widetilde{C}_2 + 5\widetilde{B}^4\widetilde{C}_3}}{2\sqrt{2}}, \\ \omega_1 &= \frac{36\widetilde{B}^4 h\widetilde{C}_2^2 + 96\widetilde{B}^6 h\widetilde{C}_2\widetilde{C}_3 + 65\widetilde{B}^8 h\widetilde{C}_3^2}{96\sqrt{2}\left(8\widetilde{C}_1 + 6\widetilde{B}^2\widetilde{C}_2 + 5\widetilde{B}^4\widetilde{C}_3\right)^{3/2}}. \end{aligned} \tag{23}$$

For a practical quartz crystal resonator, we are mostly interested in nonlinear terms on frequency which cause frequency shift. Therefore, we defined frequency shift as

$$\Omega = (\omega - \omega_F)/\omega_F , \tag{24}$$

where ω_F in the fundamental frequency of long wave thickness-shear vibration of quartz crystal plate which has been defined in (10). In this study, we consider an 8.6845 MHz resonator [1].

In the previous derivation we have the assumption that h is to be properly chosen so that the solution series (14) are convergent. The admissible values of this auxiliary parameter can be found by plotting the so-called h-curves [17]. The h-curves for Ω is shown in Fig. 1. It is clear that the series for Ω is convergent with large scope of h for practical vibration amplitude. The nonlinear material properties needed for the calculation are given by Patel [10].

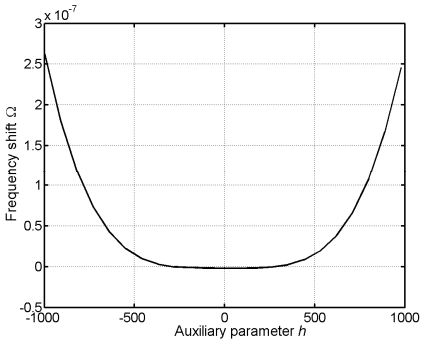


Figure 1. The h-curve for Ω with $\overline{B} = 4\times10^{-7}$.

Table 1 shows the comparison of our analytical solutions of Ω with results of perturbation method with different vibration amplitude. Numerical results show frequency shift caused by vibration amplitude is very small which indicates neither kinematic nor material nonlinearities are the main factor of frequency shifts and performance fluctuation of quartz crystal resonators.

Table 1. The frequency shift Ω with different amplitudes

Amplitude	HAM solutions	Perturbation solutions
1E-10	-1.1936620731E-16	-1.1936620731E-16
2E-9	-4.1778172561E-14	-4.1778172561E-14
3E-8	-1.1936620731E-11	-1.1936620731E-11
4E-7	-1.7904931116E-9	-1.7904931116E-9
5E-6	-2.9841557024E-7	-2.9841557024E-7
6E-5	-4.1779190805E-5	-4.1779190795E-5

V. CONCLUSIONS

The nonlinear equations of thickness-shear vibrations of quartz crystal plates have been studied by the Galerkin method and homotopy analysis method, and we obtained amplitude-frequency relation which depends on many factors such as vibration amplitude, thickness, and length of a plate. The amplitude-frequency relation also reveals that with the change of amplitude, the frequency will become unstable which is not known in the linear theory. Numerical results show neither kinematic nor material nonlinearities is the main factor of frequency shifts and performance fluctuation of quartz crystal resonators. This promotes us to study of nonlinear behaviors of quartz crystal resonators with the consideration of many complicating factors.

REFERENCES

[1] J. S. Yang. An introduction to the theory of piezoelectricity. Berlin: Springer, 2005.

[2] W. Heywang, K. Lubitz, and W. Wersing. Piezoelectricity: evolution and future of a technology. Berlin: Springer, 2008.

[3] H. F. Tiersten. Linear piezoelectric plate vibrations. New York: Plenum, 1969.

[4] R. D. Mindlin. An introduction to the mathematical theory of vibrations of elastic plates. J. S. Yang (ed.). Singapore: World Scientific, 2006.

[5] J. Wang and J. S. Yang. Higher-order theories of piezoelectric plates and applications. Applied Mechanics Reviews, 2000, 53(4), 87-99.

[6] M. S. Patel, Y.-K. Yong, and M. Tanaka. Drive level dependency in quartz resonators. International Journal of Solids and Structures, 2009, 46(9), 1856-1871.

[7] J. S. Yang. Equations for the extension and flexure of electroelastic plates under strong electric fields. International Journal of Solids and Structures, 1999, 36, 3171-3192.

[8] J. S. Yang and S. H. Guo. Effects of nonlinear elastic constants on electromechanical coupling factors. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2005, 52(12), 2303-2305.

[9] J. S. Yang. Two-dimensional equations for electroelastic plates with relatively large in-plane shear deformation and nonlinear mode coupling in resonant piezoelectric devices. Acta Mechanica, 2008, 196(1-2), 103-111.

[10] M. S. Patel. Nonlinear behavior in quartz crystal resonators and its stablility. Ph. D. dissertation, Rugter Univeristy, 2008.

[11] J. Wang, R. X. Wu, J. K. Du, D. J. Huang, H. P. Hu, and Y. T. Hu. Nonlinear Mindlin plate equations for the thickness-shear vibrations of crystal plates. Proceedings of the 2008 Symposium on Piezoelectricity, Acoustic Wave, and Applications, 87-92, 2008.

[12] R. X. Wu, J. Wang, and J. K. Du. Propagation of nonlinear waves in elastic rods and plates. Sciences in China-G, unpublished.

[13] J. Wang, R. X. Wu, Y.-K. Yong, J. K. Du, and D. J. Huang. An analysis of vibrations of quartz crystal plates with nonlinear Mindlin plate equations. Proceeding of the Joint Conference of 2009 IEEE International Frequency Control Symposium and the European Frequency and Time Forum, 2009, 450-454.

[14] J. S. Yang. Analysis of ceramic thickness shear piezoelectric gyroscopes. Journal of the Acoustical Society of America, 1997, 102, 3542-3548.

[15] A. Abd-alla and G. A. Maugin. Nonlinear phenomena in magnetostrictive elastic resonators. International Journal of Engineering Science, 1989, 27(12), 1613-1619.

[16] R. X. Wu, J. Wang, J. K. Du, and Y. T. Hu. Solutions of nonlinear thickness-shear vibrations of an infinite isotropic plate with the homotopy analysis method. Numerical Algorithms, in press.

[17] S. J. Liao. Beyond perturbation: introduction to the homotopy analysis method. Boca Raton: Chapman & Hall/CRC Press, 2003.

[18] J. Wang, J. K. Chen, and S. J. Liao. An explicit solution of the large deformation of a cantilever beam under point load at the free tip. Journal of Computational and Applied Mathematics, 2008, 202(2): 320-330.

[19] L. M. Gao, J. Wang, Z. Zhong, and J. K. Du. An analysis of surface acoustic wave propagation in functionally graded plates with homotopy analysis method. Acta Mechanica, 2009, 208: 249-258.

Motional Impedance Analysis: Bridging the 'Gap' in Dielectric Transduction

Siddharth Tallur, Tiffany J. Cheng, Suresh Sridaran, Sunil A. Bhave
OxideMEMS Laboratory, Cornell University
Ithaca, NY USA

Abstract — **This paper presents an analytical model to estimate the motional resistance for partial air gap capacitively-transduced MEMS resonators. This model serves as a link between the well formulated analytical models for conventional air gap and internal dielectric transduction schemes, thereby helping decide which scheme is optimal for a given design frequency. Using this model, we simulate and experimentally verify the motional resistance for a 303MHz polysilicon disk resonator within a 5% range of accuracy.**

I. Introduction

Semiconductor electromechanical resonators, with quality factors (Q) often >10000, provide a low-power small-footprint CMOS-compatible alternative to various electrical components in wireless communication and signal processing. On-chip fabrication of high Q micromechanical resonators can enable low-cost, miniature, low-power filter banks for radio front ends. Moreover, such resonators have promising applications in high performance microprocessors for synchronized clocking arrays with low power, jitter and skew.

Conventional air gap capacitively-transduced RF MEMS resonators typically have high motional impedances on the order of a few kΩ [1, 2]. Sub-GHz partial gap resonators with 10 nanometer air gaps achieving impedances <1kΩ have been shown [3, 4]. Extending the frequency of these resonators entails scaling of resonator dimensions, leading to increased motional impedance at higher frequencies. Dielectric electrostatic transduction has several benefits over conventional capacitive air gap transduction schemes for higher frequencies and smaller dimensions [5]. However, there is no clear understanding on which of these transduction schemes is better for a given design frequency and desired mode shape of the resonator. This work presents an analytical treatment of this problem, and insights derived from the equations presented here enable the designer to optimize resonator designs for low motional impedances.

Primary contact: Siddharth Tallur, sgt28@cornell.edu

II. Prior Modeling efforts

Motional impedance models for air gap transducers [6] and internal dielectric transduced resonators [5] are available in literature. However, partial air gap transduction lacks a good analytical Butterworth Van-Dyke (BVD) model in literature that can accurately predict resonator performance for varied applications. Designers often feel the need for a complete model for dielectric transduction that can help them determine the motional impedance for various schemes for their design of choice, thereby enabling them to pick a scheme that works best. An earlier attempt to model partial gap beam resonators [3] was limited by distortions of the beam geometry at the resonator-electrode interface, resulting in significant deviation from measured results. The next section presents a model that is built on the methodology presented in [6]. However, there are many additional factors to be considered for the case of partial air gap transduced resonators, such as motion of the dielectric, acoustic mismatch between the dielectric and resonator body, stress in the resonator etc. For this simple zero-order model we have made several assumptions that are listed in the following section.

III. Assumptions and methodology

A. Assumptions

The model presented here makes several assumptions to simplify the derivation. Consider a schematic of a beam resonator shown in fig. 1 for illustration. The resonator body comprises of a resonator body of length L and dielectric of length d_2 at both ends. The air gap is d_1 and the dielectric on the electrode has length d_e. The dielectric constant of the dielectric is ε_2 and that of air is ε_1. The various assumptions made are:

1. No acoustic mismatch between dielectric and resonator body.
2. Dielectric on the electrode is rigid.
3. Expressions for electrostatic forces are independent of the mode shape
4. Sense current calculation requires:
   - 4.1. Accounting for both displacement in x (motion at the free end) and displacement in y (motion at the interface between resonator and dielectric).
   - 4.2. Expressing x displacement in terms of y displacement.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

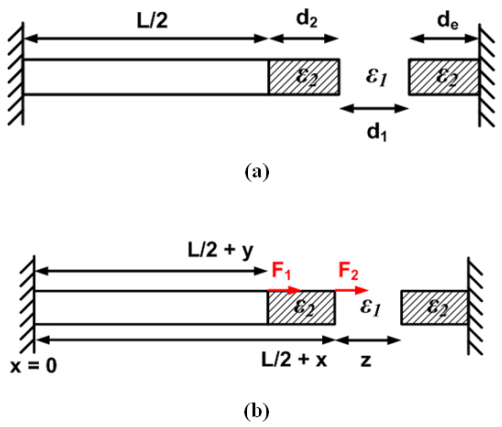


Fig 1. Schematic highlighting half the resonator geometry for a beam resonator for illustrating model assumptions and methodology

B. Methodology

The model essentially follows the derivation for motional impedance presented in [6]. As listed in step 4-5 below, we need to account for motion at the resonator-dielectric interface and air-dielectric interface also.

1. Write down the equation for mode shape in the resonator assuming no acoustic mismatch.
2. Express total capacitance between electrode and resonator body (dielectric and air gap combination) in terms of displacements.
3. Express forces in the form $F_i = \frac{V^2}{2}\frac{\partial C_{tot}}{\partial x_i}$ where V is the total applied voltage, C_{tot} is the total capacitance and x_i is the corresponding displacement.
4. Write down displacement amplitude, and hence the velocity amplitude at the free end of the resonator in terms of effective force (sum of individual forces shaped by the mode shape), stiffness and the quality factor.
5. Use these expressions to write down sense current as $i_{sense} = V_{dc}\frac{dC_{tot}}{dgap} = V_{dc}\frac{\partial C_{tot}}{\partial gap}\frac{\partial gap}{\partial t}$, where V_{dc} is the DC bias applied across the resonator.
6. Express the motional resistance R_x in terms of the applied RF voltage v_{ac} as $R_x = \frac{v_{ac}}{i_{sense}}$
7. Comparing the expression for effective force from step 4 with $F_{net} = \eta v_{ac}$, find the effective electromechanical transduction factor η.
8. Expressions for the inductance and capacitance are given by $L_x = \frac{M}{\eta^2}, C_x = \frac{\eta^2}{K}$, where M is the effective mass and K is the stiffness.

C. Equations for a beam resonator

The equations for a beam resonator can be obtained by following the methodology presented in sub-section B. For the schematic presented in fig. 1, the mode shape can be written as in equation (1), where U_0 is the amplitude of displacement, $k_n = \frac{(2n+1)\pi}{(L+2d_2)}$ is the wave number and ω is the frequency.

$$u(x) = U_0 \sin(k_n x) e^{j\omega t} \tag{1}$$

The total capacitance, C_{tot} can be written as

$$C_{tot} = \frac{\varepsilon_2\varepsilon_1 A}{\varepsilon_1 d_e + \varepsilon_2(d_1+d_2) + x(\varepsilon_1-\varepsilon_2) - \varepsilon_1 y} \tag{2}$$

The forces at resonator-dielectric and air-dielectric interfaces can be written as follows, where V is the voltage applied at the electrode. For our purposes we use expressions obtained by substituting x = d_2 and y = 0. Moreover, we have $\frac{V^2}{2} = V_{dc} v_{ac}$

$$F_1 = \frac{V^2}{2}\frac{\partial C_{tot}}{\partial y} = \frac{V^2}{2}\frac{\varepsilon_2\varepsilon_1^2 A}{[\varepsilon_1 d_e + \varepsilon_2(d_1+d_2) + x(\varepsilon_1-\varepsilon_2) - \varepsilon_1 y]^2} \tag{3}$$

$$F_2 = \frac{V^2}{2}\frac{\partial C_{tot}}{\partial x} = -\frac{V^2}{2}\frac{\varepsilon_2\varepsilon_1 A(\varepsilon_1-\varepsilon_2)}{[\varepsilon_1 d_e + \varepsilon_2(d_1+d_2) + x(\varepsilon_1-\varepsilon_2) - \varepsilon_1 y]^2} \tag{4}$$

$$F_{net} = F_1 \sin\left(\frac{k_n L}{2}\right) + F_2 \tag{5}$$

The expression for displacement can be written down as in equations (6)-(7) using expression for stiffness from [5].

$$U_0 = \frac{F}{|b\omega_0|} = \frac{F}{\left|\frac{m\omega_0^2}{Q}\right|} = \frac{FQ}{K} \tag{6}$$

$$\therefore U_0 = \frac{2Q(L+2d_2)}{Y(2n+1)^2\pi^2}\frac{\varepsilon_2\varepsilon_1 V_{dc} v_{ac}\left\{\varepsilon_1 \sin\left(\frac{k_n L}{2}\right) - (\varepsilon_1-\varepsilon_2)\right\}}{[\varepsilon_2 d_1 + \varepsilon_1 d_2 + \varepsilon_1 d_e]^2} \tag{7}$$

Then the sense current can be written as

$$i_{sense} = V_{dc}\frac{dC_{tot}}{dgap} = V_{dc}\frac{\partial C_{tot}}{\partial gap}\frac{\partial gap}{\partial t} = V_{dc}\frac{\partial C_{tot}}{\partial y}\frac{\partial y}{\partial t} \tag{8}$$

While carrying out this derivation, we express x in terms of y. The reason for this is that motion at the free end also contributes to the sense current which wouldn't be captured if we only carry out a partial differentiation of C_{tot} with y.

$$x = d_2 + \frac{y}{\sin\left(\frac{k_n L}{2}\right)}, \frac{\partial y}{\partial t} = 2\pi f_n U_0 \sin\left(\frac{k_n L}{2}\right) \tag{9}$$

Thus, we can get an expression for R_x as follows:

$$R_x = \frac{v_{ac}}{i_{sense}} = \frac{(2n+1)\pi\sqrt{Y\rho}}{2QV_{dc}^2\varepsilon_1^2\varepsilon_2^2 A}\frac{[\varepsilon_2 d_1+\varepsilon_1 d_2+\varepsilon_1 d_e]^4}{\left\{\varepsilon_1 \sin\left(\frac{k_n L}{2}\right)-(\varepsilon_1-\varepsilon_2)\right\}^2} \quad (10)$$

We obtain η from (5), and thus can express L_x and C_x as follows:

$$L_x = \frac{\rho(L+2d_2)[\varepsilon_2 d_1+\varepsilon_1 d_2+\varepsilon_1 d_e]^4}{\varepsilon_1^2\varepsilon_2^2 V_{dc}^2 A\left\{\varepsilon_1 \sin\left(\frac{k_n L}{2}\right)-(\varepsilon_1-\varepsilon_2)\right\}^2} \quad (11)$$

$$C_x = \frac{2\varepsilon_1^2\varepsilon_2^2 V_{dc}^2 A(L+2d_2)\left\{\varepsilon_1 \sin\left(\frac{k_n L}{2}\right)-(\varepsilon_1-\varepsilon_2)\right\}^2}{Y\pi^2(2n+1)^2[\varepsilon_2 d_1+\varepsilon_1 d_2+\varepsilon_1 d_e]^4} \quad (12)$$

D. Equations for a disk resonator

Beam resonators have significant mode distortion at the resonator-electrode interface [3] that are not accounted for in this model. However, we can establish the validity of this model by deriving a similar set of equations for disk resonators which do not have mode distortion.

Consider the disk resonator geometry in fig. 2. The mode shape for disk resonators can be written as [8] in (13) where $k_0 = \omega_0(R+d_2)\sqrt{\frac{\rho(1-\nu^2)}{E}}$

$$u(r) = A\frac{k_0}{R}J_1\left(\frac{k_0 r}{R}\right)e^{j\omega t} \quad (13)$$

The total capacitance can be written as follows:

$$C_{tot} = \frac{t}{\frac{1}{\theta\varepsilon_2}\ln\left(\frac{R+d_2+d_1+d_e}{R+d_2+d_1}\right)+\frac{1}{\theta\varepsilon_1}\ln\left(\frac{R+d_2+d_1}{R+d_2+x}\right)+\frac{1}{2\pi\varepsilon_2}\ln\left(\frac{R+d_2+x}{R+y}\right)} \quad (14)$$

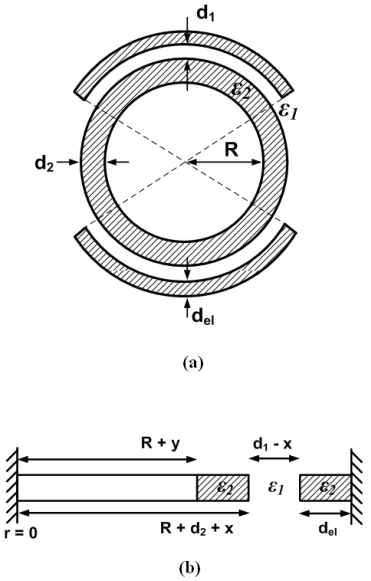


Fig 2. Schematic for a partial gap disk resonator highlighting (a) the geometry, (b) displacements at various interfaces

One can derive the stiffness (and hence effective mass) from the modeling presented in [8]

$$K = \pi\rho t(R+d_2)^2\omega_0^2\frac{J_2^2(k_0)}{J_1(k_0)} \quad (15)$$

Also, the forces and hence the effective force can be written down as in equations (16)-(18)

$$F_1|_{y=0,x=0} = \frac{V^2}{2}\frac{\partial C_{tot}}{\partial y} = V_{dc}v_{ac}\frac{\partial C_{tot}}{\partial y} \quad (16)$$

$$F_2|_{y=0,x=0} = \frac{V^2}{2}\frac{\partial C_{tot}}{\partial x} = V_{dc}v_{ac}\frac{\partial C_{tot}}{\partial x} \quad (17)$$

$$F_{net} = \frac{F_1 J_1\left(\frac{k_0 R}{R+d_2}\right)}{J_1(k_0)} + F_2 \quad (18)$$

The displacement amplitude can then be written down as follows:

$$U_0 = \frac{FQ}{K} = \frac{QV_{dc}v_{ac}\left[\frac{J_1\left(\frac{k_0 R}{R+d_2}\right)}{J_1(k_0)}\frac{\partial C_{tot}}{\partial y}+\frac{\partial C_{tot}}{\partial x}\right]_{x=0,y=0}}{\pi\rho t(R+d_2)^2\omega_0^2\frac{J_2^2(k_0)}{J_1(k_0)}} \quad (19)$$

Similar to (8)-(9), we obtain

$$i_{sense} = V_{dc}\frac{\partial C_{tot}}{\partial y}\frac{\partial y}{\partial t} \quad (20)$$

$$x = \frac{yJ_1(k_0)}{J_1\left(\frac{k_0 R}{R+d_2}\right)},\ \frac{dy}{dt} = \omega_0 U_0\frac{J_1\left(\frac{k_0 R}{R+d_2}\right)}{J_1(k_0)} \quad (21)$$

Thus we can obtain following expressions for R_x, L_x and C_x

$$R_x = \frac{\pi\rho t(R+d_2)^2\omega_0 J_2^2(k_0)}{QV_{dc}^2\left[\frac{J_1\left(\frac{k_0 R}{R+d_2}\right)}{J_1(k_0)}\frac{\partial C_{tot}}{\partial y}+\frac{\partial C_{tot}}{\partial x}\right]_{x=0,y=0}\left[\frac{dC_{tot}}{dy}\right]_{x=\frac{yJ_1(k_0)}{J_1\left(\frac{k_0 R}{R+d_2}\right)},y=0}J_1\left(\frac{k_0 R}{R+d_2}\right)} \quad (22)$$

$$L_x = \frac{\pi\rho t(R+d_2)^2 J_2^2(k_0)}{J_1(k_0)\left\{V_{dc}\left[\frac{J_1\left(\frac{k_0 R}{R+d_2}\right)}{J_1(k_0)}\frac{\partial C_{tot}}{\partial y}+\frac{\partial C_{tot}}{\partial x}\right]_{x=0,y=0}\right\}^2} \quad (23)$$

$$C_x = \frac{J_1(k_0)\left\{V_{dc}\left[\frac{J_1\left(\frac{k_0 R}{R+d_2}\right)}{J_1(k_0)}\frac{\partial C_{tot}}{\partial y}+\frac{\partial C_{tot}}{\partial x}\right]_{x=0,y=0}\right\}^2}{\omega_0^2\pi\rho t(R+d_2)^2 J_2^2(k_0)} \quad (24)$$

IV. DISCUSSIONS

We compare the model for a beam resonator in the limiting cases of air gap transduction and internal dielectric transduction. The equation for air gap transduction is

presented in (25) [6], and the equation derived from our model is presented in (26).

$$R_x = \frac{\sqrt{KM}}{Q\eta^2} = \frac{(2n+1)\pi d^4\sqrt{Y\rho}}{2QV_{dc}^2\varepsilon_0^2 A} \quad (25)$$

$$R_x = \frac{v_{ac}}{i_{sense}} = \frac{(2n+1)\pi d_1^4\sqrt{Y\rho}}{2QV_{dc}^2\varepsilon_1^2 A} \quad (26)$$

Similarly, equations (27) and (28) present equations for the fundamental mode for internal dielectric transduced resonators in literature [5] and from our model respectively.

$$R_x = \frac{\pi g^4\sqrt{Y\rho}}{2QV_{dc}^2\varepsilon_2^2 A\left\{\sin\left(\frac{k_0 L}{2}\right)\right\}^2} \quad (27)$$

$$R_x = \frac{\pi(d_2+d_e)^4\sqrt{Y\rho}}{2QV_{dc}^2\varepsilon_2^2 A\left\{\sin\left(\frac{k_0 L}{2}\right)\right\}^2} \quad (28)$$

The model converges to the well established models for both cases, thereby serving as a 'bridge' between them.

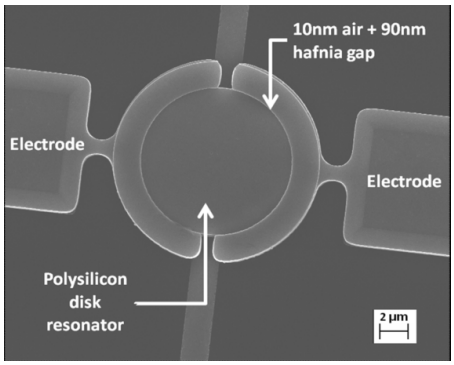


Fig 3. Scanning electron micrograph of the radial breathing mode disk resonator actuated through a 10nm air + 90nm hafnia gap fabricated as in [3]

TABLE I. COMPARISON OF PREDICTED RESISTANCE FROM MODEL TO MEASURED RESISTANCE FOR A WINE GLASS MODE DISK RESONATOR [7]

Process	*Measured R_x (Ω) [7]*	*Predicted R_x (Ω)*
a) 94nm air gap after HF release	2.62k	2.24k
b) After partially filling gap with 30.7nm HfO_2 and sintering at 400°C for 3 minutes	966	869
c) Same conditions as (b) with lower V_{DC}	6.5k	5.23k
d) After sintering at 400°C for another 3 minutes	685	516

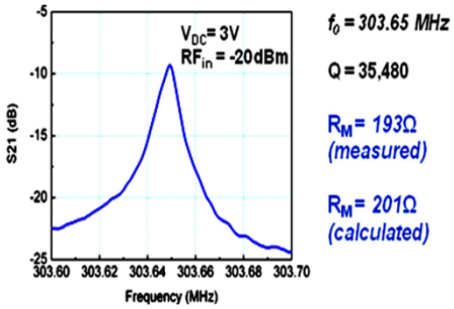


Fig 4. Transmission response of the resonator shown in fig. 3 highlighting measured resistance value

V. COMPARISON OF MODEL TO MEASURED RESULTS

We compared predictions made from our model to measured motional resistance for a radial breathing mode disk resonator designed at OxideMEMS laboratory. The resonator comprises of a poly silicon disk resonator with a radius of 8.83µm actuated through a 10nm air + 90nm hafnia gap. Fig. 3 shows a scanning electron micrograph of this resonator. As shown in the transmission response of the resonator in fig. 4, the predicted motional resistance from the model matches the measured resistance value to within 5% accuracy. To further corroborate the validity of the model, we compared predictions from the model to a wine-glass mode disk resonator reported earlier [7] in table 1.

The model thus compares quite well for measured resistance for these disk resonator geometries. For higher accuracy, the model must include the effects of strain in the dielectric and acoustic mismatch [9] to make it useful for varied applications [10] of partial air gap transduced resonators.

ACKNOWLEDGMENT

The authors extend their thanks to Prof. Ehsan Afshari at Cornell University for interesting discussions on modeling electrostatic forces and NSF career award for support.

REFERENCES

[1] S. Pourkamali, G.K. Ho, F. Ayazi, *IEEE Transactions on Electron Devices,* vol.54, no.8, pp.2024-2030.

[2] Y.-W. Lin, S.-S. Li, Z. Ren, C.T.-C. Nguyen, *Tech. Digest IEDM 2005*, pp. 278-281.

[3] T. Cheng and S.A. Bhave, *Proc. MEMS 2010,* pp. 695-698..

[4] M. Akgul, B. Kim, Z. Ren, C.T.-C. Nguyen, *Hilton Head, SC*, pp.467-470, June 6-10, 2010.

[5] D. Weinstein and S.A. Bhave, *J. Microelectromech. Syst.*, 18.6, pp. 1401-1408, 2009.

[6] V. Kaajakari, A.T. Alastalo, T. Mattila, *IEEE Transactions on UFFC,* 53 (12), pp. 2484-89, 2006.

[7] L.-W. Hung, Z.A. Jacobson, Z. Ren, A. Javey, C.T.-C. Nguyen, *Hilton Head, SC*, pp.208-211, June 1-5, 2008.

[8] Z. Hao, S. Pourkamali, F. Ayazi , *IEEE J. Microelectromech. Syst.* 13, pp. 1054–1062, 2004

[9] E. Hwang, S.A. Bhave, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control,* 57(7), pp. 1664-1672, 2010.

[10] R. Melamud, B. Kim, M. Hopcroft, S. Chandorkar, M. Agarwal, C. Jha, and T. Kenny, *Proc. IEEE Int. Conf.MEMS, Kobe, Japan, 2007*, pp. 199–202.

Lateral Field Excitation of membrane-based Aluminum Nitride resonators

M. Gorisse, A. Reinhardt, C. Billard, M. Borel, E. Defaÿ
CEA, Leti, Minatec Campus
Grenoble, France
Marie.gorisse@cea.fr

T. Bertaud, T. Lacrevaz, C. Bermond
IMEP-LAHC, UMR CNRS
Université de Savoie
Le Bourget du Lac, France

***Abstract*—This paper reports the fabrication of Lateral Field Excitation resonators on a free-standing Aluminum Nitride membrane. We present a very simple fabrication process, and discuss the electrical measurements. In particular, the influence of the electrode width on the generation of parasitic longitudinal waves is demonstrated. We also provide measurements of the temperature dependence of the resonance frequency of the thickness-shear mode, which is of -14 ppm/°C, being much lower than the temperature dependence of the thickness-extensional mode in the same material stack.**

I. INTRODUCTION

Bulk Acoustic Wave resonators have now come to maturity for telecommunication applications, especially in RF front-ends. They have also demonstrated a large potential for sensor application, for their large frequency sensitivity versus added mass. They come mainly in two configurations, differing on how the acoustic waves are trapped within the piezoelectric film and prevented from radiating within the substrate. This is either achieved by leaving an air/solid interface in membrane-based Film Bulk Acoustic Wave Resonators or by forming an acoustic Bragg mirror underneath Solidly Mounted Resonators. Both configurations exploit the resonance of bulk acoustic waves generated by a vertical electric field applied by the means of electrodes positioned at the top and bottom of the piezoelectric field. Alternatively, Lateral Field Excitation (LFE) [1] resonators exploit vertically propagating bulk acoustic waves generated by an electric field parallel to the piezoelectric layer, generated by pairs of coplanar electrodes located on top of the piezoelectric material. Such structures, which should not be mistaken with the same denomination introduced recently for Lamb wave resonators [2], were initially proposed for quartz resonators [1], [3], but have recently been proposed in the thin film form. In particular, Corso *et al.* [4] used Zinc Oxide (ZnO) films and ensured an acoustic isolation by using an acoustic Bragg mirror.

LFE offer some advantages over bulk acoustic wave resonators. First, in piezoelectric materials of hexagonal class, like ZnO or Aluminum Nitride (AlN), they enable the excitation of shear waves which are otherwise not excited by a vertical electric field. These kind of waves are said theoretically to suffer from lower acoustic losses, both because they exhibit less viscoelastic damping, as they do not induce volume changes in the material and because they do not normally couple to pressure waves in the atmosphere surrounding the resonator. Moreover, no metal, which usually exhibit relatively large acoustic losses, lie in the acoustic path. Therefore, LFE resonators are usually expected to exhibit improved quality factors compared to conventional bulk wave resonators. More practically, they are technologically simpler to fabricate, requiring only one metallization level. The fact that they do not require a bottom electrode can be utterly interesting as the growth of piezoelectric film atop electrodes requires using materials with compatible crystalline lattices, for example, when depositing Aluminum Nitride (AlN) films, using Platinum or Molybdenum bottom electrodes, which suffer from poor electrical conductivity. Without the need for a bottom electrode, piezoelectric films can directly be deposited atop the substrate and take for example benefit of its crystalline properties, thus enabling the use of epitaxially grown films which are otherwise tedious to use [5].

In this paper, we focus on AlN based resonators acoustically isolated by forming free standing membranes. Such a structure offers the advantage of being very easy to fabricate, as it does not require the deposition of multiple materials. In section II, we detail the fabrication process used to fabricate such LFE resonators. We then present electrical characterizations in section III and discuss them in section IV. Finally, in section V, we present the influence of temperature on the electrical response of these resonators and use it to determine the effect of temperature on shear bulk waves in AlN.

II. RESONATOR FABRICATION

The devices presented here consist of two square electrodes deposited side-by-side on top of a piezoelectric free-standing membrane. The electrodes are 230x212 µm² rectangles separated by 2, 4 or 8 µm. Images of a device are presented in figure 1.

The fabrication of resonators, summarized in Figure 2, starts with the thermal oxidation of a high resistivity silicon wafer to get a 500 nm-thick SiO_2. A 2 µm thick AlN film is then deposited by reactive pulsed DC sputtering. 200 nm-thick aluminum is then deposited and patterned using dry etching. An additional silicon dioxide layer is deposited and patterned

978-1-61284-110-6/11/$26.00 ©2011 IEEE

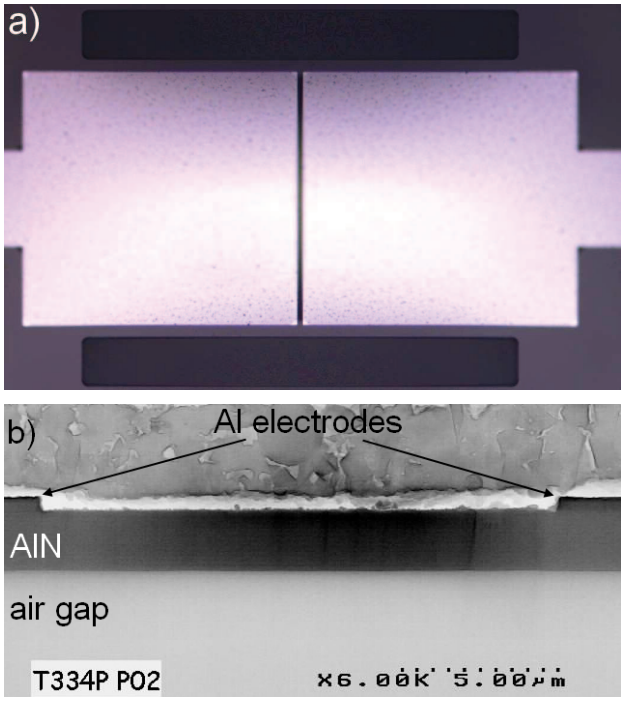


Figure 1: a) Optical image of the LFE device and b) scanning electron micrograph of the cross section of a LFE device

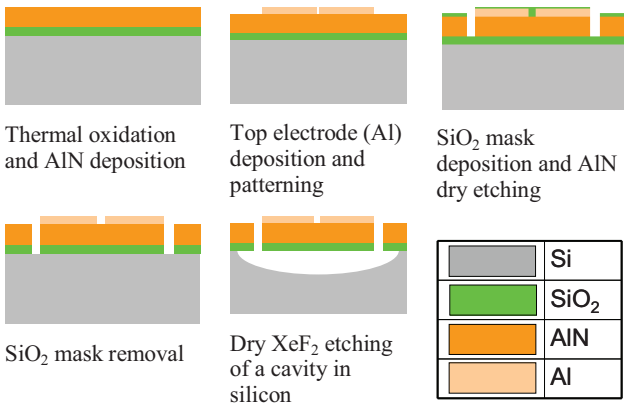


Figure 2: Process flow for the fabrication of the LFE devices

by dry etching to form a mineral mask for the etching of release holes in the aluminum nitride layer.

This oxide mask, as well as the thermal oxide located at the bottom of the release holes through the AlN film is then dry etched.

Finally, a cavity is formed below resonators by an isotropic chemical etching of the silicon substrate using XeF_2. During this step, part of the SiO_2 is removed, especially on the edges of the device. The SEM image of Figure 1.b) is taken on the edge of the device, explaining the lack of SiO_2.

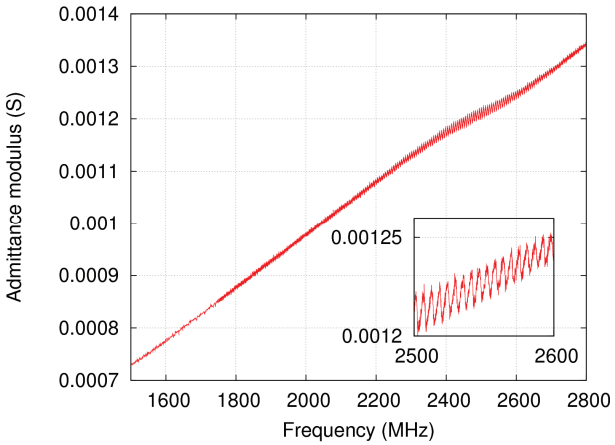


Figure 3: RF measurements of an unreleased LFE device with a 2 µm spacing

III. RF CHARACTERIZATION

Devices are characterized using a vector network analyzer and coplanar RF probes. RF measurements were performed between 10 MHz and 20 GHz, with more accurate frequency steps of 100 kHz between 1.5 GHz and 3.1 GHz to optimize the measurement accuracy in this band. A full 2-port Short-Open-Load-Thru (SOLT) calibration was done. A semi-automatic RF prober was used to ensure repeatable contact resistance between the different devices. Measurements are performed using two-port probes, the second port being shorted, hence corresponding to a signal applied on one of the electrodes when the other one is grounded. For each device presented in the following, we removed the contribution from pads and ground lines by measuring a reference device similar to the measured one but without the resonator electrodes (so as to feature an open circuit) and substracting this measurement from the one of the resonator. The admittance is presented using its modulus.

The first measurement we present here is a High-overtone Bulk Acoustic Resonator (HBAR), obtained before the final fabrication step consisting in releasing the structure. At this stage, the resonator is comprised of the piezoelectric and SiO_2 layers directly located on top of the substrate; the top electrode is the same as in the released structure. Figure 3 presents the electrical measurements of these unreleased devices. First, we observe periodic peaks separated by 5.7 MHz. This frequency spacing between resonances corresponds to the spacing between overtones of the thickness-extensional mode of an AlN layer deposited on top of a 725 µm Si substrate, hence demonstrating the HBAR behavior of the device. The envelope of resonance peaks having its maximum around 2.53 GHz is characteristic of the generation of longitudinal waves by the AlN film. At this stage, no electrical response related to shear waves could be determined.

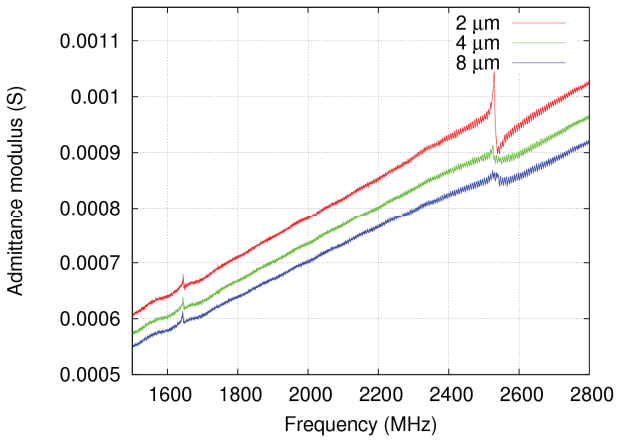


Figure 4: RF measurements of LFE devices with varying spacing

The structures are then released, leaving an AlN/SiO_2 membrane, with their top electrodes. Figure 4 shows the electrical response of the resonators inter-electrodes spacings of 2 to 8 µm. On these curves we can see three different contributions. The first is the resonance located around 1.65 GHz, which corresponds to the first thickness-shear resonance in the AlN film, and hence is the actual LFE resonance. The second one around 2.52 GHz is characterisistic of the thickness-extensional mode in the membrane, being the proof of the generation of also longitudinal waves in the piezoelectric film. The last contribution, represented by small periodic peaks, corresponds to the HBAR resonance we presented above. The LFE and HBAR resonances seem not to depend upon the distance between the electrodes, while the thickness-extensional resonance is attenuated when the inter-electrodes gap increases.

IV. Discussion

To confirm the assumptions related to the nature of the waves excited in the membrane, we calculated the electrical response of our stack using a Mason model. We used a device with an inter-electrode spacing of 2 µm for the fit to get the better resonances. Figure 5 a) shows the fit of the low amplitude multiple resonances using parameters for longitudinal waves. Frequencies fit very well with measurements, revealing a periodicity of about 5.7 MHz, hence corresponding to overtones of the thickness-extensional mode observed on the HBAR structure, confirming our hypothesis.

Using a similar Mason model, also with parameters for the longitudinal wave, we will now focus on the resonance located at 2.52 GHz. Here we model only the membrane made of SiO_2, AlN and Al. To fit correctly this resonance, we needed to dramatically reduce the piezoelectric properties of the piezoelectric material. This was performed by adding an additional capacitance in series with the piezoelectric layer in the electrical model. The result is presented in figure 5b). The position of the thickness-extensional resonance is fitted by

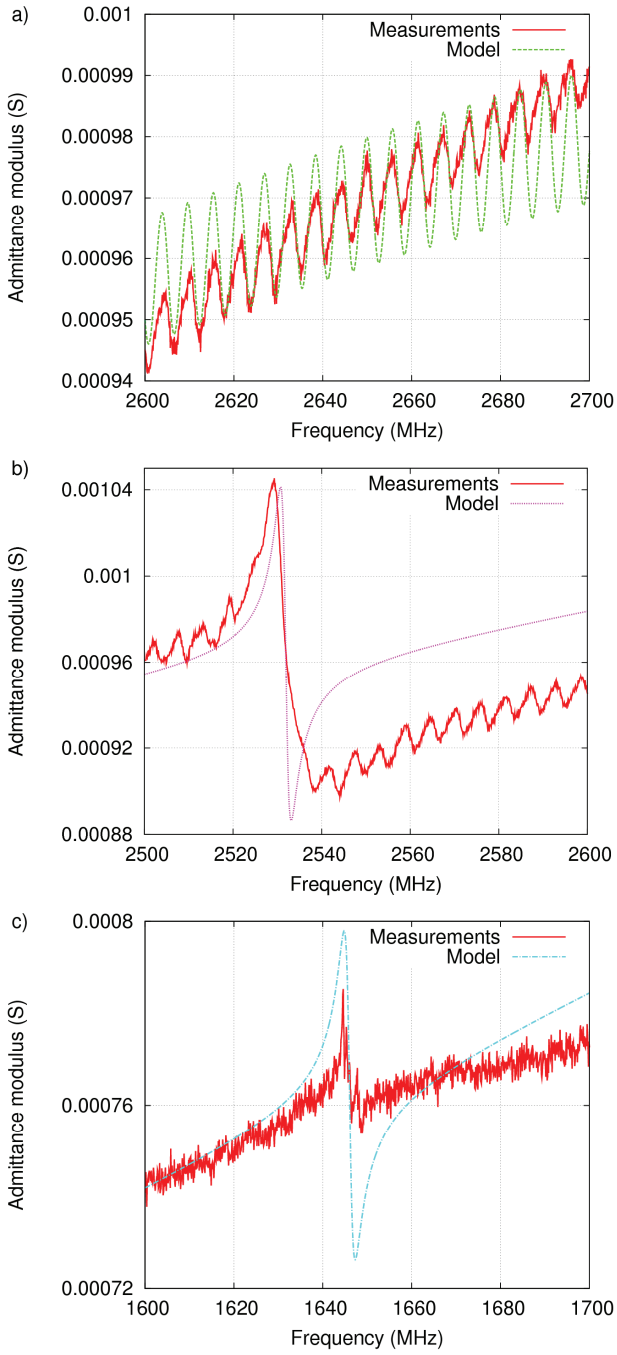


Figure 5: Mason fits a) of the HBAR resonance, b) of the FBAR resonance and c) of the LFE resonance.

varying the thickness of the SiO_2 layer. We deposited 500 nm of SiO_2, but the chemical product of the reaction between Si and XeF_2 is reacting slowly with SiO_2, leaving a thinner layer [6]. This fit gave about 290 nm of remaining SiO_2, but it is over-evaluated (probably around 240 nm).

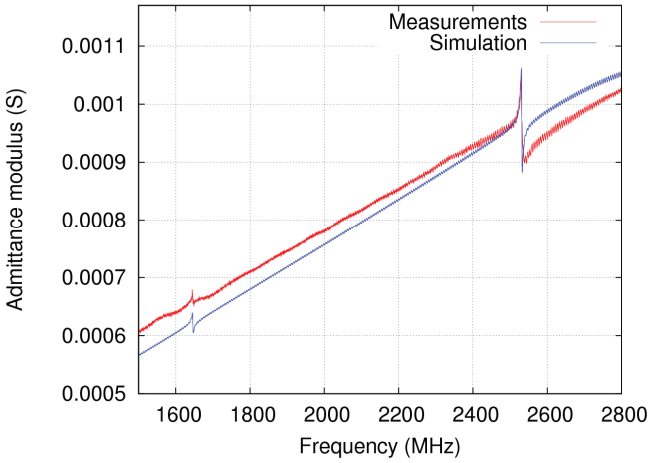


Figure 6: Complete simulation of the device

Finally, we look at the thickness-shear resonance. We used the acoustic velocity of shear waves in AlN and an estimation of the capacitance formed by the two top electrodes. The frequency is closely related to the thickness of the AlN layer. By fitting the frequency, we conclude that the AlN thickness is about 1.78 µm, which is close to what was observed experimentally on SEM pictures and through reflectometric measurements, which is around 1.87 nm.

We combined all these contributions together, assuming that each of them is related to parts of the device, and connecting each model in parallel. Each contribution was ponderated by its respective static capacitance value, which contributed to the overall electric response. The result is presented in Figure 6. The simulation is qualitatively very close to the measurements, presenting the same kind of resonances. Because of its high electromechanical factor, the relative contribution of the thickness-extensional mode needed to remain very low (close to 0.5 % of the total static capacitance), even when using a series capacitor to decrease it. The major contribution was attributed to the thickness-shear mode, and the remaining to the high-overtone resonator response. The resulting electrical schematic is shown in Figure 7.

To provide a physical meaning to each contribution, we performed Finite Element (FEM) simulations. Figure 8 shows that, as expected, most of the electric field strength is located in the gap between the two electrodes, corresponding to the thickness-shear resonance. However, although it may be considered horizontal in most of the inter-electrode gap, simulations reveal that it is vertically aligned under the electrodes, where it is normal to the equipotential surfaces. Therefore, under the electrodes, the piezoelectric material is excited by a vertical electric field, from which the longitudinal waves originate. Although the electric potential drop across the thickness of the AlN film remains low, the surface over which this happens is relatively large, explaining the relatively large

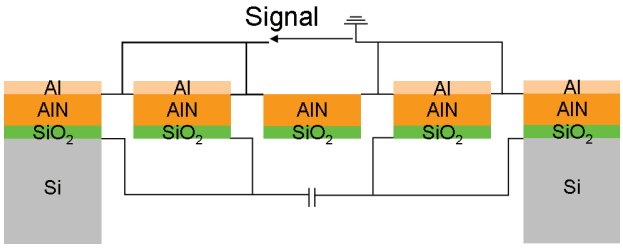


Figure 7: Equivalent circuit for the LFE resonator

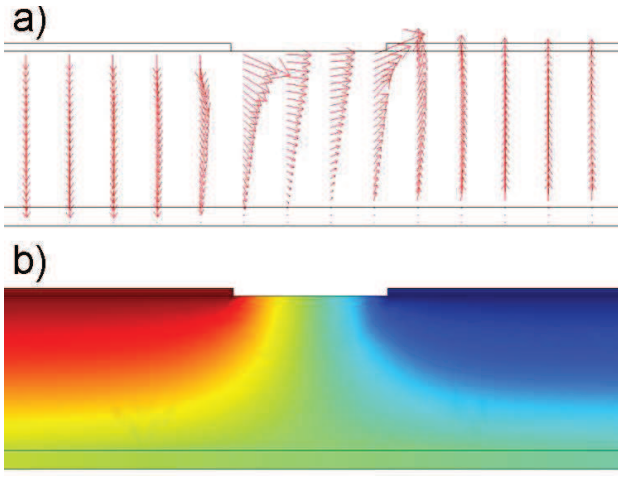


Figure 8: Electric field (a) and electric potential (b) distribution obtained from FEM simulations

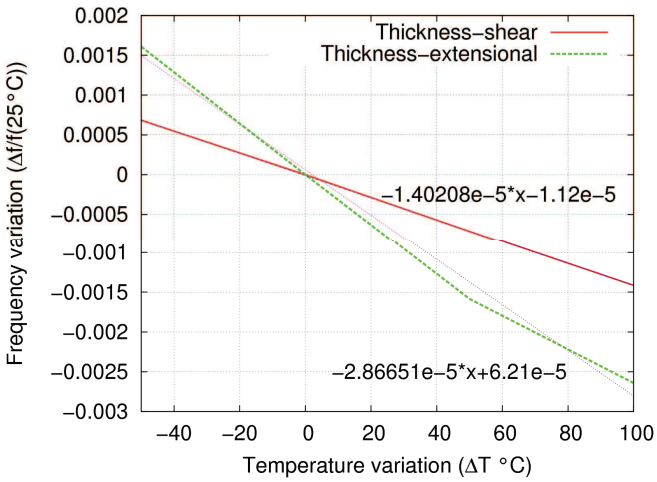


Figure 9: Relative frequency shifts of the thickness-shear and thickness extensional resonances with temperatures

contribution of the thickness-extensional resonance. When the spacing between the two top electrodes is reduced, the electric field distribution in the inter-electrodes gap is only weakly affected, while the strength of the electric field under the electrodes is much more affected. This is the reason why the thickness-extensional mode seems more sensitive to the spacing between electrodes.

The generation of longitudinal waves due to a vertical electric field occurs also at the edges of the resonator, where electrodes are no more located on top of the air-gap: longitudinal waves propagate then both in the piezoelectric film and in the substrate, generating overtone resonances.

V. TEMPERATURE BEHAVIOUR

After having identified each contribution, we measured the frequency shifts of the thickness-extensional and thickness-shear resonances with temperature using a 4 kHz step around each resonance. Their evolution is shown in Figure 9. The temperature coefficient of variation of the thickness extensional resonance is measured to be of –29 ppm/°C, being in agreement with the values usually considered for FBAR [7] resonators. For the thickness-shear mode, the coefficient of temperature is measured to be of –14 ppm/°C, being significantly smaller than the thermal sensitivity of the thickness-extensional mode. This result is consistent with data available in the literature [8].

VI. CONCLUSION

This paper reports the first example so far of Lateral Field Excitation (LFE) resonator fabricated on a free-standing AlN membrane. The electrode shape were not fully optimized, and in particular were very large, leading to the generation of parasitic longitudinal waves due to vertically oriented electric field out of the inter-electrode gap. These parasitic contributions are noticed on the electric response and tend to decrease the performances of the thickness-shear mode. Future work will look at reducing them by reducing the width of the electrodes. Nevertheless, we also reported the characterization of the thermal dependence of frequency of the thickness-shear mode, which was found to be of –14 ppm/°C, being much lower than that of the thickness-extensional mode usually exploited in AlN.

REFERENCES

[1] R. Bechmann, “Parallel field excitation of thickness modes of quartz plates”, Proceedings of the 1960 International Frequency Control Symposium, pp. 68-88 (1960).

[2] C. Zuo, J. Van der Spiegel and G. Piazza, “1.05-GHz CMOS oscillator based on lateral-field-excited piezoelectric AlN contour-mode MEMS resonators”, IEEE Trans. Ultrason. Ferroelect. and Frequ. Contr. vol 57, pp. 82-87 (2010).

[3] A.W. Warner, “Use of parallel-field excitation in the design of quartz crystal units”, Proceedings of the 1963 International Frequency Control Symposium, pp. 248-266 (1963).

[4] C.D. Corso, A. Dickherber and W.D. Hunt, “Lateral field excitation of thickness shear mode waves in a thin film ZnO solidly mounted resonator”, Journal of Applied Physics 101, 054514 (2007)

[5] M. Pijolat, J.C. Moreno, A. Reinhardt, C. Billard, D. Mercier, M. Aïd, F. Semond, S. Ballandras, E. Defaÿ, “16 GHz acoustic resonator with ultra thin epitaxial Aluminum Nitride films”, Proceedings of the 2011 International Frequency Control Symposium (2011).

[6] J.F. Veyan, M.D. Halls, S. Rangan, D. Aureau, X.M. Yan and Y.J. Chabal, “XeF_2-induced removal of SiO_2 near Si surfaces at 300 K: an unexpected proximity effect”, Journal of Applied Physics 108, 114914 (2010).

[7] A. Artieda, P. Muralt, “Fabrication and properties of high-Q AlN/SiO_2 composite TFBAR’s for above IC oscillators”, Proceedings of the 2007 IEEE Ultrasonics Symposium, pp. 1164-1167 (2007).

[8] R.R. Reeber, K. Wang, “High temperature elastic constant prediction of some group-III nitrides”, MRS Internet Journal Nitride Semiconductor Research vol. 6, n°3 (2001).

The Study of Activation Energy(Ea) by Aging and High Temperature Storage for Quartz Resonators' Life Evaluation

Chun-Nan Shen, Min-Chiang Chao, Xiao-Wei Yang, Colin Chang
TXC(NINGBO) CORPORATION
Ningbo Zhejiang, China 315800
E-mail: obi-wan@txc.com.tw

Abstract—**This paper studied the evaluation of activation energy(Ea) based on the experimental data from quartz resonators' biased aging and high temperature storage. 5.0 mm × 3.2 mm metal seam sealed quartz resonators were used in this study. To assure the results are trustable, two approach methods were used for Ea evaluation. The first method is data fit by Arrhenius model for biased aging data. The second method is MTBF evaluation based on high temperature storage data. Both methods conclude the close Ea value, about 0.578eV, and which means 85℃ for 7days accelerated aging is equivalent as 0.83 year at 25℃.**

I. INTRODUCTION

For cost and efficiency consideration, electronic device manufacturers usually cannot take a long time room temperature testing for product life evaluation. Instead, most of the electronic devices are through the reliability acceleration tests [1, 2, 3] to estimate the device life. The predominant method to estimate the equivalent life of passive devices is based on Arrhenius's law. Arrhenius proposed the model to describe how temperature affects the materials inter diffusion as following equation.

$$AFT = e^{\left[\frac{Ea}{K_B}\cdot\left(\frac{1}{Tu}-\frac{1}{Ta}\right)\right]} \quad (1)$$

Where AFT is the temperature acceleration factor, which means equivalent life ratio, K_B = 8.616 × 10^{-5}(eV/°K) — Boltzmann's constant, Ta(°K) is the temperature of acceleration environment, Tu(°K) is the device operated temperature, and Ea is the activation energy. The common reliability accelerating temperatures of quartz resonator are 125℃, 105℃ and 85℃. The AFT (or the life estimation) will be affected by Ea value a lot, as shown in Table I. For the quartz resonator's Ea studies, some researches were already done more than 20 years ago [4, 5, 7]. However, the device size and structure changed a lot during past decades. It is necessary to re-evaluate the Ea value for today's quartz crystal resonators. To assure the results are trustable, two approach methods were used for Ea evaluation. The first method is data fit by Arrhenius model for biased aging data. The second method is MTBF evaluation based on high temperature storage data.

TABLE I. PRODUCT ACTIVATION ENERGY(EA) VERSUS TEMPERATURE ACCLEARTION FACTOR (AFT) AT 125℃

T_L(℃)	T_H(℃)	Ea(eV)	K(eV/ °K)	Tu(°K)	Ta(°K)	AFT
25	125	0.7000	$8.617*10^{-5}$	298	398	943.17
25	125	0.6000	$8.617*10^{-5}$	298	398	354.53
25	125	0.5000	$8.617*10^{-5}$	298	398	133.26
25	125	0.4000	$8.617*10^{-5}$	298	398	50.09
25	125	0.3000	$8.617*10^{-5}$	298	398	18.83

Like common microelectronic devices, the function of aging frequency change can be expressed as Equation (2) which relates with current, temperature, and time [6]

$$\frac{\Delta f}{f}(i,T,t) = R(i)\cdot R(T)\cdot R(t) \quad (2)$$

For the sake of simplicity, this study we focused on Ea for thermal effect only. Which means we need to minimize the R(i) effect in experiments and neglect it in calculations. 3 elevated temperatures, 85℃, 105℃, 125℃ were selected to make thermal effects stronger, and no room temperature condition in this study. After neglect electrical current effect (or drive level effect), Equation (2) can be re-state as following

$$\frac{\Delta f}{f}(T,t) = R(T)\cdot \ln(1+bt) \quad (3)$$

Where $R(T) = C\cdot e^{\left(\frac{-Ea}{K_B\cdot T}\right)}$, b is constant, and for a specified time, Equation (3) can be expressed as the relationship of frequency change versus to temperature in Equation (4).

$$\ln\frac{\Delta f}{f} = \ln C - \frac{Ea}{K_B}\cdot\frac{1}{T} \quad (4)$$

II. Approach A: Activation Energy for Aging Testing

A. Aging Experiment

We took 60 pieces samples for each elevated temperature, 85℃, 105℃, 125℃. 5.0mm × 3.2mm metal seam sealed, 40MHz fundamental quartz crystal resonators were used. Each sample was working in separated oscillation circuit simultaneously. Frequency were measured daily for 1000hrs (about 42days). The results are shown as Fig. 1.

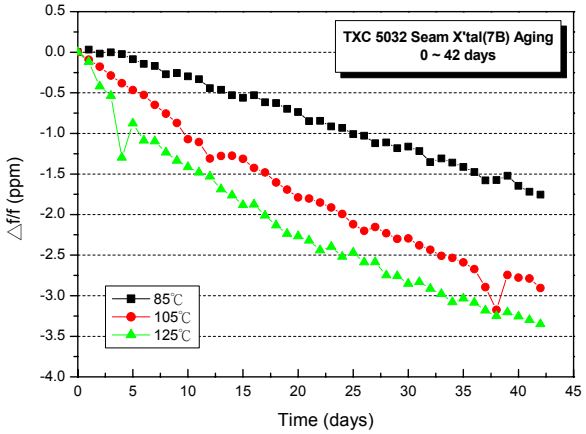


Figure 1. The aging frequency change of 5.0mm × 3.2mm metal sealed quartz resonator for 1000 hrs

TABLE II. EA FIT FROM FIGURE 1 FOR SELECTED DAYS

Days	7	14	21	31	42
Ea	0.578	0.372	0.311	0.270	0.200

B. Activation Energy Estimation through Arrhenius Prediction Modle

In Equation (3), if bt >> 1 it can be simplified as Equation (5) and Equation (6).

$$\frac{\Delta f}{f} = R(T)\ln b + R(T)\ln t \quad (5)$$

$$\frac{\Delta f}{f} = C\cdot e^{-\frac{Ea}{K_B\cdot T}}\cdot \ln t \quad (6)$$

For a lower temperature and time, T_L and t_L, we can get the frequency change from Equation (6). Similarly, we can have another frequency change for a higher temperature and time, T_H and t_H.

$$if\ \delta f_L = \frac{\Delta f}{f}(T_L, t_L), \text{and } \delta f_H = \frac{\Delta f}{f}(T_H, t_H)$$

$$\Delta = \frac{\delta f_H}{\delta f_L} = \frac{\ln t_H}{\ln t_L}\cdot e^{\frac{Ea}{K_B}\left(\frac{1}{T_L}-\frac{1}{T_H}\right)} = \frac{\ln t_H}{\ln t_L}\mathrm{A} \quad (7)$$

$$\Delta = \frac{\delta f_H}{\delta f_L} = \frac{\ln t_H}{\ln t_L}\cdot \exp\frac{1}{KB}\left(\frac{1}{TL}-\frac{1}{TH}\right)Ea \quad (8)$$

For now, we can calculate the Δ ratio from Equation (8) by specified Ea, and we can also calculate the Δ ratio from experiment data, and compare these two results to verify the Ea is suitable or not. For T_H=105 ℃, t_H=7 days, and T_L=85℃, t_L = 7, 14, 21, 31, 42 days, we can get the comparison results as Table III, and the ratio is close when Ea = 0.578 eV.

TABLE III. THE ARRHENIUS VERIFICATION MODEL OF 105℃ VERSUS TO 85℃

105 vs 85 Verification						
Formula Prediction by Real Ea						Real Ratio
	7	14	21	31	42	NA
Ea	0.578013	0.372427	0.310908	0.269896	0.200391	NA
△7/7	3.212913	2.070151	1.728199	1.500231	1.113883	3.84529
△7/14	2.369043	1.526427	1.274289	1.106196	0.821322	2.411959
△7/21	2.053537	1.32314	1.104581	0.958875	0.71194	2.123997
△7/31	1.820636	1.173076	0.979305	0.850124	0.631195	1.954337
△7/42	1.672711	1.077765	0.899737	0.781052	0.579911	1.655865

We compared different temperature and time conditions (125℃, 7days vs. 105℃, 42days and 105℃, 7days vs. 85℃, 42days) and found the ratio will closest when Ea = 0.578eV as shown in Fig. 2.

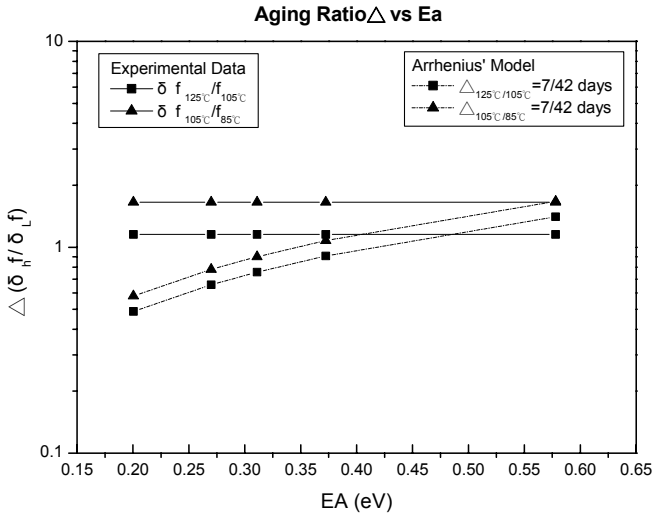


Figure 2. The verification model of 7days at higher tempearure compare to 42days at lower temperature

III. Approach B: Activition Engery Based on MTBF Estimation with High Temperature Storage Data

The classical Arrhenius equation show as Equation (9), where t is the time, T is the temperature, replace t as the definition of MTBF in reliability, Equation (9) could further express as Equation (10).

$$\ln t = a + b \cdot \frac{1}{T} \tag{9}$$

$$\ln(MTBF) = \ln\left(\frac{n \times t}{r}\right) = \frac{Ea}{K_B} \cdot \frac{1}{T} \tag{10}$$

Where n is the sample number, t is the experimental time, r is the number of failure device. TABLE IV is the experimental results of high temperature storage [8]. The failure number could be calculated by various device specification, but the meaningful MTBF must satisfied the normal electronics device's failure principle, which means the failure number should between 1 and total samples. Based on it, we selected the blue area as meaningful MTBF area and used as Ea estimation. The results were showed as Table V, the average Ea is 0.569, which is close to 0.578 as estimated by approach A. Because high temperature storage had only thermal effect, no drive level effect inside, and the Ea is close to elevated temperature biased aging test, it can also show thermal effect is dominate in elevated high temperature biased aging.

TABLE IV. The Failure Nomber of Different Device Specification for High Tempeerature Storage

Temp (℃)	Device Houre (hr)			Failure Number						
	Samples	Time(H)	Total Time	±4ppm ±4ohm	±4.5ppm ±4.5ohm	±4.8ppm ±4.8ohm	±5ppm ±5ohm	±5.5ppm ±5.5ohm	±5.8ppm ±5.8ohm	±6ppm ±6ohm
85℃	30	1074	32220	18	8	**1**	**1**	**0**	**0**	0
105℃	30	1074	32220	30	26	**20**	**18**	**6**	**2**	0
125℃	30	1074	32220	30	28	**26**	**25**	**20**	**15**	11
150℃	30	1074	32220	30	30	**29**	**27**	**21**	**20**	14
85℃	30	1921	57630	30	25	19	15	4	**1**	**1**
105℃	30	1921	57630	30	29	29	29	21	**14**	**10**
125℃	30	1921	57630	29	29	29	27	19	**16**	**12**
150℃	30	1921	57630	30	30	29	29	25	**22**	**16**

TABLE V. The Averaged Ea of High Temperature Storage is 0.569(eV)

Ea(eV)					
±4.5ppm ±4.5ohm	±4.8ppm ±4.8ohm	±5ppm ±5ohm	±5.5ppm ±5.5ohm	±5.8ppm ±5.8ohm	±6ppm ±6ohm
	0.6313	**0.6220**	**0.3767**	**0.6946**	
				0.5735	**0.5180**
Average = 0.5693					

IV. Conclusion

We studied the evaluation of activation energy (Ea) based on the experimental data from quartz resonators' biased aging and high temperature storage. 5.0 mm × 3.2 mm metal seam sealed quartz resonators with fundamental frequency 40MHz were used in this study. To avoid the interference between thermal and drive level effect, we selected 3 elevated temperatures in aging test. To assure the results are trustable, two approach methods were used for Ea evaluation. The first method is data fit by Arrhenius model for biased aging data. The second method is MTBF evaluation based on high temperature storage data. Both methods conclude the close Ea value, about 0.578eV (for thermal effect), and which means 85℃ for 7days accelerated aging is equivalent as 0.83 year at 25℃. The aging mechanism and the effects by drive level or other factors still need more studies in the future.

Acknowledgment

This study was supported by government torch plan item (No.2007GH010123) of The People's Republic of China.

References

[1] Zeng-ZhaoZhang, Yong Pan, "Reliability Prediction of Electronic Devices", 1st ed., Science Press: China, August 2007.

[2] Xi-Sen Wen, "Application and Theory of Reliability Accelerated Experiment", 1st ed., Science Press China, August 2007.

[3] Ying Gu, "Mathematic of Reliability Engineering", 1st ed., Publishing House of Electronics Industry: Beijing, August 2004.

[4] R. Filler et al, "The aging of resonators and oscillators under various test conditions", *Proc. 38th AFCS* , **444**, 1987

[5] A. A. Feinberg, "Parametric failure rate model for quartz crystal device aging with application to surface acoustic wave filters", *Proc. 41th AFCS*, **360**, 1987

[6] D. Epstein, "Application and use of acceleration factors in microelectronics testing", *Solid State Tenchnology*, **116**, 1982

[7] M. R. Miljkovic, G. L. Trifunovic, V. J. Brajovic, *42th AFCS*, **404**, 1988.

[8] Standardization Administration of China, "Constant Stress Life Tests and Acceleration Life Test -- General Rule", GB 2689.1-81 , October 1981.

High Frequency High-Order Rayleigh Modes in ZnO/GaAs

J. Pedrós
Cavendish Laboratory, University of Cambridge
CB3 0HE Cambridge, United Kingdom
E-mail: jp495@cam.ac.uk

C. J. B. Ford, C. H. W. Barnes, J. P. Griffiths, G. A. C. Jones
Cavendish Laboratory, University of Cambridge
CB3 0HE Cambridge, United Kingdom

L. García-Gancedo
Department of Engineering, University of Cambridge
CB3 0FA Cambridge, United Kingdom

A. J. Flewitt
Department of Engineering, University of Cambridge
CB3 0FA Cambridge, United Kingdom

F. Calle
Instituto de Sistemas Optoelectrónicos y Microtecnología, Universidad Politécnica de Madrid, 28040 Madrid, Spain

***Abstract*—Strong high-order Rayleigh or Sezawa modes, in addition to the fundamental Rayleigh mode, have been observed in ZnO/GaAs(001) systems along the [110] propagation direction of GaAs. The dispersion of the different acoustic waves has been calculated and compared to the experimental data. The bandwidth and impedance matching characteristics of the multimode SAW delay lines operating at high frequencies (2.5-3.5 GHz regime) have been investigated.**

I. Introduction

ZnO films are known to enhance the piezoelectric coupling of GaAs substrates, facilitating the integration of surface acoustic wave (SAW) devices with the GaAs electronics. In spite of the small sound velocity mismatch between the two materials, this slow-on-fast structure supports several guided waves in the overlayer. However, only the fundamental Rayleigh wave has been studied so far [1,2].

The dynamic modulation of the band diagram induced by the piezoelectric field of a SAW in GaAs-based systems has allowed the development of single-electron transistors using the split-gate technique [3]. More recently, these constrictions have been combined with lateral *n-p* junctions pursuing a high frequency single-photon source [4]. Other approaches to SAW-assisted anti-bunched photon emitters rely on the controlled injection of excitons into quantum dots [5] or the modulation of their emission by means of the SAW strain fields [6]. All these applications would benefit strongly not only from the enhanced piezoelectric coupling provided by the ZnO but also from the capacity of modulating the optoelectronic systems using waves with different piezoelectric and strain field depth profiles to those of the fundamental Rayleigh mode.

In this paper we report on the characteristics and dispersion of the different acoustic waves supported by the ZnO/GaAs heterostructure and on the performance of the SAW delay lines used to generate them.

II. ZnO Films And Device Design

ZnO films of different thickness *H*, ranging from 1 to 2.5 μm, were deposited on semi-insulating GaAs(001) substrates by high target utilization sputtering (HiTUS) [7] at room temperature. In this technique, an argon plasma is generated in a side chamber and steered onto the target, avoiding the direct contact of the plasma with the substrate. This prevents the ion bombardment of the substrate, providing high quality films with very low stress and defect density even at high deposition rates and at room temperature [7].

The sputtering parameters were optimized in order to obtain highly resistive films oriented along the (001) crystal axis, conditions that ensure a strong piezoelectricity. A Zn target (99.999 %) was exposed to the Ar plasma with a constant target power of 800 W. The flows of Ar and O_2 were 65 and 41 sccm, respectively, providing a deposition rate of ~50 nm/min. The crystal quality of the sputtered ZnO films was assessed by X-ray diffractometry, confirming the unique orientation of the films along the (001) axis. The resistivity of all the films was very high (10^{10} Ω m).

SAW delay lines, formed by split-finger interdigital transducers (IDTs), were patterned by e-beam lithography on

This work has been partially supported by the EPSRC, grant number EP/F063865/1. J. Pedrós aknowledges the support from the Marie Curie Intra-European Fellowship within the 7th European Community Framework Programme. The ICTS Programme from the Spanish Ministerio de Ciencia e Innovación is also acknowledged.

the ZnO films along the [110] direction of the underlying GaAs(001) substrates. The IDTs were formed by Ti/Au (10/10 nm) contacts and had a period λ of 1 μm and a metallization ratio of 0.5, i.e. finger width and pitch of 125 nm. The IDTs had an aperture W of 60 μm, whereas the number of finger pairs N_p was varied in order to modify the bandwidth of the passband. The delay lengths were 1 and 2 mm. The S parameters of the devices were measured with a coplanar probe station connected to a network analyzer, and in some cases (when specifically stated) the devices were also characterized once mounted on a sample holder with SMA connectors.

III. RESULTS

A. *High-order Rayleigh or Sezawa modes*

Fig. 1 shows the insertion loss of identical SAW delay lines fabricated on a series of ZnO films of different thickness on GaAs(001) substrates. The spectra of the devices on films with thickness H ranging from 1 to 1.4 μm present two clear resonances, whereas the device on the thickest film (H=2.5 μm) supports three. The lowest-frequency resonance, at 2.53 GHz for all the devices, corresponds to the fundamental Rayleigh mode (labeled as R). The resonances at higher frequencies are associated with high-order Rayleigh modes, also called Sezawa modes (labeled as S_i, i≥1). These are guided waves in the overlayer that arise since the ZnO/GaAs structure forms a slow-on-fast system, i.e. the bulk transverse velocity in the substrate is larger than in the overlayer. The frequency of the first Sezawa mode (S_1) decreases from 3.20 to 2.86 GHz as the film thickness increases from 1 to 2.5 μm (as indicated by the arrow in Fig.1), whereas the second Sezawa mode (S_2) appears only for the thickest film, at a frequency of 3.08 GHz.

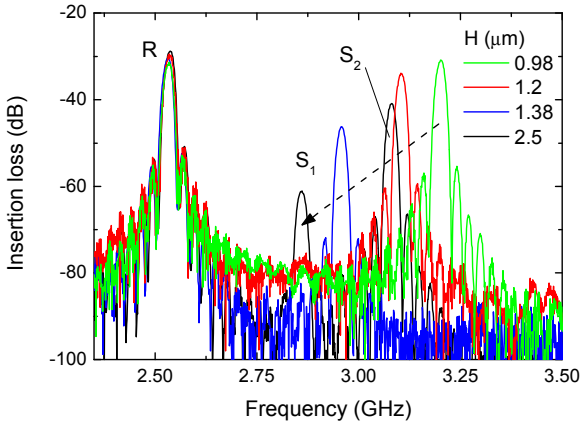


Figure 1. Insertion loss of identical SAW delay lines on ZnO films of different thickness H on GaAs. R, S_1, and S_2 denote the Rayleigh and the first and second Sezawa modes, respectively.

B. *Dispersion of the ZnO/GaAs structure*

The identification of the acoustic modes shown in Fig. 1 is based on the comparison of the experimental data with the numerical calculation of the dispersion of the ZnO/GaAs structure. The calculation is based on a Green's function formalism which was developed to calculate Brillouin spectra [8]. The material parameters of ZnO and GaAs used for the calculation have been taken from Refs. [9] and [10], respectively.

The results of the simulations for the ZnO/GaAs(001) heterostructure with propagation along the GaAs [110] direction are summarized in Fig. 2. ZnO has a wurzite structure (crystal class 6mm) with isotropic elastic properties in the (00.1) plane, so that the propagation direction is referred to that of the underlying cubic structure of GaAs (crystal class -43m) that presents a four-fold symmetry in the elastic properties in the (001) plane. The gray scale in Fig. 2 depicts the magnitude of the shear vertical component of a wave having the given velocity and $kH=2\pi H/\lambda$ value. The experimental data (circles) have been determined from the resonance frequency of the IDTs (Fig. 1). The velocity of the R mode decreases from its value in GaAs for kH=0, to that in ZnO for approximately kH>7. Additionally, S_i modes emerge nearly periodically beyond their threshold values of kH. The velocity of the S_i modes decreases monotonically with increasing kH. The numerical results are in good agreement with the experimental data. The small discrepancy on the S_i modes in the thickest ZnO film could originate from a slight misalignment of the delay line pattern from the [110] direction. The broadening of the R mode is an artifact of the calculation when resolving simultaneously the intense R mode and the less intense S_i modes.

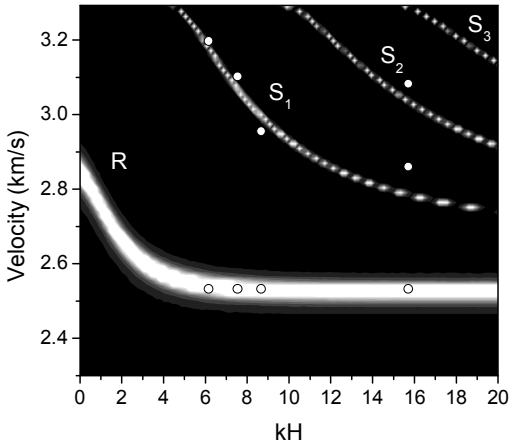


Figure 2. Dispersion of the Rayleigh (R) and Sezawa (S_i) modes in the ZnO/GaAs(001) structure for propagation along the [110] direction of GaAs. The brighter the scale is, the larger the calculated intensity of the mode. The experimental data are shown by the circles.

C. *Device bandwidth*

Figs. 3(a) and 3(b) depict, respectively, the insertion and return loss of a series of SAW delay lines with varying number of finger pairs (N_p) fabricated on a 2.5 μm-thick ZnO film on GaAs(001). The resonances of the R, S_1, and S_2 modes are clearly observed in both the transmission and the reflection spectra for all the cases except for the S_1 mode in the reflection spectra for the lower N_p cases, where it is hardly distinguished from the background.

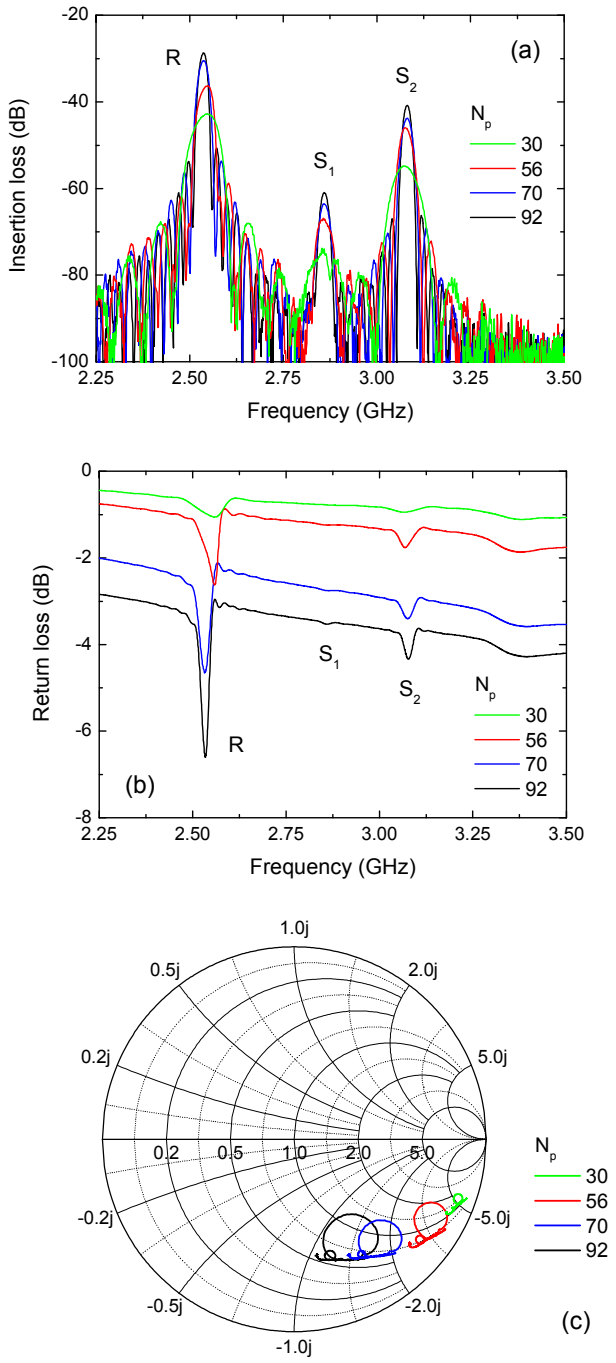


Figure 3. (a) Insertion loss, (b) return loss, and (c) Smith chart (normalized to 50Ω) of SAW delay lines with varying number of finger pairs N_p on a 2.5-μm thick ZnO film on GaAs. R, S_1, and S_2 denote the Rayleigh and the first and second Sezawa modes, respectively.

The variation of N_p modifies the losses of the devices by means of two different mechanisms. On the one hand, the passband of the devices narrows as N_p increases and this translates into a reduction of the losses. This correlation is clearly observed in Figs. 3(a) and 3(b). On the other hand, the increase in N_p produces an increase in the capacitance of the IDT, decreasing its reactance and thus enhancing the impedance matching condition of the device. Fig. 3(c) shows the Smith chart of the impedance of the devices. The traces, in the capacitive half of the chart, move approximately along curves of constant resistance, except at the mode resonances, where a loop is formed. As N_p increases, the traces present larger loops, i.e. stronger resonances, and they shift towards smaller values of capacitive reactance.

The evolution of the bandwith with N_p for the different mode resonances shown in Fig. 3(a) is summarized in Fig. 4. Both, the bandwidth between first nulls on either side of the resonance frequency (BW_{nn}) and between points 3 dB below the main lobe maximum (BW_3) have been extracted from the experimental data in Fig. 3(a). The experimental bandwidths in Fig 4 (labeled as $BW_{nn\ exp}$ and $BW_{3\ exp}$) are compared to their theoretically expected values (labeled as $BW_{nn\ th}$ and $BW_{3\ th}$). The theoretical bandwidth between first nulls is given by the expression $BW_{nn} = 2f_0/N_p$, where f_0 is the resonance frequency of the mode. The theoretical value of the bandwidth at 3 dB is $BW_3 \approx 0.31BW_{nn}$, the characteristic value of a square sinc function at that level. The comparison of the experimental and theoretical values of BW_3 shows an overall good agreement, whereas the experimental values of BW_{nn} are systematically lower than the theoretically predicted ones. This discrepancy is likely to be related to the error in determining the position of the nulls. This error is especially large in the case of the S_1 mode where first nulls appear at a very high insertion loss, as can be noticed in Fig. 3(a).

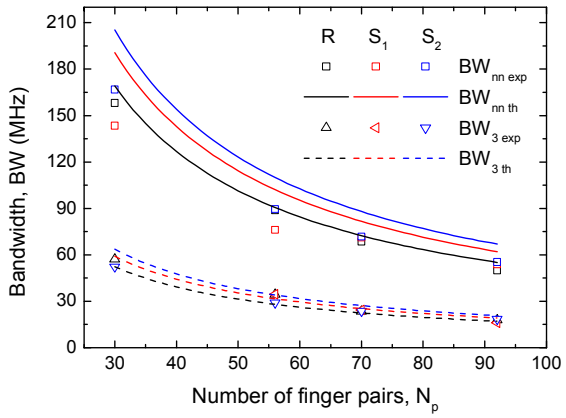


Figure 4. Evolution of the bandwidth of the Rayleigh (R), first Sezawa (S_1), and second Sezawa (S_2) mode resonances with the number of finger pairs in SAW delay lines on a 2.5-μm thick ZnO film on GaAs. See text for details.

D. *Impedance matching*

A stub tuner has been used to impedance match one of the IDTs forming the SAW delay lines under study in order to determine the maximum potential of the devices. The selected device was patterned on a 2.5 μm-thick ZnO film on GaAs(001) formed by IDTs with W=60 μm, λ =1 μm, and N_p=70. The device was mounted on a sample holder with SMA connectors. The contribution of the measurement set-up, including an SMA cable and the stub tuner in its initial contracted position, has been eliminated by means of the calibration.

Fig. 5 shows the return loss spectra of an IDT before and after using the stub tuner to impedance match it, whereas the inset presents the data in the format of a Smith chart, where the variation of the impedance with frequency is depicted. The comparison of the spectra in Fig. 5 and its inset for the untunned IDT with those of a similar IDT (N_p=70) in Figs. 3(b) and 3(c) indicates the contribution of the sample holder. The traces, which are in the capacitive half of the chart when measured on wafer with probes [see Fig. 3(c)], cross over from inductive to capacitive when measured on housing, as shown in the inset of Fig. 5. In particular, the resonance of the R mode lies in the inductive half of the chart. The inductance measured is likely to originate from the wire bondings used to connect the device to the sample holder.

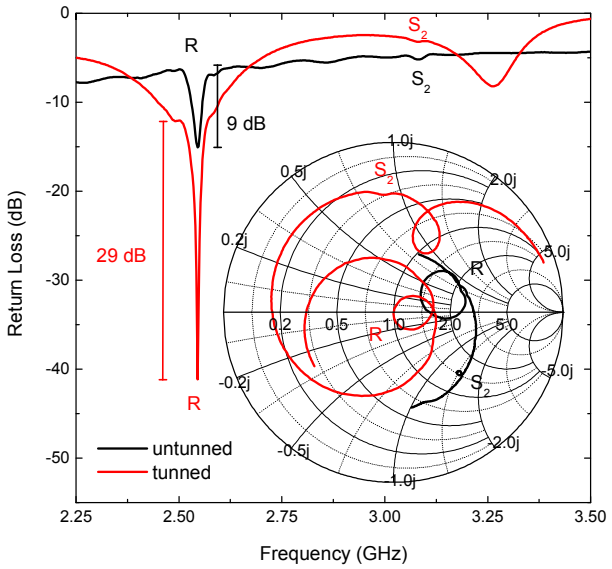


Figure 5. Return loss of an IDT with N_p =70 on a 2.5-μm thick ZnO film on GaAs (device mounted on a sample holder). The inset shows the Smith chart (normalized to 50Ω) of the return loss. R and S_2 denote the Rayleigh and second Sezawa modes, respectively.

When the stub tuner is used to impedance match the R mode, as indicated by its resonance centered at 1 on the real axis (i.e. 50 Ω) of the Smith chart (inset of Fig. 5), the deep associated to that mode in the return loss spectrum (Fig. 5) is enhanced by more than 25 dB. Nonetheless, the stub tuner introduces an additional modulation of the background of approximately 8-10 dB in amplitude and a period of 715 MHz. One of the minima of this modulation coincides with the resonance of the R mode, broadening its base. If that contribution is subtracted, the net enhancement of the magnitude of the R mode amounts to 20 dB, as indicated in Fig. 5.

IV. Conclusions

Strongly oriented and highly resistive ZnO films on GaAs substrates have been obtained by sputtering at room temperature. The different acoustic waves propagating along the [110] direction of GaAs in ZnO/GaAs(001) systems have been experimentally and theoretically investigated. In addition to the fundamental Rayleigh mode, strong high-order Rayleigh or Sezawa modes propagate confined in the overlayer. In particular, the insertion loss of the S_1 mode can be as low as that of the R mode for certain thickness-to-wavelength ratios. The bandwidth and impedance matching characteristics of the different resonances in SAW delay lines operating at high frequency (2.5-3.5 GHz regime) have been investigated.

References

[1] Y. Kim, W. D. Hunt, F. S. Hickernell, R. J. Higgins, and C.-K. Jen, "ZnO films on {001}-cut <110>-propagating GaAs substrates for surface acoustic wave device applications" IEEE Trans. Ultrason. Ferroelec. Cont., vol. 42, pp. 351- 361, 1995.

[2] V. Y. Zhang, J. E. Lefebvre, and T. Gryba, "SAW characteristics in a layered ZnO/GaAs structure for design of integrated SAW filters" Proc. IEEE Ultrason. Symp., pp. 261-264, 2001.

[3] J. M. Shilton, V. I. Talyanskii, M. Pepper, J. E. F. Frost, C. J. B. Ford, C. G. Smith, and G. A. C. Jones, "High-frequency single-electron transport in quasi-one-dimensional GaAs channel induced by surface acoustic waves", J. Phys. Condens. Matter, vol. 8, pp. L531-L539, 1996.

[4] J. R. Gell, P. Atkinson, S. P. Bremner, F. Sfigakis, M. Kataoka, D. Anderson, G. A. C. Jones, C. H. W. Barnes, D. A. Ritchie, M. B. Ward, C. E. Norman, and A. J. Shields, "Surface-acoustic-wave-driven luminescence from lateral *p-n* junction", Appl. Phys. Lett., vol. 89, pp. 243505-1-4, 2006.

[5] O. D. D. Couto Jr., S. Lazic, F. Iikawa, J. A. H. Stotz, U. Jahn, R. Hey, and P. V. Santos, "Photon anti-bunching in acoustically pumped quantum dot" Nat. Photonics, vol. 3, pp. 645-648, 2009.

[6] J. R. Gell, M. B. Ward, R. J. Young, R. M. Stevenson, P. Atkinson, D. Anderson, G. A. C. Jones, D. A. Ritchie, and A. J. Shields, "Modulation of single quantum dot energy levels by a surface-acoustic-wave", Appl. Phys. Lett., vol. 93, pp. 018115-1-4, 2008.

[7] L. García-Gancedo, J. Pedrós, A. J. Flewitt, W. I. Milne, G. M. Ashley, J. Luo, and C. J. B. Ford, "Ultrafast sputtered ZnO thin films with high k_T for acoustic wave device applications" Proc. IEEE Ultrasonics Symp. 2010, in press.

[8] X. Zhang, J. D. Comins, A. G. Every, P. R. Stoddart, W. Pang, and T. E. Derry, "Surface Brillouin scattering study of the suface excitations in amorphous silicon layers produced by ion bombardment", Phys. Rev. B, vol. 58, pp. 13677- 13685, 1998.

[9] G. Carlotti, G. Socino, A. Petri, and E. Verona, "Acoustic investigation of the elastic properties of ZnO films" Appl. Phys. Lett., vol. 51, pp. 1889-1891, 1987.

[10] A. J. Slobodnik, Microwave Acoustics Handbook, 1973.

UHF CMOS-MEMS Bulk Acoustic Wave Resonator

J. Giner, A. Uranga, E. Marigó, J.L. Muñoz-Gamarra and N. Barniol
Electronic Engineering Department. Universitat Autònoma de Barcelona,
08193Barcelona, Spain
Joanjosep.giner@uab.cat

***Abstract*—This paper presents a UHF CMOS MEMS resonator working at 384MHz at its bulky mechanical mode. The novel topology proposed allows us to design the resonator taking into account the coupling area of the movable resonator as an important parameter in the design. The structure is fabricated using a commercial CMOS technology. This fact allows to integrate directly the microelectromechanical structure with the electronic. This work presents the measurement of both, the stand alone resonator and the resonator with an amplification circuitry.**

I. INTRODUCTION

Microelectromechanical resonators are widely used in RF applications, both at HF, VHF and UHF ranges. In particular, CMOS MEMS are a promising solution for the implementation of RF front-ends due the facility of MEMS integration, reduced area and low power consumption. Very competitive devices have been reported using these fully integrated structures not only for RF applications but also for sensing. A CMOS metal cantilever, electrostatically excited, with an exceptional 1 ag/Hz mass sensitivity is used as frequency-determining element of an oscillato [1-2]. A CMOS polysilicon Clampled Clamped shaped beam with a Q of 4400 [3] is an example of CMOS MEMS resonators used for RF filters in the range of HF. In this work, two resonators are electrically coupled in order to obtain a IF filter. In [4] a squared frame resonator based on four FF-Beams anchored by their nodal points was fabricated by an stack of metal and intermetal oxide in a collaborative way to increased the coupling area and achieve a quality factor Q of around 1400. Other interesting solutions can be found in [5] where flexural modes and a stack of CMOS metals are used to obtain a resonator working at HF range. Higher frequency CMOS MEMS flexural resonators have been demonstrated in [6] to reach 200 MHz with a 40 nm gaps. However, the use of the flexural approach has the intrinsic problem of the scaling. An increase of frequency implies a length reduction of the beam and therefore the reduction of the coupling area. Bulk modes offer the solution to the scaling problems and reach very good results at VHF and UHF ranges as it has been proven in [6-9].

In this paper a Bulk Acoustic Resonator working at 380 MHz is presented. The LBAR resonator structure presented in [12] suffers for a limitation in the coupling area. An increase of this area, in order to achieve a better coupling, changes the resonance frequency due to variations in the resonator width.

This paper shows a novel geometry that allows at the same time to preserve the coupling area and increase the resonance frequency.

II. DESIGN

An LBAR resonator working in its longitudinal vibration mode has been selected to implement the novel structure. Figure 1 shows the diagram of the resonator, whose topology is similar to the suspensions founded in [13]. A maximum displacement occurs at the end of the structure while the center of the resonator presents zero movement. This center point is selected as anchorage point.

The design technique is based on the merge of two imaginary ellipses, with major and minor axis of *a* and *b* microns respectively, over a rectangle presenting and area of WxL. The LBAR is defined by the area that separates the two ellipses. As it will be seen, the novel structure allows to predefine the length (W) of the coupling area, and fix the rest of the dimensions to achieve the desired resonant frequency.

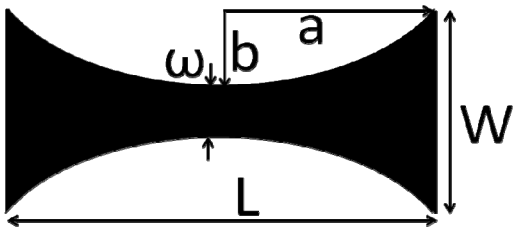


Figure 1. Diagram of the resonator, *b,* corresponds on the minor axis and *a to* the major axes. WxL defines the rectangle over the ellipses.

Equation 1 allows to obtain the frequency as a function of the stiffness of the structure for the desired mode, K_x and the mass, *M.* The mass is obtained by subtracting the mass of the imaginary ellipse ($M=a\cdot b\cdot\pi\cdot t\cdot\rho$) to the mass of the rectangle ($L\cdot W\cdot t\cdot\rho$) in equation (2), being t and ρ the thickness and density of the resonator. The stiffness is obtained using

978-1-61284-110-6/11/$26.00 ©2011 IEEE

expression (3) presented in [13], where E is the young modulus.

The design process starts fixing the length of the coupling surface (*W*) and therefore, the coupling area in each arm of the resonator. Design rules of the technology impose the minimum width of the device near the anchors (*w)*. Therefore, the minor axis of the ellipse is determined by the subtraction of *W* and ω. Finally, the major axis (*a*) length of each arm is determined by the resonant frequency, following expression 1.

The new resonator has been fabricated using the polysilicon capacitance module present in the commercial CMOS technology of the Austria Microsystems (AMS) of 350 nm. The resonator has been dimensioned to operate in the free band of 434 MHz. The previous described design process has been followed to implement the structure. In particular, a coupling length of W=3µm with a minimum width ω=850nm has been fixed. Figure 2 represents the dependence of the major axis with the resonant frequency. Final dimensions are detailed in Table 1

TABLEI TABLE OF DIMENSION OF THE RESONATOR

A	Area of the rectangle (LxW)	21 pm^2
Ae	Area of the ellipse	12.9 pm^2
ρ	PolySi density	2230 kg/cm^3
L	Length of the resonator	7 µm
W	Coupling lenght	3 µm
a	Ellipse major axis	3.5 µm
b	Ellipse minor axis	1.075 µm
K_x	Resonator stiffness	46.764 N/m
E	Young modulus	160 Gpa
ω	Minimum width of the resonator	850 nm
t	PolySi thickness	280 nm
M	Mass of the resonator	6.21 pg

$$f = \frac{1}{2\pi}\sqrt{\frac{K_x}{M}} \tag{1}$$

$$M = (A - A_e)t\rho = \left[2a(2b+\omega) - \pi ab\right]t\rho \tag{2}$$

$$K_x = \frac{4Ebt}{\frac{4a(2b+w)}{\sqrt{w(4b+w)}}\arctan\sqrt{1+\frac{4b}{w}} - a\pi} \tag{3}$$

III. FABRICATION

The new resonator has been fabricated using the polysilicon capacitance module where the movable structure is fabricated using the 280 nm thickness polysilicon 1 layer

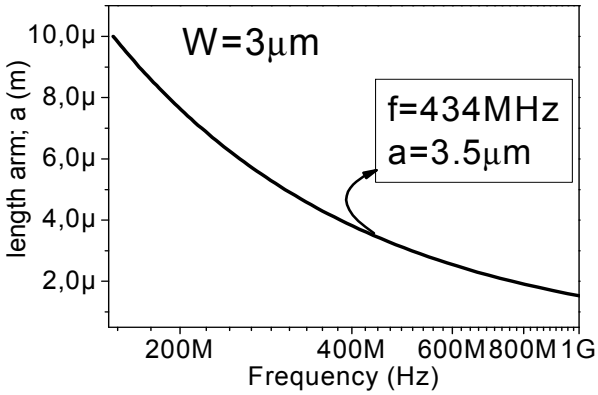


Figure 2.Length of semi resonator vs Frequency (*a*) for a fixed length of the coupling area, *W*

while the two drivers are fabricated using the polysilicon 2 layer. The layout is drawn with a 0 gap between drivers and resonator in order to take advantage of the fabrication process, using the interpoly oxide as spacer [9]. Therefore, it is expected to achieve a gap of 40 nm, corresponding to the oxide between the two polysilicon layers.

To enable the release of the resonator a pad window is opened over the structure, in order to remove the pasivation layer. Vias 3 and 2 are added to eliminate the different inter

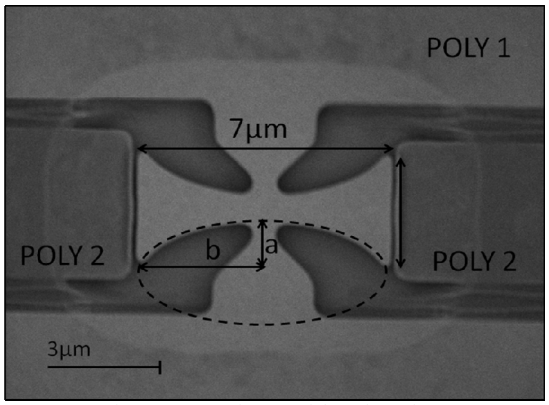


Figure 3. SEM image of the released resonator

metal oxides that would cover the MEMS. A mask less wet etching is performed using an HF based solution to release the resonator. Figure 3 shows a SEM image of the released resonator.

IV. ELECTRICAL CHARACTERIZATION

Two versions of the resonator have been implemented. The first one corresponds to the stand alone resonator while the second implementation adds a transimpedance amplifier in charge of amplifying the signal provided by the MEMS. This second implementation will allows discussing the effect of the presence of an amplifier on the characterization of the MEMS. Fig.7 shows an optical image of the MEM resonator, the circuitry and pads.

In both cases the measurement set up is shown in fig. 4. A biasing voltage, which allows to read the current provided by the movement, is applied over the resonator as well as the excitation ac signal. A Bias-tee has been included. To enlarge the output signal, the two electrodes are joined together to read the motional current. The output is applied directly to the Network analyzer or to the transimpedance amplifier. Scattering parameter S21 is used to characterize the LBAR.

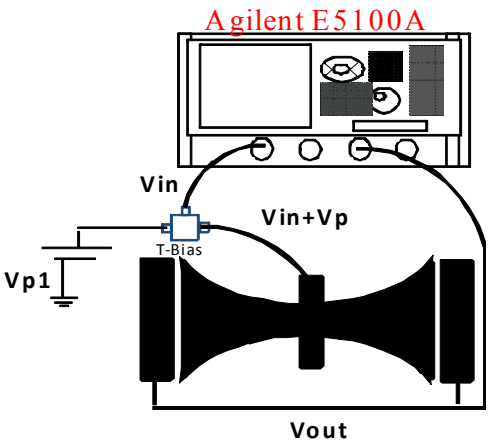


Figure 4. Electrical characterization set up, for stand alone resonator.

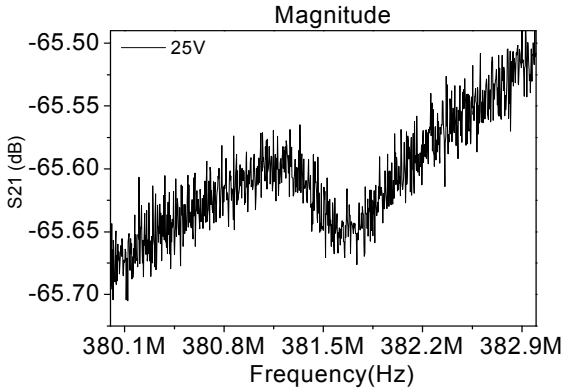


Figure 5. Transmission coefficient S21 for the stand alone resonator in air conditions.

Figure 5 shows the test results corresponding to the stand alone resonator. The resonance frequency is shown at 384 MHz with a level of around 65.6 dB. The small value of the peak is attributed to the presence of the parasitic capacitance between the movable structure and the drivers that mask the motional current.

The measurements of the fully integrated (see fig.6) prototype are shown in the figure 7. As it is expected, the coupling level is increased up to 39.2, due the 30 dB gain provided by the amplifier. In this case the resonance frequency is located at 387 Mhz due to the tolerances in the fabrication process. Contrary to other cases, it can be observed that there is not benefit of using the amplifier in the reduction of the parasitic current. This fact is attributed to the one port set-up measurement. Although the amplifier allows to eliminate the capacitance between the excitation and read-out pads, that exist in the stand-alone resonator, the dominant capacitance, corresponding to the physical capacitance between the movable polysilicon 1 layer and the electrodes polysilicon 2, is still present.

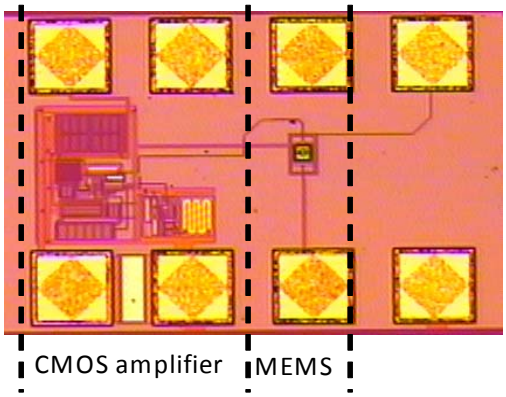


Figure 6. Optical image of the fully integrated prototype. The output of the mechanical device is connected to the input of the transimepedance amplifier

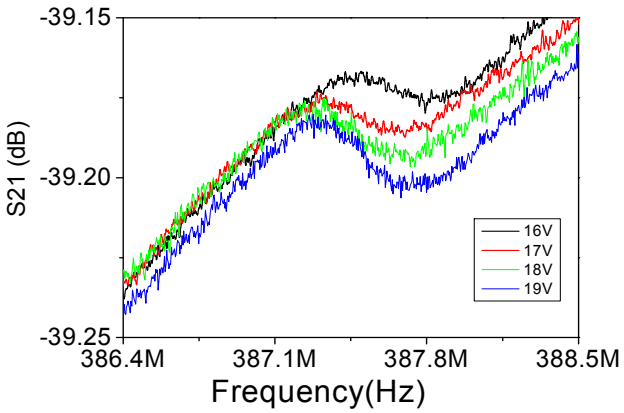


Figure 7. Transmission coefficient S21 for fully integrated resonator and CMOS amplifier.

V. Conclusions

We have presented a novel LBAR resonator that allows in a very simple manner to design a high frequency resonator, making independent the coupling area and resonance frequency.

To demonstrate the viability of the structure, a resonant micro-electromechanical LBAR working at 380 MHz with, and without CMOS on-chip readout circuitry has been implemented and tested.

The test results confirm the functionality of the LBAR, although a reduction of the parasitic capacitance is desired to enhance the resonance peak. The differential integrated fabrication with a dummy LBAR could solve the problem of the parasitic currents as were demonstrated in [3].

Acknowledgments

This work has been totally founded by the Spanish Government through the project Nemesys (TEC2009-09008/MIC)

REFERENCES

[1] J. Verd, A. Uranga, G. Abadal, J. Teva, F. Torres, F. Perez-Murano, J. Fraxedas, J. Esteve, and N. Barniol, "Monolithic mass sensor fabricated using a conventional technology with attogram resolution in air conditions," *Applied Physics Letters,* vol. 91, pp. 013501-013501-3, 2007.

[2] J. Verd, A. Uranga, G. Abadal, J. L. Teva, F. Torres, J. L. Lopez, F. Perez-Murano, J. Esteve, and N. Barniol, "Monolithic CMOS MEMS Oscillator Circuit for Sensing in the Attogram Range," *Electron Device Letters, IEEE,* vol. 29, pp. 146-148, año

[3] J. L. Lopez, J. Verd, A. Uranga, J. Giner, G. Murillo, F. Torres, G. Abadal, and N. Barniol, "A CMOS-MEMS RF-Tunable Bandpass Filter Based on Two High- Q 22-MHz Polysilicon Clamped-Clamped Beam Resonators," *Electron Device Letters, IEEE,* vol. 30, pp. 718-720, 2009.

[4] L. Chiung-Cheng, C. Fang, and G. K. Fedder, "Integrated HF CMOS-MEMS square-frame resonators with on-chip electronics and electrothermal narrow gap mechanism," in *Solid-State Sensors, Actuators and Microsystems, 2005. Digest of Technical Papers. TRANSDUCERS '05. The 13th International Conference on*, 2005, pp. 2074-2077 Vol. 2

[5] C. Wen-Chien, F. Weileun, and L. Sheng-Shian, "A generalized foundry CMOS platform for capacitively-transduced resonators monolithically integrated with circuits" *Journal of Micromechanics and Microengineergin, N.6, Vol. 21, 2011*.

[6] J.L.Lopez, J.Verd, J.Teva, G.Murillo, J.Giner, F.Torres, A.Uranga, G. Abadal, and N.Barniol, "Integration of RF-MEMS Resonators on Submicrometric Commercial CMOS Technologies," *Journal of Micromechanical and Microengineering,* vol. 19, p. 10, 2009.

[7] X. Yuan, L. Sheng-Shian, L. Yu-Wei, R. Zeying, and C. T. C. Nguyen, "UHF micromechanical extensional wine-glass mode ring resonators," in *Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International*, 2003, pp. 39.2.1-39.2.4

[8] L. Sheng-Shian, L. Yu-Wei, X. Yuan, R. Zeying, and C. T. C. Nguyen, "Micromechanical "hollow-disk" ring resonators," in *Micro Electro Mechanical Systems, 2004. 17th IEEE International Conference on. (MEMS)*, 2004, pp. 821-824

[9] T. Mattila, J. Kiihamäki, T. Lamminmäki, O. Jaakkola, P. Rantakari, A. Oja, H. Seppä, H. Kattelus, and I. Tittonen, "A 12 MHz micromechanical bulk acoustic mode oscillator," *Sensors and Actuators A: Physical,* vol. 101, pp. 1-9, 2002.

[10] N. Lobotniu and E. Garcia, *Mechanics of Microelectromechanical Systems*: Kluwer Academic Publishers, 2005.

The Properties of M^{th} Order Differences and Their Relationship to M^{th} Order Random Stability

Victor S. Reinhardt, Raytheon Space and Airborne Systems, El Segundo, California, USA

Abstract—Allan and Hadamard variances are well-known examples of M^{th} order difference (Δ) variances. These are mean square (MS) averages over data of the M^{th} order Δ-measure $\Delta(\tau)^M v(t)$, where $\Delta(\tau)v(t) = v(t+\tau) - v(t)$. Many of the important properties of Allan and Hadamard variances are specific examples of the more general properties of M^{th} order Δ-measures. This paper first provides a useful compendium of these properties along with proofs. Allan and Hadamard variances are also known as MS measures of M^{th} order random stability (or instability). But what is M^{th} order random stability? The second part of this paper offers a definition of such stability as the prediction error *solely due to the random error or noise in the data* at an $(M+1)^{th}$ point from an M-parameter calibration function that is passed through M previously measured data points. Using this definition and the properties of Δ-measures, it is shown that M^{th} order Δ-variances are natural measures of M^{th} order MS random stability when the deterministic drift in the data has $(M-1)^{th}$ or lower order polynomial behavior, the calibration function is an $(M-1)^{th}$ polynomial, and the measured data and predicted value are all spaced by τ. When the drift (aging plus environmental variation) does not have this behavior, it is further shown that Δ-variances are biased measures of such polynomial random stability, even when one uses fitting techniques to remove the drift from the data. This drift removal process is shown to alter the spectral kernel $K_{stat}(f)$ that relates the Δ-variance in question to the power spectral density (PSD) of the random noise in the data, even if the fit perfectly removes the direct drift effects. Methods are presented for computing such altered $K_{stat}(f)$ and are illustrated by computing the changes in $K_{stat}(f)$ that occur due to various drift-removal techniques in the unmodified Allan variance. These results are then used to explain why biases in drift-removed Δ-variances can change so drastically as a function of power law PSD type and fit methodology. Finally, it is shown that M^{th} order Δ-measures that are natural measures of random stability in the presence of intrinsic aging can become highly biased as a result of environmental drift removal.

I. INTRODUCTION

Allan [1] and Hadamard [2] variances are well-known examples of M^{th} order difference variances or Δ-variances. These variances are proportional to the mean square (MS) of the M^{th} order forward difference or Δ-measure $\Delta(\tau)^M v(t)$ [3]. Here, v(t) is some data variable such as x(t) the time error or y(t) the fractional frequency error [1], and the operator $\Delta(\tau)^M$ is the M^{th} application of the 1^{st} order forward difference operator $\Delta(\tau)$ defined by $\Delta(\tau)v(t) = v(t+\tau) - v(t)$ [3]. Many of the important properties of Allan and Hadamard variances are specific examples of the more general properties of Δ-measures and their variances. Some of these properties have been well-known since the time of Newton [3][4], but others are less well-known [5]. This paper provides a useful compendium of these properties as they relate to error measures such as Allan and Hadamard variances along with their proofs.

Allan and Hadamard variances are also known as MS measures of M^{th} order random stability (or instability) [1][2]. But exactly what does one mean by such stability? The second part of this paper offers a definition of M^{th} order random stability as the prediction error *solely due to the random error or noise in the data* at an $(M+1)^{th}$ point from an M-parameter calibration function that is passed through M previously measured data points. Using this definition and the properties of Δ-measures, the paper investigates nature of biases in M^{th} order Δ-variances used as measures of polynomial M^{th} order random stability (defined as M^{th} order random stability when the calibration function is an $(M-1)^{th}$ order polynomial and all points are equally spaced by τ). Here, bias is used in the sense of the mean difference between the Δ-variance value and that of the above definition of random stability.

It is shown that M^{th} order Δ-variances are natural bias-free measures of such M^{th} order polynomial stability when the deterministic drift in the data (intrinsic aging in v(t) over time plus environmentally induced changes in v(t) over time) is an $(M-1)^{th}$ or lower or polynomial. When this is not the case, it is also shown that attempting to remove the drift using data fitting techniques will leave a residual bias that alters the spectral kernel $K_{stat}(f)$ that relates the Δ-variance in question to the power spectral density (PSD) of the random noise in the data, even if the fit perfectly removes the direct drift effects. Methods are presented for computing such altered $K_{stat}(f)$ and are illustrated by computing the changes in $K_{stat}(f)$ that occur due to various drift-removal techniques in the unmodified Allan variance. These results are then used to explain why biases in drift-removed Δ-variances can change so drastically as a function of power law PSD type [1][6] and fit methodology. Finally, it is shown that M^{th} order Δ-measures that are natural measures of random stability in the presence of intrinsic aging can become highly biased as a result of environmental drift removal.

A. Basic Terminology

Figure 1 and Table I describe the basic terminology and truth model to be used in this paper. v(t) in (1) is any

continuous data variable, such as x(t) the time error or y(t) the fractional frequency error [1]. Here, t is the ideal error free continuous observation time, and $v(t_n)$ is a set of data samples at the error free sampling times t_n over the data interval T as described in (2) and (3). In our truth model, v(t) is additively partitioned into random error or noise $v_r(t)$ (4) and drift $v_c(t)$ (5), which is the deterministic component of v(t). Here, changes in $v_c(t)$ over time are due both to intrinsic aging, which is a pure function of time, and independent environmental parameters $\upsilon(t)$ (6), which mediate the change in $v_c(t)$ over time. Note in (5) that $\upsilon(t)$ is subsumed into the t-variable in $v_c(t)$. This simply assumes that $\upsilon(t)$ does not change with ensemble averaging over a specific data interval T; that is, $\upsilon(t)$ is deterministic. In the paper, we will also assume that $v_r(t)$ is both zero mean (ZM) and wide-sense stationary (WSS), but the paper's results can be straightforwardly generalized to include noise that is neither ZM nor WSS [7]. Finally, we note that drift is in general neither ZM nor WSS.

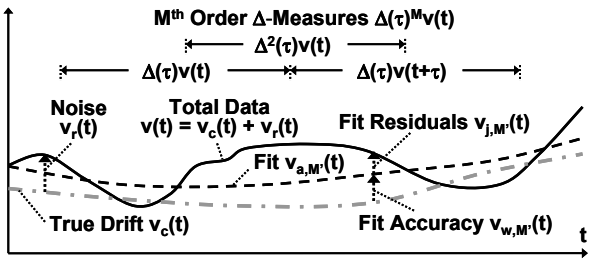


Figure 1. Basic Terminology

TABLE I. BASIC TERMINOLOGY

Total continuous data variable: $v(t) = v_c(t) + v_r(t)$	(1)
Sampled data: $t = t_n$ [Unless otherwise specified sampling is uniform]	(2)
Uniform sampling: $t_n = t_0 + nt_s$ [n= 0 to N-1 and $t_s = T/N$]	(3)
Random error or noise: $v_r(t)$	(4)
True deterministic drift: $v_c(t) = v_c(t,\upsilon(t))$	(5)
Environmental variables included in $v_c(t)$: $\upsilon(t)$	(6)
Fitted estimate of $v_c(t)$ from the data set $v(t_n)$: $v_{a,M'}(t,v(t_n))$	(7)
M'-parameter model function (before fit) [1]: $v_{a,M'}(t,\mathbf{a})$ [$\mathbf{a} = (a_0 \dots a_{M'-1})^\dagger$]	(8)
True fit accuracy: $v_{w,M'}(t_{n'},v(t_n)) = v_{a,M'}(t_{n'},v(t_n)) - v_c(t_{n'})$	(9)
Fit residuals or data precision: $v_{j,M'}(t_{n'},v(t_n)) = v(t_{n'}) - v_{a,M'}(t_{n'},v(t_n))$	(10)
Theoretical mean square (MS) errors [2]: $\sigma_{\varsigma,M'}(v(t_n))^2 = \mathcal{E}\Sigma_{n=0}^{N-1} v_{\varsigma,M'}(t_n)^2$	(11)
M^{th} order Δ-measure of $v(t_n)$: $\Delta(\tau)^M v(t_n)$ where $\Delta(\tau)v(t_n) = v(t_n+\tau) - v(t_n)$	(12)
$\tau = m_\tau t_s$ is assumed, where m_τ is an integer.	(13)
Theoretical M^{th} order Δ-variance: $\sigma_{\Delta,M}(\tau,v(t_n))^2 = \lambda_M^{-1}\mathcal{E}(\Delta(\tau)^M v(t_n))^2$	(14)
$\lambda_M = \Sigma_{m=0}^{M} c(M,m)^2 = c(2M,M)$ where $c(m,n) = m!/(n!(m-n)!)$	(15)

[1] $(\dots)^\dagger$ = hermitian transpose (transpose for real matrices).

[2] ς = c or r and $\mathcal{E}$ = ensemble averaging operator.

In our truth model, some as yet unspecified fitting technique has generated a data fit $v_{a,M'}(t,v(t_n))$ (7) based on the data set $v(t_n)$ and an adjustable M' **a**-parameter model function $v_{a,M'}(t,\mathbf{a})$ (8). Note that the explicit reference to $v(t_n)$ in $v_{a,M'}(t,v(t_n))$ is important to distinguish it from $v_{a,M'}(t,v_r(t_n))$ and $v_{a,M'}(t,v_c(t_n))$, which are fits based on $v_r(t_n)$ or $v_c(t_n)$ alone as the data set. Two basic error measures associated with the fitting process are the true fit accuracy $v_{w,M'}(t_{n'},v(t_n))$ defined in (9) and the fit error residuals or data precision $v_{j,M'}(t_{n'},v(t_n))$ defined in (10). $\sigma_{\varsigma,M'}(v(t_n))^2$ (ς = w or r) defined in (11) is the ensemble ($\mathcal{E}$) averaged (theoretical) mean square (MS) of $v_{\varsigma,M'}(t_{n'},v(t_n))$ at the point $t_{n'}$. For further detail on this model, see [7].

(12) and (14) are the definitions of $\Delta(\tau)^M v(t_n)$ the M^{th} order Δ-measure of $v(t_n)$ and $\sigma_{\Delta,M}(\tau,v(t_n))^2$ its $\mathcal{E}$-averaged variance to be called the theoretical Δ-variance. Here, the normalization factor λ_M in (15) is chosen so that $\sigma_{\Delta,M}(\tau,v_r(t_n))^2$ equals $\mathcal{E}v_r(t_n)^2$ for $v_r(t_n)$ samples that are uncorrelated with each other [5]. Note that $t_n + M\tau$ must be within the data set for the definition to apply. Further note that $\Delta(\tau)^M v(t)$ is not in general ZM or WSS, since $\Delta(\tau)^M v_c(t)$ is not in general ZM or WSS. Thus, $\sigma_{\Delta,M}(\tau,v(t))^2$ is not in general a true variance in the sense that the ensemble mean has been removed, but we will call it so for simplicity. On the other hand, $\Delta(\tau)^M v_r(t)$ is assumed to be ZM and WSS, and thus $\sigma_{\Delta,M}(\tau,v_r(t))^2$ is a true variance. Finally, we note that $\sigma_{\Delta,2}(\tau,y(t))^2 = \sigma_y(\tau)^2$ the theoretical unmodified Allan variance of y(t) [1], $\sigma_{\Delta,3}(\tau,x(t))^2 = \sigma_x(\tau)^2 = (\tau^2/3)\sigma_y(\tau)^2$ the theoretical unmodified Allan variance of x(t) [1], and $\sigma_{\Delta,2}(\tau,y(t))^2 = {}_H\sigma_y(\tau)^2$ the theoretical unmodified Hadamard variance of y(t) [2].

II. A COMPENDIUM OF M^{TH} ORDER Δ-MEASURE PROPERTIES

This section contains a compendium of the properties of M^{th} order Δ-measures or M^{th} order forward finite differences $\Delta^M(\tau)v(t)$ and their Δ-variances as they relate to stability and fitting error measures. Explicit proofs are provided for properties that do not follow directly from previous ones to make the compendium more complete. Finally, note there are also backward, mean, and divided differences not discussed here [3,8].

The theory of finite differences, which is also called the calculus of finite differences, has a very long history [3,4]. References begin about the time of Newton [4], continuing with an exhaustive 1860 treatise by Boole [3], and on through the 20th century [8]. Finite differences were also utilized in the earliest computing engines, such as the Babbage engine [9], to turn multiplications into additions in the generation of polynomials. The historical material emphasizes such functional generation and the numerical solution of difference equations. Thus, despite the long history of the field, there is still material relating to stability and fitting error that seems to be only recently added [5].

Table II lists basic terminology and properties of M^{th} order Δ-measures [3][5][8]. In the table, secondary properties straightforwardly follow from definitions, so explicit proofs will not be provided. Note that the Fourier properties in the table are important for deriving the spectral properties of Δ-variances. These Fourier properties are generated here by equating time-domain operators (i.e., $\Delta(\tau)^M$) to time-domain impulse response functions (i.e., $h_{\Delta,M}(t)$) and then Fourier transforming these impulse response functions. Finally, two properties of note not in the table are that both $\Theta(\tau)$ and $\Delta(\tau)$ can be represented in terms of the differentiation operator $\mathcal{D}_t = d/dt$ as $\Theta(\tau) = \exp(\tau\mathcal{D}_t)$ and $\Delta(\tau) = \exp(\tau\mathcal{D}_t) - 1$ when τ is not function of t [2]. Fourier transforming $\mathcal{D}_t$ is an alternative method for generating the table's Fourier properties.

TABLE II. BASIC PROPERTIES OF M^TH ORDER Δ-MEASURES

(Forward) τ-displacement operator [1]:

$\Theta(\tau)v(t) = v(t+\tau) = \int dt' h_\Theta(t-t')v(t') = h_\Theta(t-t') \otimes v(t')$ (16)

Equivalent t-domain impulse response: $h_\Theta(t) = \delta(t+\tau)$ (17)

f-domain response function [2]: $H_\Theta(f) = \mathcal{F}_{f,t} h_\Theta(t) = \exp(j\omega\tau)$ (18)

$V_\Theta(f) = \mathcal{F}_{f,t} v_\Theta(t) = \mathcal{F}_{f,t} \Theta(\tau)v(t) = H_\Theta(f)V(f)$ where $V(f) = \mathcal{F}_{f,t} v(t)$ (19)

M^th order displacement: $\Theta(\tau)^m = \Theta(m\tau)$ (20)

Equivalent t-domain impulse response: $h_{\Theta,m}(t) = \delta(t+m\tau)$ (21)

f-domain response function: $H_{\Theta,m}(f) = \mathcal{F}_{f,t} h_{\Theta,m}(t) = \exp(jm\omega\tau)$ (22)

M^th order (forward) difference operator: $\Delta(\tau)^M = (\Theta(\tau)-\Theta(0))^M$ (23)

1st order Δ-measure: $\Delta(\tau)v(t) = v(t+\tau) - v(t) = (\Theta(\tau)-\Theta(0))v(t)$ (24)

Binomial expansion: $\Delta(\tau)^M = \Sigma_{m=0}^M (-1)^{M-m} c(M,m) \Theta(\tau)^m$ (25)

Falling Factorial: $z^{(m)} = z(z-1)\ldots(z-m+1)$ [Note $m^{(m)} = m!$] (26)

$\Delta(\tau)^M t^{M'} = M'^{(M)} t^{(M'-M)}$ if $M \le M'$ otherwise $\Delta(\tau)^M t^{M'} = 0$ (27)

Equivalent t-domain impulse response:

$h_{\Delta,M}(t) = \Sigma_{m=0}^M (-1)^m c(M,m) \delta(t+m\tau)$ (28)

f-domain response function: $H_{\Delta,M}(f) = \Sigma_{m=0}^M (-1)^{M-m} c(M,m) \exp(j\omega\tau)^m$

$H_{\Delta,M}(f) = (\exp(j\omega\tau) -1)^M = (2j)^M \exp(j\pi f\tau)\sin(\pi f\tau)^M$ (29)

[1] Integration limits are $-\infty$ to $+\infty$ unless otherwise stated. Finite limiting factors are assumed to be included in the integrand. $\otimes$ = convolution.

[2] Fourier Tansform: $\mathcal{F}_{f,t} v(t) = \int df \exp(-j\omega t) v(t)$ where $\omega = 2\pi f$.

TABLE III. FURTHER PROPERTIES OF M^TH ORDER Δ-MEASURES

M^th order drift insensitivity: $\Delta(\tau)^M v_{ply,M'}(t,\mathbf{a}) = 0$ if $M' \le M$ (30)

$(M-1)^{th}$ order polynomial: $v_{ply,M}(t,\mathbf{a}) = \Sigma_{m=0}^{M-1} a_m (t-t_0)^m$ (31)

Prediction error [1] at $t_{m'}$ (M+1 total data points)

$v_{e,M,M'}(t_{m'},v(t_m)) = v(t_{m'}) - v_{cal,M'}(t_{m'},v(t_m))$ (32)

where $v_{cal,M'}(t_{m'},v(t_m))$ is fit to M measured points $v(t_m)$ $m \ne m'$ ($t_{m'}$ not in fit)

$\sigma_{e,M,M'}(t_{m'},v(t_m))^2 = \mathcal{E} v_{e,M,M'}(t_{m'},v(t_m))^2$ (33)

Theorem: $v_{e,M,M}(t_{m'},v(t_m)) = \Delta(\tau)^M v(t_0) (-1)^{M-m'}/c(M,m')$ (34)

when $v_{cal,M}(t_{m'},v(t_m)) = v_{poly,M}(t_{m'},v(t_m))$ that passes through M measured $v(t_m)$, all points are spaced by τ, and $t_{m'}$ is not included in the measured t_m.

$\sigma_{e,M,M}(v(t_{m'}))^2 = \lambda_M \sigma_{\Delta,M}(\tau,v(t))^2/c(M,m')^2$ (35)

Theorem: $v_{j,M}(t_m) = \Delta(T/M)^M v(t_0)(-1)^{M-m} c(M,m)/\lambda_M$ (36)

when $v_{a,M}(t_{m'},v(t_m)) = v_{poly,M}(t_{m'},v(t_m))$ is generated with a USLF over M+1 measured points spaced by τ and $t_{m'}$ is included in the measured t_m.

$\sigma_{j,M}(v)^2 = \sigma_{\Delta,M}(\tau,v(t_0))^2$ (37)

WSS spectral representation² of σ_{stat}^2: $\sigma_{stat}^2 = \int df K_{stat}(f)|H_s(f)|^2 L_v(f)$ (38)

$K_{stat}(f) = |H_{\Delta,M}(f)|^2$ for $\sigma_{stat}^2 = \sigma_{\Delta,M}(\tau,v_r(t_0))^2$ (39)

[1] $v_{e,M,M'}(t_{m'},v(t_m))$ and $v_{j,M'}(t_{m'},v(t_m))$ have similar formulas but not the same meaning.

[2] $H_s(f)$ represents system filtering that is included in the t-domain variables $v(t)$, $v_r(t)$, and $v_e(t)$ [5][7]. Thus, the single sideband power spectral density (PSD) of $v_r(t)$ is $|H_s(f)|^2 L_v(f)$, where $L_v(f)$ is the pre-filtered PSD of the pre-filtered noise process, i.e., for neg-p noise this PSD is $L_v(f) = K_p|f|^p$ [5][6][7].

Table III lists further properties of M^th order Δ-measures. The first property (30) is the well-known insensitivity of M^th order Δ-measures to $(M-1)^{th}$ or lower order polynomial drift. This follows directly from (27). The next property (34) relates Δ-measures to polynomial prediction error after an ideal M-point calibration. General prediction error $v_{e,M,M'}(t_{m'},v(t_m))$ is defined in (32) as the difference at an *unmeasured* point $t_{m'}$ between the true value $v(t_{m'})$ and a prediction based on $v_{cal,M'}(t_{m'},v(t_m))$ an M'-parameter *calibration* function somehow generated from M *measured* calibration points $v(t_m)$. Note that $v_{cal,M'}(t_{m'},v(t_m))$ is essentially the same as the fit $v_{a,M'}(t_{n'},v(t_n))$, except that the points and functions may be different. As shown in Figure 2, when $t_{m'}$ is the last point t_M, the prediction error is an extrapolation error, and when $t_{m'}$ is an internal point, the prediction error is an interpolation error. Finally, an ideal M-point calibration $v_{cal,M}(t,v(t_m))$ is defined as an M-parameter calibration function that exactly passes through all M measured points. Note that this is always possible here for a well-formed function, because there are zero degrees of freedom (0-DOF) in the fit.

Theorem (34) states that the prediction error at an unmeasured $(M+1)^{th}$ point $t_{m'}$ using an ideal $(M-1)^{th}$ order polynomial calibration is proportional to $\Delta(\tau)^M v(t_0)$ when all points are spaced by τ. The proof of (34) is as follows [5]. First from (25), note that

$$\Delta(\tau)^M v(t_0) = \Sigma_{m \ne m'}^M (-1)^{M-m} c(M,m) \Theta(\tau)^m v(t_0) + (-1)^{M-m'} c(M,m') v(t_{m'}) \quad (40)$$

Second, note that $v_{ply,M}(t_m,\mathbf{a})$ can always be made to pass through M data points, because it is a 0-DOF fit. Thus, one can substitute $v_{ply,M}(t_{m'},v(t_m))$, the $(M-1)^{th}$ order polynomial that passes though the measured $v(t_m)$, for $v(t_m)$ in the right term of (40) to obtain

$$\Delta(\tau)^M v(t_0) = \Sigma_{m \ne m'}^M (-1)^{M-m} c(M,m) \Theta(\tau)^m v_{ply,M}(t_0,v(t_m)) + (-1)^{M-m'} c(M,m') v(t_{m'}) \quad (41)$$

Reversing the binomial expansion, we can then write this as

$$\Delta(\tau)^M v(t_0) = \Delta(\tau)^M v_{ply,M}(t_0,v(t_m)) + (-1)^{M-m'} c(M,m')(v(t_{m'}) - v_{ply,M}(t_{m'})) \quad (42)$$

Finally from (30), we note that $\Delta(\tau)^M v_{ply,M}(t_0,v(t_m)) = 0$, which proves (34). For extrapolation error, also note that (34) has the simple form $v_{e,M,M}(t_M,v(t_m)) = \Delta(\tau)^M v(t_0)$.

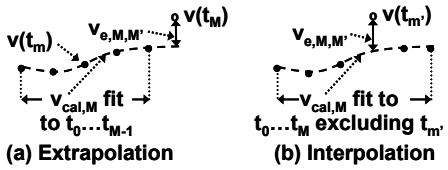


Figure 2. Extrapolation and interpolation prediction error.

(36) was first demonstrated by the author using combination of analytical and Monti Carlo techniques [5], but Charles Greenhall later provided the author with a totally analytical proof in an unpublished communication [10]. We reproduce this proof here as follows with some modifications.

TABLE IV. MATRIX EQUATIONS FOR M+1 DATA POINTS

$\mathbf{v} = (v(t_0),\ldots,v(t_M))^\dagger$ (43)

$\mathbf{v}_j = (v_j(t_0),\ldots,v_j(t_M))^\dagger$ (44)

$\mathbf{v}_{ply} = (v_a(t_0),\ldots,v_a(t_M))^{(T)} = \mathbf{ua}$ (45)

$\mathbf{a} = (a_0,\ldots,a_{M-1})^\dagger$ (46)

$\mathbf{u}_{nm} = (t_n - t_0')^m$ [n = 0 to M, m = 0 to M-1, t_0' arbitrary] (47)

Un-weighted least squares fit (ULSF) solution:

$\mathbf{a} = (\mathbf{u}^\dagger\mathbf{u})\mathbf{u}^\dagger\mathbf{v}$ $\mathbf{v}_{ply} = \mathbf{g}_w\mathbf{v}$ $\mathbf{v}_j = (\mathbf{I} - \mathbf{g}_w)\mathbf{v}$ $\mathbf{g}_w = \mathbf{u}(\mathbf{u}^\dagger\mathbf{u})\mathbf{u}^\dagger$ (48)

$\Delta(T/M)^M v(t_0) = \boldsymbol{\delta}^\dagger\mathbf{v}$ [from (25)] (49)

$\boldsymbol{\delta} = ((-1)^M c(M,0),\ldots,(-1)^{M-n} c(M,n),\ldots c(M,M))^\dagger$ (50)

$(\boldsymbol{\delta}^\dagger\mathbf{u})_m = \Delta(T/M)^M t_0^{m-1} = 0$ [from (27)] (51)

$\mathbf{v}_j = \mathbf{g}_\delta\mathbf{v}$ [(36) in matrix notation] $\mathbf{g}_\delta = \boldsymbol{\delta}(\boldsymbol{\delta}^\dagger\boldsymbol{\delta})^{-1}\boldsymbol{\delta}^\dagger = \lambda_M^{-1}\boldsymbol{\delta}\boldsymbol{\delta}^\dagger$ (52)

$\lambda_M = \boldsymbol{\delta}^\dagger\boldsymbol{\delta}$ [from (15) and (50)] (53)

Table IV defines the matrix notation used to prove (36). In the table, we write the M+1 the values of $v(t_m)$, $v_j(t_m)$, and

$v_{ply,M'}(t_{m'},v(t_m))$ as M+1 dimensional vectors $\mathbf{v}$, $\mathbf{v}_j$, and $\mathbf{v}_{ply}$ (43), (44), and (45). In (45), $v_{ply,M'}(t_{m'},v(t_m))$ is defined using both the M-dimensional parameter vector $\mathbf{a}$ (46) and a power law basis function matrix $\mathbf{u}$ (47). In this notation, an un-weighted least squares fit (ULSF) solution is given by (48) [11]. In this form, it is well known that $\mathbf{g}_w$ the solution matrix for $\mathbf{v}_{ply}$ in terms $\mathbf{v}$ is an operator that projects $\mathbf{v}$ into an M-dimensional manifold defined by the set of basis vectors in $\mathbf{u}$ [11]; that is, $\mathbf{g}_w\mathbf{v}_{ply} = \mathbf{v}_{ply}$ for any $\mathbf{a}$. It is also well-known that $\mathbf{v}_j = (\mathbf{I} - \mathbf{g}_w)\mathbf{v}$, and $(\mathbf{I} - \mathbf{g}_w)$ is an operator that projects $\mathbf{v}$ into the normal to this $\mathbf{u}$-defined manifold (the orthogonality theorem) [11]. Since, there are only M+1 dimensions in the total space, $\mathbf{v}_j$ can thus be represented by a constant b times a single vector orthogonal to all the vectors in $\mathbf{u}$.

To find such a vector, consider $\Delta(T/M)^M v(t_0)$ given in (49) as $\boldsymbol{\delta}^\dagger\mathbf{v}$, where $\boldsymbol{\delta}$ is defined in (50). From (27), we note in (51) that $\boldsymbol{\delta}^\dagger\mathbf{u} = \mathbf{0}$. Thus, $\boldsymbol{\delta}$ is a vector orthogonal to the basis $\mathbf{u}$, and we can write $\mathbf{v}_j$ as $b\boldsymbol{\delta}$. We further note that $\mathbf{g}_\delta$ in (52) is the projection operator into the direction of $\boldsymbol{\delta}$; that is, $\mathbf{g}_\delta\mathbf{v}_j = b\boldsymbol{\delta}(\boldsymbol{\delta}^\dagger\boldsymbol{\delta})^{-1}\boldsymbol{\delta}^\dagger\boldsymbol{\delta} = b\boldsymbol{\delta} = \mathbf{v}_j$. Finally, noting $\lambda_M = \boldsymbol{\delta}^\dagger\boldsymbol{\delta}$ in (53), we prove (36).

The final property of M^{th} Order Δ-Measures in Table III relates to the wide-sense stationary (WSS) spectral integral representation of σ_{stat}^2 [5][7], where σ_{stat}^2 is some MS statistic of $\Delta(\tau)^M v_r(t)$. This spectral representation is given here and in Table III (38) as

$$\sigma_{stat}^2 = \int df\, K_{stat}(f)|H_s(f)|^2 L_v(f) \tag{38}$$

where

(a) $L_v(f)$ is the pre-system-filtered single sideband power spectral density (PSD) of $v_r(t)$ [5][7] (see Table III Note 2),

(b) $H_s(f)$ is a system response function [5][7], which is a generalization of a high frequency spectral cut-off f_h [1], and

(c) $K_{stat}(f)$ is a spectral kernel that defines the spectral properties of the statistic independent of $H_s(f)$ [5][7].

For $\sigma_{stat}^2 = \sigma_{\Delta,M}(\tau,v_r(t_0))^2$, (39) notes that $K_{stat}(f) = |H_{\Delta,M}(f)|^2$ [5]. This $|H_{\Delta,M}(f)|^2$ has the well-known f^{2M} highpass filtering property of an M^{th} order Δ-variance with respect to $L_v(f)$ [1][5][7]. We finally note that $K_{stat}(f)$ for $\sigma_{\Delta,M}(\tau,v_r(t_0))^2$ is a response function squared, because $\Delta(\tau)^M$ is a time translation invariant operator. We finally note that $K_{stat}(f)$ cannot be represented as a response function squared for all error measures or statistics [5][7].

III. M^{TH} Order Random Stability (Instability)

Intuitively, random stability (or instability) is a measure of the temporal variation of v(t) solely due to the random data component $v_r(t)$. In this section, we precisely define the concept of M^{th} order random stability and explore the issues associated with implementing this definition when $v_c(t)$ as well as $v_r(t)$ is present in data.

Table V (54) defines general M^{th} order random stability $v_{s,M}(t_{m'},v_r(t_m))$ as the prediction error $v_{e,M,M}(t_{m'},v_r(t_m))$ when

(a) the data consists solely of random error $v_r(t_n)$, and

(b) the M-parameter calibration function $v_{cal,M}(t,v_r(t_m))$ is fit to pass through M measured calibration points $v_r(t_m)$.

Similarly, Table V (56) is the equation for polynomial M^{th} order random stability, which is defined as M^{th} order random stability when

(a) the model function $v_{cal,M}(t,\mathbf{a})$ is $v_{ply,M}(t,\mathbf{a})$, and

(b) all points are equally spaced by τ.

Note here that (34) is also used to write polynomial random stability in terms of an M^{th} order Δ-measure. Their MS forms are given in (55) and (57) (using (35)). Furthermore, applying (30) to (56) and (57), we note in (58) that $\Delta(\tau)^M v(t_n)$ and $\sigma_{\Delta,M}(\tau,v(t))^2$ the Δ-measure and Δ-variance acting on the raw (non-fit removed) data are unbiased measures of M^{th} order polynomial random stability when the drift is of the form $v_{ply,M'}(t,\mathbf{a})$ for $M' \le M$ over the data interval T (because $\Delta(\tau)^M v_c(t) = 0$). This is a generalization of the well-known fact that Hadamard variances are insensitive to constant frequency drift [2].

TABLE V. M^{TH} ORDER RANDOM INSTABILITY

General M^{th} order random stability at uncalibrated point $t_{m'}$	
$v_{s,M}(t_{m'},v_r(t_m)) = v_{e,M,M}(t_{m'},v_r(t_m))$	(54)
$\sigma_{s,M}(t_{m'},v_r(t_m))^2 = \mathcal{E}v_{s,M}(t_{m'},v_r(t_m))^2$	(55)
Polynomial M^{th} order random stability, $v_{cal,M}(t,\mathbf{a}) = v_{ply,M}(t,\mathbf{a})$, $t_m = t_0+m\tau$	
$v_{s,M}(\tau,t_{m'},v_r(t_m)) = \Delta(\tau)^M v_r(t_0)\,(-1)^{M-m'}/c(M,m')$	(56)
$\sigma_{s,M}(\tau,t_{m'},v_r(t_m))^2 = \sigma_{\Delta,M}(\tau,v_r(t_0))^2\,\lambda_M/c(M,m')^2$	(57)
When $v_c(t) = v_{ply,M'}(t,a)$ for $M' \le M$	
$\Delta(\tau)^M v(t_0)$ and $\sigma_{\Delta,M}(\tau,v(t_0))^2$ (acting on data with no fit removal) are natural unbiased measures $v_{s,M}(\tau,t_{m'},v_r(t_m))$ and $\sigma_{s,M}(\tau,t_{m'},v_r(t_m))^2$.	(58)
When $v_c(t) \ne v_{ply,M'}(t,a)$ for $M' \le M$ & linear fit over N points t_n	
$\Delta(\tau)^M v_{j,M'}(t_{n'},v(t_n)) = \Delta(\tau)^M v_r(t_{n'}) - v_{c\text{-}bias}(t_{n'}) - v_{r\text{-}bias}(t_{n'})$	(59)
$v_{c\text{-}bias}(t) = \Delta(\tau)^M v_{w,M'}(t,v_c(t_n))$ = bias due to $v_{a,M'}(t,v_c(t_n)) \ne v_c(t)$ over T	
$v_{r\text{-}bias}(t) = \Delta(\tau)^M v_{ply,M'}(t,v_r(t_n))$ = bias that changes kernel $K_{stat}(f)$	
Note $v_{ply,M'}(t_{n'},v(t_n)) = v_{ply,M'}(t_{n'},v_r(t_n)) + v_{ply,M'}(t_{n'},v_c(t_n))$	(60)

When $v_c(t)$ is not equal to $v_{ply,M'}(t,\mathbf{a})$ for $M' \le M$ over T, $\Delta(\tau)^M v_c(t_n)$ is not zero, and thus $\Delta(\tau)^M v(t_n)$ and its various MS statistics $\sigma_{stat}(v(t_n))^2$ are no longer unbiased measures of M^{th} order polynomial random stability. This is a generalization of the well-known fact that Allan variances acting on raw data are sensitive to frequency drift in the data and are thus biased measures of purely random stability. When utilizing Δ-variances as measures of random stability in the presence of such drift, standard practice is to minimize the bias by attempting to remove the drift with some fit $v_{a,M'}(t_{n'},v(t_n))$ and then utilizing $\Delta(\tau)^M v_{j,M'}(t_{n'},v(t_n))$ to generate the Δ-variance [1]. However, it is recognized that this is an imperfect process, and drift removal methodologies have been developed to minimize the bias for constant frequency drift [12].

In such imperfect cases, Table V (59) separates the bias or difference between $\Delta(\tau)^M v_{j,M'}(t_{n'},v(t_n))$ and $\Delta(\tau)^M v_r(t_{n'})$ the true measure of polynomial random stability into two terms as follows. $v_{c\text{-}bias}(t)$ in (59) is the direct bias component that occurs when the fitted model function $v_{a,M'}(t,v_c(t_n))$ does not completely track $v_c(t)$ over T. The form shown in (59) in terms of $v_{w,M'}(t,v_c(t_n))$ is obtained by using (9), (10), and (60). In this case, we also note that (38) no longer completely

specifies σ_{stat}^2; that is, one must add a second spectral integral term that is a function of the PSD of $v_c(t)$ (or its non-WSS equivalent) to completely specify σ_{stat}^2 [5][7]. The second bias component in (59) $v_{r\text{-}bias}(t)$ is the bias effect of the drift removal fitting process on $K_{stat}(f)$. It is noted that the fitting process changes $K_{stat}(f)$ from its non-fit-removed form even when the drift removal makes $v_{c\text{-}bias}(t) = 0$, because the fitting process filters $v_r(t)$ as well as $v_c(t)$ [5][7].

To analytically compute this fit-removed $K_{stat}(f)$, one can use the spectral theory developed in [5] and [7]. In addition, there is a simple numerical method for generating $K_{stat}(f)$ for any σ_{stat}^2 of the form $\Sigma_{n'}\xi_{n'}v_{stat}(t_{n'},v(t_n))^2$ as follows. One first generates code to compute $v_{stat}(t_{n'},v(t_n))$ starting with two input data sets $v(t_n) = \cos(2\pi f_o t_n)$ and $v(t_n) = \sin(2\pi f_o t_n)$ over the desired range of f_o. One then computes $\sigma_{stat}^2(f_o)$ for each f_o from these $v(t_n)$ and adds the resultant $\sigma_{stat}^2(f_o)$ values together to produce a net $\sigma_{stat}^2(f_o)$. We note for this input that $\sigma_{stat}^2(f_o) = K_{stat}(f_0)$, because

(a) $v(t_n)$ includes the system filtering, so $\sigma_{stat}(f_o)^2 = \int df\, K_{stat}(f)L_v(f)$,

(b) $L_v(f)$ in this case consists of Dirac δ-functions at $\pm f_o$,

(c) $K_{stat}(f)$ is symmetric about f, and

(d) summing $\sigma_{stat}(f_o)^2$ for cosine and sine inputs is equivalent to randomizing the phase of the input to make the output result WSS.

In using this numerical technique, a word of caution is warranted. Numerical truncation error can generate misleading results at small f when a high degree of cancellation from many terms is responsible for the result. Thus, one should always check for proper theoretical asymptotic behavior before relying on the numerical result.

As an illustration, let us now consider the $K_{stat}(f)$-biases introduced by various drift removal methods in the generation of the overlapping statistic of $\sigma_{\Delta,3}(\tau,x_j(t_{n'},x(t_n)))^2$, the drift-removed $\sigma_x(\tau)^2$ the unmodified Allan variance of x(t) [1]. The drift estimators we will consider are

(a) a 2nd order polynomial ULSF,

(b) the Greenhall frequency drift estimator $\hat{c}(r)$ [12], and

(c) a 6th order polynomial ULSF.

$K_{stat}(f)$ results for these drift removal methods are shown in Figure 3 for $\tau/T = 0.5$, where bias effects on $K_{stat}(f)$ are maximized. We note that bias effects on $K_{stat}(f)$ for all these drift estimators quickly disappear as τ/T becomes smaller than 0.5. For reference, the non-drift removed $K_{stat}(f)$ is also shown.

In Figure 3, only the lower portion of the spectrum near $f\tau = 0.5$ is shown. This is because drift-removed $K_{stat}(f)$ bias changes are negligible at higher spectral frequencies. In [12], Greenhall noted that the effect of kernel bias on the Allan variance varies drastically with the noise type. From Figure 3, one can see why this occurs; as p in $L_x(f) \propto |f|^p$ changes from white noise to more negative power law orders, the region near $f = 0.5/\tau$, where the maximum $K_{stat}(f)$ bias change occurs, is more and more heavily weighted in the σ_{stat}^2 integral (38). Note for white noise, that bias changes to σ_{stat}^2 are minimal if $f_h >> 1/\tau$, because white noise equally weights the whole spectrum up to f_h. In fact, for this white noise case in the limit of $f_h\tau \to \infty$, the fit-removed σ_x^2 is the same as the non-fit removed one. This is the source of the misconception that fit removal for large N can completely remove frequency drift effects without biasing σ_x^2.

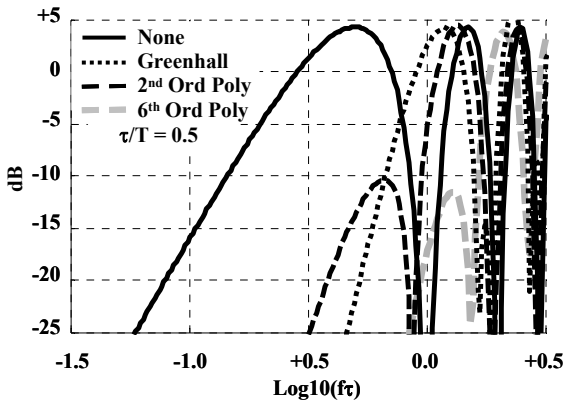


Figure 3. The Kernels $K_{stat}(f)$ for the unmodified Allan variance of x(t) with various forms of drift removal at $\tau/T = 0.5$.

Figure 3 shows the alteration in $K_{stat}(f)$ for 6th order polynomial ULSF removal, as well as 2nd order polynomial ULSF removal. This case is important in understanding bias effects when dealing with long term data where a constant frequency drift model is no longer valid. This is true for long term crystal oscillator aging, which has long term frequency aging behavior given by a sum of $A\cdot\ln(1+Bt)$ terms, and for environmentally induced drift large T, which typically has quite complicated behavior over long observation intervals.

One can see from Figure 3 that $K_{stat}(f)$ bias changes due to 6th order polynomial ULSF removal are quite drastic for large τ/T. One can also show that this 6th order polynomial case is typical for polynomial removal with orders greater than 3. We note that such (M'-1)th order polynomial USLF drift removal highpass filters the non-fit removed $K_{stat}(f)$ with $|f|^{2M'}$ HP behavior and a highpass knee of about $f = 1/T$ [5][7]. An importance consequence of this is that a Δ-variance that is a natural random stability measure without drift removal for intrinsic aging will have significant $K_{stat}(f)$ bias, if environmental factors contribute significantly to the drift over T and have complicated temporal behavior.

A final approach utilized in removing drift from long term data is piece-wise drift removal over portions T/N' of the total data interval T. One can show, however, that this causes the highpass filtering knee frequency that alters the drift-removed $K_{stat}(f)$ to move from 1/T to N'/T. This not only increases the $K_{stat}(f)$ bias for τ values on the order of 0.5T, but also causes this increase to remain appreciable for τ values down to 0.5T/N'.

IV. Summary and Conclusions

This paper has introduced a physical definition of Mth order random stability as the prediction error from an ideal M-

point calibration solely due to random noise. This definition allows one to define M^{th} order random stability, not only for polynomial fit prediction with uniform sampling intervals, but also for non-polynomial fit prediction with non-uniform sampling intervals. Using this definition and the properties of M^{th} order Δ-measures derived in the text, we have shown that M^{th} order Δ-measures and their variances without drift removal are natural measures of M^{th} order polynomial random stability when the drift over T is in the form of an $(M-1)^{th}$ or lower order polynomial. We have also provided a quantitative method for determining the bias in a fit-removed M^{th} order Δ-measure used to estimate M^{th} order polynomial random stability when the drift over T is not in the form of an $(M-1)^{th}$ or lower order polynomial. An important conclusion of the paper is that a Δ-variance that is a natural random stability measure without drift removal for intrinsic aging will have significant $K_{stat}(f)$ bias, if environmental factors contribute significantly to the drift over T and have complicated temporal behavior.

REFERENCES

[1] IEEE 1139-2008: Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology—Random Instabilities, IEEE, 2008.

[2] R. A. Baugh, "Frequency Modulation Analysis with the Hadamard Variance," in Proc. 25th Annual Frequency Control Symposium (IEEE), 1971, pp. 222–225.

[3] George Boole, A Treatise On The Calculus of Finite Differences, 2nd ed., New York: Chelsea, 1872.

[4] Isaac Newton, The Mathematical Principals of Natural Philosophy, First Paperbound Ed, New York: The Citadel Press, 1964 (orig. 1687), Book III, Lemma V, Case 1.

[5] V. S. Reinhardt, "How extracting information from data highpass filters its additive noise," Proc. 39th PTTI Systems and Applications Meeting, Dec., 2007, pp. 559-580.

[6] V. S. Reinhardt, "Modeling Negative Power Law Noise," Proc. 2008 IEEE IFCS, pp. 685-592.

[7] V. S. Reinhardt, "The Profound Impact of Negative Power Law Noise on Statistical Estimation," IEEE Trans. UFFC, vol. 57, no. 1, pp. 74-81, Jan. 2010.

[8] K. S. Miller, An Introduction to the Calculus of Finite Differences and Difference Equations, New York: Dover, 1966.

[9] http://en.wikipedia.org/wiki/Charles_Babbage, 22 February 2011.

[10] C. A. Greenhall, Private communication, June, 2009.

[11] S. M. Kay, Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory, Saddle River: Prentice Hall, 1993.

[12] C. A. Greenhall, "A Frequency-Drift Estimator and Its Removal from Modified Allan Variance," Proc. 1997 IEEE IFCS, pp. 428-432.

Novel Concept of Sub-Picosecond Timing System and its Applications in Fundamental Metrology

Ivan Prochazka and Jan Kodet
Czech Technical University
Prague, Czech Republic
E-mail: prochazk@cesnet.cz

Petr Panek
Institute of Photonics and Electronics
Academy of Sciencies of the Czech Republic
Prague, Czech Republic

Ulrich Schreiber
Wettzell Observatory
Technical University Munich
Bad Koetzting, Germany

***Abstract*— We are presenting a novel concept of the time interval measurement which uses a surface acoustic wave (SAW) filter as a time interpolator. Based on our previous results we have designed and constructed a timing device which allows registration of times of arrival of pulses with a single shot precision of 800 fs RMS, the time interpolation nonlinearity bellow 200 fs, and the stability in sense of TDEV <10 fs for averaging times from 300 s to 3000 s. Two identical timing devices above have been used in an experiment, in which the frequency outputs of two Hydrogen masers were compared. The experiment proofed that one of possible applications of the timing device is a comparison of high performance low-noise frequency sources and measurement of their frequency stability in the time-domain. The noise background of the timing system was negligible for averaging intervals on the order of 100 ms and longer.**

I. INTRODUCTION

A novel time interval measurement method that makes use of a SAW filter as a time interpolator has been introduced in [1]. The method is based on the fact that a transversal SAW filter excited by a short pulse generates a finite-time signal with highly suppressed spectra outside a narrow frequency band. It results from the sampling theorem that if the responses to two excitations are sampled at clock ticks, they can be precisely reconstructed from a finite number of samples and then compared so as to determine the time interval between the two excitations.

A detailed analysis of measurement errors of this method has been given in [1] and [2]. It results from this analysis that the method excels in time interpolation efficiency since the time interpolation error relative to clock period is very small. It means that an accurate measurement can be achieved even with relatively low clock frequency. Also random errors caused by noises generated in different parts of the timing device are very low thanks to averaging effect as the interpolation process is not based on the only one but many observations. A special care has been given to minimization of non-linear effects in the timing device [3]. The experimental results [4], [5] proofed the new concept. The extreme single-shot precision, time interpolation linearity as well as the timing stability were in agreement with the expectations.

II. TIMING DEVICE

Based on the previous results we have designed and constructed a timing device which allows measurement and registration of times of arrival of input pulses with respect to a local time base. The block scheme of the timing device is in Fig. 1. An input pulse excites a SAW filter whose output is sampled at clock ticks and converted by an analog-to-digital converter (ADC) and buffered. Once the response to the input pulse has been processed, an internally generated pulse is applied to the same circuit. This internal pulse is generated synchronously with the time base reference. After both responses have died out, they are reconstructed from the samples and compared by means of cross-correlation to determine the time interval between the two pulses. All the processing is integrated into a simple algorithm based on the fast Fourier transform. The device uses the clock sampling frequency of 100 MHz. The SAW filter central frequency is of 525 MHz, and the filter bandwidth is of 18 MHz.

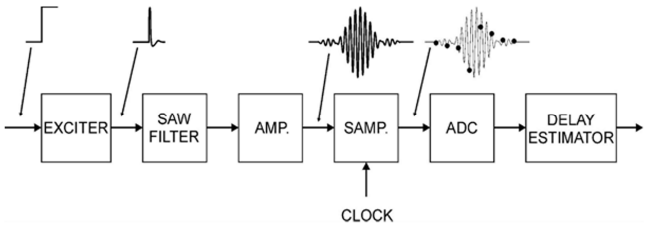


Figure 1. The block scheme of the timing device.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

The application of the internally generated time mark ensures calibrating the internal delays on a shot-by-shot basis hence the extreme timing stability and reproducibility of the measurements is achieved.

The internal time base is based on an external frequency multiplier providing spectrally pure 200 MHz sine wave multiplied from a 10 MHz frequency reference. The conversion of the 200 MHz sine wave to a pulse clock signal and its distribution to the ADC and the excitation circuit use high speed SiGe comparators and buffers.

To be able to check the measurement performance and do different experiments a reference time mark generator has been included into the timing device. It generates low jitter pulses synchronously to the local time base. The repetition frequency of these pulses is $100\ \text{MHz}/2^{17} = 763$ Hz and their jitter is below 300 fs RMS which is practically negligible with respect to the single-shot precision of the timing device itself.

The timing device is constructed on a single PCB (Fig. 2) which is installed in an Aluminum box (Fig. 3) to ensure the proper passive heat transfer and high timing stability. The box dimensions are 250 x 140 x 45 mm. Two identical timing devices described above together with a power supply unit are placed in a 19" case and create a two-channel NPET timing device (Fig. 4). The other 19" case includes the frequency multiplier 10 MHz to 200 MHz. The NPET communicates with a personal computer via USB. The input signals are connected through SMA connectors.

The performance of the timing device has been extensively tested using the reference time marks generated synchronously to the local time base as described above. The resulting histogram of the TOAs is in Fig. 5. The single-shot precision evaluated from these measurements was nearly 800 fs RMS.

Thanks to the device concept and construction an excellent timing stability has been obtained. Operating the device in a common laboratory environment without temperature stabilization, the TDEV better than 4 fs was routinely achieved for averaging times from 300 s to 3000 s (Fig. 6).

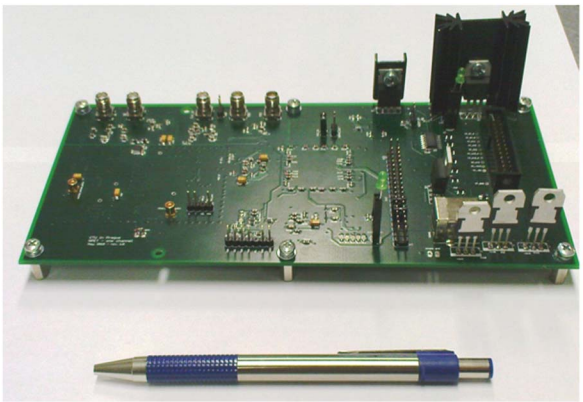


Figure 2. The timing device is constructed on a single PCB.

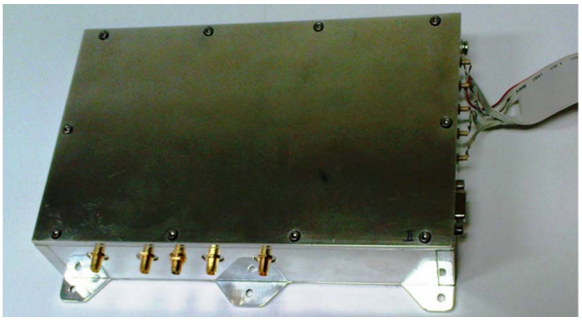


Figure 3. The timing device PCB is installed in an Aluminium box.

Figure 4. Two identical timing devices together with a power supply unit are placed in a 19" case and create a two-channel NPET timing device. The other 19" case includes the frequency multiplier 10 MHz to 200 MHz.

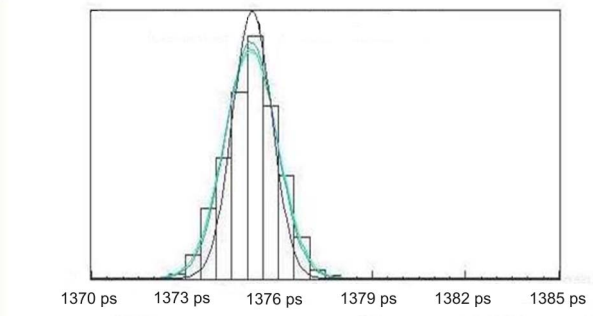


Figure 5. The resulting histogram of TOAs, 500 fs per bin. The single-shot precision evaluated from these measurements is nearly 800 fs RMS.

The timing stability can also be demonstrated on the 256 s moving average of the TOAs (Fig. 7) which did not leave the +/- 4 fs range within the period of 3 hours.

We further completed a series of measurements focused on the linearity of the time interpolation. In this case we used an external pulse generator not synchronous with the local time base. The pulses were connected to inputs of two identical timing devices through cables with different delays and the difference of TOAs in dependence on the TOA was measured. Several such measurements with different differential delay of the input pulses were completed. The resulting nonlinearity was evaluated below 200 fs.

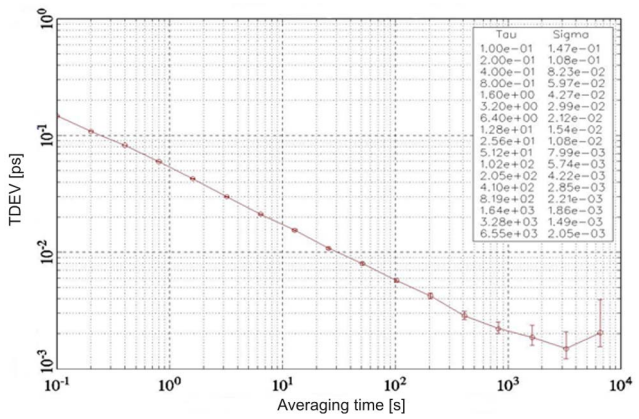


Figure 6. Operating the device in a common laboratory environment, the TDEV better than 4 fs was routinely achieved for averaging times from 300 s to 3000 s. The instability can be considered as a white phase noise for averaging times up to ~1000 s.

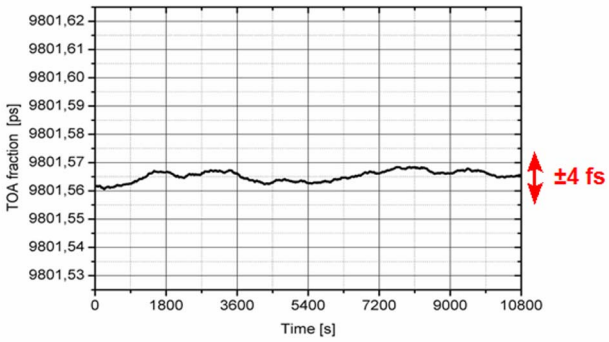


Figure 7. The 256 s moving average of the TOAs did not leave the ± 4 fs range within the period of 3 hours.

III. EXPERIMENTS

Two timing devices described above have been used in an experiment in which the frequencies of two H-masers were compared. The experiment was performed in the underground time and frequency laboratory at the Geodetic Observatory Wettzell, Germany. The experimental setup is in Fig. 8. There are from left to right the hydrogen maser EFOS SN 60, the data acquisition PC #1, the NPET timing device (channels #1 and #2), frequency multiplier #1, and the data acquisition PC#2. The other maser EFOS SN 39 and the frequency multiplier #2 were located in the next room.

At first both the timing devices were connected to the same frequency reference (H-maser SN 39 and frequency multiplier #1) according to the block scheme in Fig. 9a. The aim of this measurement was to prove the jitter and time stability of the experimental setup including the cables. The jitter was on the expected level of 800 fs RMS. The resulting TDEV is in Fig. 10.

Figure 8. The experimental setup. There are from left to right the hydrogen maser EFOS SN 60, the data acquisition PC #1, the NPET timing device (channels #1 and #2), the frequency multiplier #1, and the data acquisition PC#2. The other maser EFOS SN 39 and frequency multiplier #2 were located in the next room.

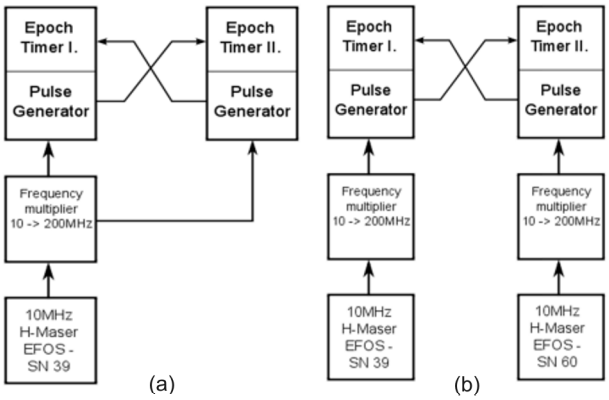


Figure 9. The block scheme for the first measurement (a) when both the timing devices were connected to the same frequency reference and the second measurement (b) when each timing device was connected to another H-maser.

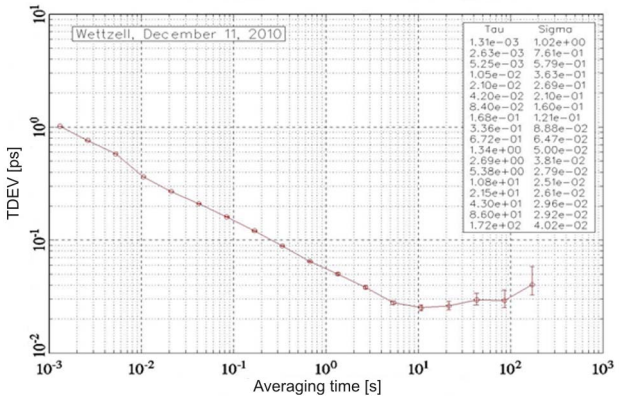


Figure 10. The resulting TDEV of the experimental setup when both the timing devices were connected to the same frequency reference.

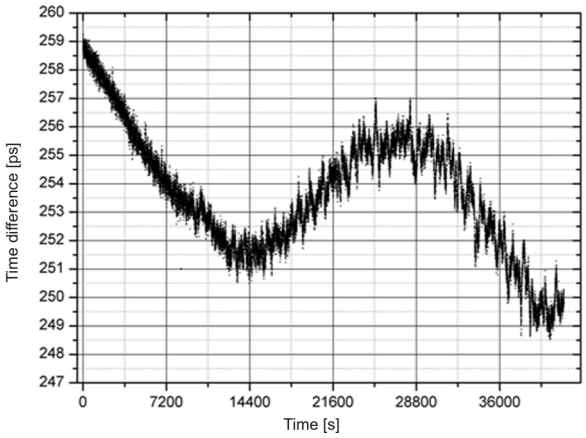


Figure 11. The record of the phase difference between two H-masers.

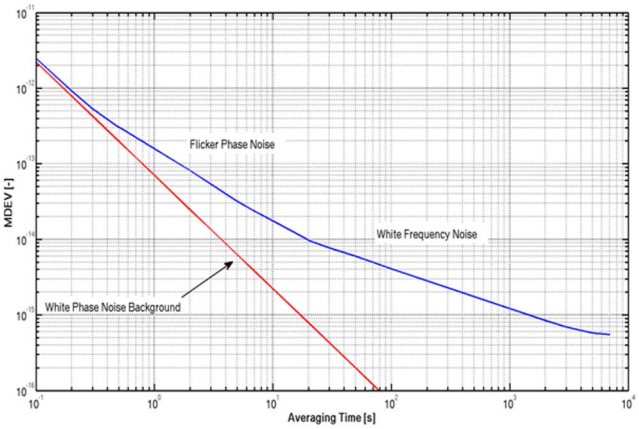


Figure 12. The evaluated MDEV of the phase difference between two H-masers. The red line represents the white phase noise background caused by the timing device. It is negligible for averaging times greater than ~400 ms. For longer averaging times the MDEV follows the slopes of flicker phase noise and white frequency noise as expect

After that each timing device was connected to another frequency reference according to the block scheme in Fig. 9b and the phase difference between the H-masers was measured. The resulting record of the phase difference is in Fig. 11 and the evaluated MDEV is in Fig. 12. The white phase noise background caused by the timing device (red line) was computed from the results of the previous measurement. It is negligible for averaging times greater than ~400 ms. For longer averaging times the MDEV follows the slopes of flicker phase noise and white frequency noise as expected [6].

IV. Conclusion

Two timing devices of a novel concept have been used in a timing experiment in which the frequency outputs of two H-masers were compared. The experiment proofed that one of possible applications of the timing device is a comparison of high performance low-noise frequency sources and measurement of their frequency stability in the time-domain for a wide range of averaging intervals. The noise background of the timing system was negligible for averaging intervals on the order of 100 ms and longer. The excellent timing precision near 800 fs RMS and stability well below 10 fs of the novel principle timing system has been demonstrated.

Acknowledgment

This work has been carried out at the Czech Technical University in Prague. The research and development of solid state photon counting detectors and their applications is supported MSM6840770015.

References

[1] Panek P., "Time Interval Measurement Based on SAW Filter Excitation," IEEE Trans. Instr. Meas, vol. 57, no. 11, pp. 2582-2588, Nov 2008.

[2] Panek, P.: “Random Errors in Time Interval Measurement Based on SAW Filter Excitation. IEEE Transactions on Instrumentation and Measurement”, vol. 57, no. 6, pp. 1244-1250, Jun 2008.

[3] Prochazka, I. - Panek, P.: “Nonlinear effects in the time measurement device based on surface acoustic wave filter excitation”. Review of Scientific Instruments. 2009, vol. 80, no. 7, p. 076102.

[4] Panek, P.; Prochazka, I.: “Time interval measurement device based on surface acoustic wave filter excitation, providing 1 ps precision and stability”. Review of Scientific Instruments, vol. 78, no. 9, Sep 2007.

[5] Panek, P. - Prochazka, I. - Kodet, J.: “Time measurement device with four femtosecond stability”. Metrologia. 2010, vol. 47, no. 5, p. L13-L16.

[6] Riehle, F., “Frequency Standards: Basics and Applications", Wiley-VCH, 2005, ISBN: 3527402306, p. 540.

Acceleration "G" Compensated for VCOCXO Based on Digital Controller

QingXiao Shan，Yang Jun, JianYun chen, Tang Qian, LongZhe Ji
Mechatronic and Automation school
National University of Defense Technology
Changsha, Hunan, China
Email:chorely@yahoo.cn

Abstract- *The paper proposes a digital compensation by imposing real-time compensation voltage to Vin Pin of VCOXCO while device vibrating. The embedded digital compensating circuit includes a MEMS three-axes acceleration sensor, a FPGA and a D/A converter. G-sensitivity vector can be computed based on the MEMS sensor. A proportion digital compensation model is built to generate voltage which can compensate frequency drift and perturbation caused by acceleration.*

□. INTRODUCTION

Quartz crystal oscillator is sensitive to temperature and acceleration[1]. OCXO can maintain temperature stable to increase stability of output frequency, so OCXO is widely applied at precise equipment. Most OCXO are VCOCXO whose frequency can be adjusted from 10^{-7} to 10^{-5} by controlling voltage. OCXO is sensitive to acceleration which causes equipment performance deterioration at moving applications. The range of typical g-sensitivities spans from 1×10^{-10} Hz/g for precision SC cut to greater than 1×10^{-7} Hz/g for a low cost AT cut.

At present, to decrease g sensitivity, the way used most is mechanic cushion which is useful only to high frequency vibration. It can't operate effectively when the device suffers low frequency vibration or continual g movement. Additionally, mechanic cushion increases the volume and weight of devices, which is harmful to miniature devices.

Gagnepain uses two resonators in series electrically and reverses parallel connection mechanically, so g sensitivity was counteracted[2]. Owing to lack of repeatability in the acceleration sensitivity, it's difficult to make alignment. Greenray corp manufactures dual-resonator oscillators using AT-strip quartz whose g sensitivity is repeatable in both magnitude and direction, so the g sensitivity can be below 1×10-11.[3]

Przyjemski[4] and Emmons[5] used an accelerometer aligned with the g sensitivity direction. The accelerometer signal was fed into a tune circuit to decrease g sensitivity. As the circuit was analog, it's hard to be linear at all point. Some scholars[6] proposed three accelerometers mounting with the oscillator and aligned on mutually perpendicular axes and a DSP was applied to generate compensation voltage to VCOCXO. But it was complicated to install three accelerometers and it would take some time for DSP to compute and generate compensation value.

This paper proposes real time digital compensation based on FPGA process. A MEMS three axes acceleration sensor whose package is less than 3mm×3mm×0.9mm is applied to get the acceleration vector. A FPGA is employed which can realize computation in one clock period. Because of the integration and miniaturization of digital device, the whole compensation system is flexible, low-cost and easy to be applied.

Ⅱ. PRINCIPLE OF DIGITAL COMPENSATION

The research is carried out on the basis of Voltage Controlled OCXO for several reasons. First, the temperature of OCXO is stable so temperature sensitivity can be substantially reduced. Second, with the big dimension, digital controller can be embedded into OCXO. At last it's easy to realize the compensation by adjusting the "Vin" pin of VCOCXO. So little change is needed to be made to the original OCXO.

A. Frequency domain analysis

When sinusoidal vibration is imposed on OCXO, there will be vibration-induced sideband in the carrier, the ratio of the power can be denoted as follows[1]:

$$L_V^n(dBc) = 20\log(J_n(\beta)/J_0(\beta)) \tag{1}$$

Where $J_n(\beta)$ is the Bessel function of nth sideband.

$$\beta = \Delta f / f_V = (\vec{\Gamma}\cdot\vec{A}) f_0 / f_v \tag{2}$$

is the modulation index.

Where f_0 : output frequency of resonator; f_v : vibration frequency; $\vec{\Gamma}$ is g sensitivity vector and $\vec{A}$ is acceleration g vector.

Several approximations can be made if $\beta < 0.1$. The approximation is

$$J_0(\beta) = 1, \quad J_1(\beta) = \beta / 2, \quad J_n(\beta) = 0 , \text{ so}$$

$$L_v^1(\beta) \approx 20 * \log((\vec{\Gamma} \cdot \vec{A}) f_0 / (2 f_v)) \tag{3}$$

B. Building acceleration sensitivity vector model

When acceleration exists, the quartz resonator will generate stress which would corrupt its RF outputs. The package type of resonator would also affect g-sensitivity. Resonators would be distorted when large-scale g, which causes non-linear g-sensitivity, is imposed on.

The direction of g-sensitivity vector was fixed for each OCXO. A MEMS three dimension g sensor was installed at the package of OCXO to sense the acceleration. Supposing that axes of g sensor was x,y,z and g sensitivity angle of each axis was $\varphi_x, \varphi_y, \varphi_z$, it's shown as Fig.1.

The value of g sensitivity vector can be described as

$$a_m = a_x \cdot \cos\varphi_x + a_y \cdot \cos\varphi_y + a_z \cdot \cos\varphi_z \tag{4}$$

a_x, a_y, a_z was measured by sensor, so if $\varphi_x, \varphi_y, \varphi_z$ were known, a_m could be computed.

If the vibration was imposed only along one axis, g sensitivity would be only caused by one axis. Supposed that vibration was imposed on x axis, so

$$a_m = a_x \cdot \cos\varphi_x = \vec{A} \tag{5}$$

We imposed sinusoid vibration at fixed frequency, so from equation (2),(3) we could get

$$\cos\varphi_x = \vec{A} / a_x = 10^{L_v^1(\beta)/20} \cdot 2 f_v / (f_0 \cdot k_a) \tag{6}$$

Where g-sensitivity coefficient is k_a .

By using this method, $\cos\varphi_x, \cos\varphi_y, \cos\varphi_z$ could be obtained. Fig.2 shows g-sensitivity test platform. We impose sine wave vibration upon one axis. Output of resonator was multiplied to increase the acceleration effect. The ratio power $L_v^1(\beta)$ could be provided by a spectrum analyzer.

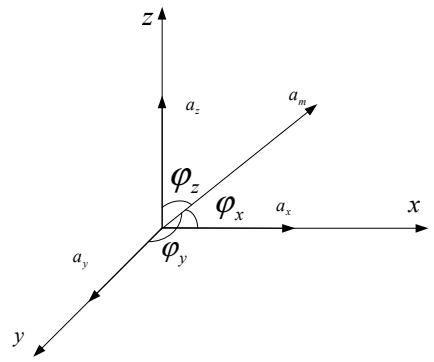


Figure 1. G-sensitivity vector with three axes of sensor

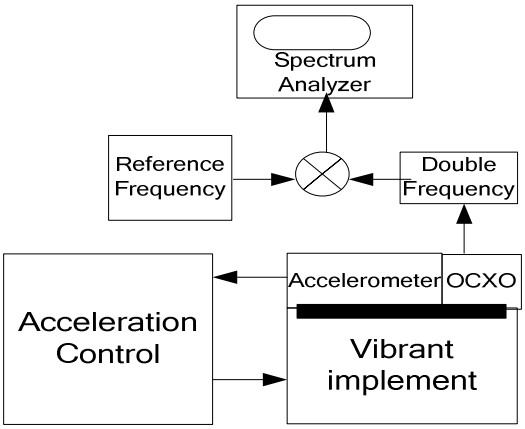


Figure 2. G-sensitivity test platform

C. Proportion digital real-time compensation

Supposed that voltage control coefficient was k_v (Hz/v), g-sensitivity coefficient was k_a and the real-time generated compensation voltage was

$$\Delta v = k_p \cdot \frac{k_a \cdot a_m}{k_v} \tag{7}$$

Where k_p was proportion coefficient which would determine the compensation effect.

The algorithm diagram is shown in Fig.3. Fig.4 shows the compensation result. f_{v1} is frequency drift caused by acceleration, f_{v2} is frequency drift caused by imposed voltage Δv which counteracts f_{v1} . f_{vc1} is frequency drift after compensation which is much less than itself before compensation.

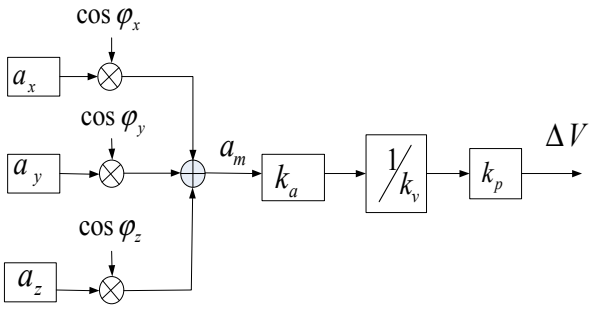


Figure 3. Algorithm diagram

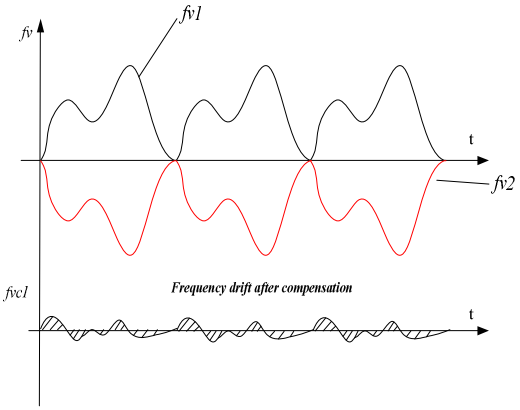


Figure 4. Frequency drift after compensation

D. Time delay effect

Suppose that a time delay Δt exists at compensation for it costs time to sense and compensate. If sinusoidal wave vibration is imposed, the RF output after compensation can be described as follows

$$V(t) = V_0 \cos(2\pi f_0 t + (f_0 \cdot \vec{\Gamma} \cdot \vec{A} / f_v) \sin 2\pi f_v t - (f_0 \cdot k_v \cdot \Delta v / f_v) \sin 2\pi f_v (t - \Delta t)) \quad (8)$$

When there is an ideal compensation, $k_v \cdot \Delta v = \vec{\Gamma} \cdot \vec{A}$, so

$$V(t) = V_0 \cos(2\pi f_0 t + (f_0 \cdot \vec{\Gamma} \cdot \vec{A} / f_v)(\sin 2\pi f_v t - \sin 2\pi f_v (t - \Delta t))) \quad (9)$$

III. DIGITAL COMPENSATION SYSTEM DESIGN

For acceleration device experience varies with high amplitude, it is hard to maintain linear in any work state when analog circuit is employed. What is more, it is hard to compute the sensitivity vector with analog circuit.

We designed digital compensation circuit based on FPGA circuit shown as Fig.5. A MEMS triaxial sensor was applied which could output three axes acceleration values and a FPGA with 45nm low-power copper process technology is also employed. FPGA got the value from sensor by SPI interface. A serial D/A converter was applied to generate the compensation voltage. Dc-dc converter and voltage regulating module provided necessary voltage for FPGA. The board size of digital circuit was designed based on OCXO, so it could be installed just on the pin of OCXO shown as Fig.6. Furthermore, the total circuit could be embedded in the OCXO, which made it easy being widely applied.

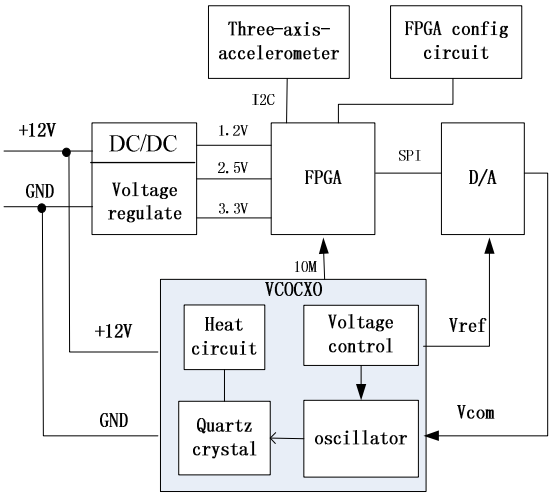


Figure 5. Digital compensation scheme

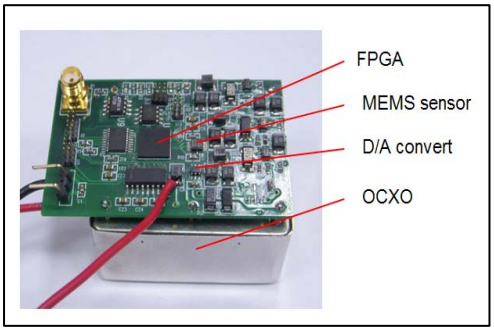


Figure 6. Experiment prototype

IV. EXPERIMENT

An embodiment of proposed compensation circuit is designed which consists of a MEMS acceleration sensor, a FPGA system and a 12 bit D/A converter. The parameter of OCXO can be described as follow: $f_0 = 10MHz$, $k_g = 5.06 \times 10^{-9}$, $k_v = 8 \times 10^{-8}$. With the 12bit D/A, the frequency controlling resolution is $k_v / 4096 = 1.95 \times 10^{-11}$ which is much less than k_g.

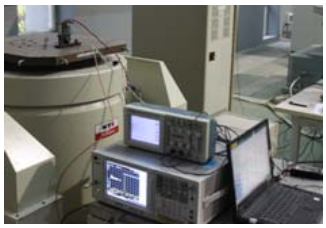
And experiment platform is also built. It includes:

1) DC power supply;
2) digital storage oscilloscope TDS1012;

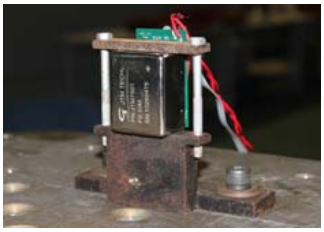
3) Spectrum Analyzer E4443A;

4) vibration controlling system;

The experiment environment is shown in Fig.7(a). The OCXO is installed on the platform of vibration system as shown in Fig.7(b). Compensation delay is 7.5us.



(a) Experiment platform

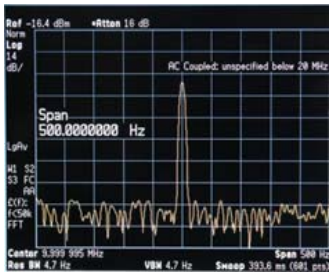


(b) OCXO fixed by clamp

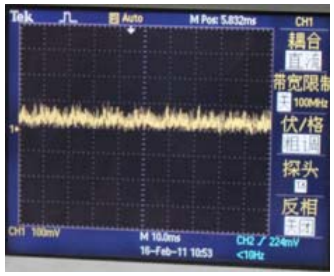
Figure 7. Experiment environment

The output spectrum of OCXO at rest is shown in Fig.8(a). Then 15Hz 4g sine wave vibration is imposed and output spectrum is shown in Fig.8(c). Fig.8(b) shows the compensation voltage and Fig.8(d) shows the output spectrum after compensation while kp=4. We can see 11.9dB attenuation is gotten. Then 30Hz, 300Hz experiments were also carried out. From recorded data we can see that g sensitivity is much decreased after compensation.

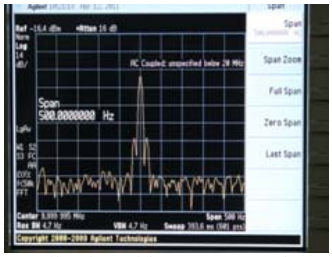
We can see from Tab.1 that the digital compensation gets good result when low frequency vibration is imposed. But it is less effective when high frequency vibration is imposed. The reason is that delay may cause this problem.



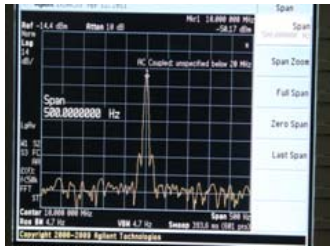
(a) Output at rest



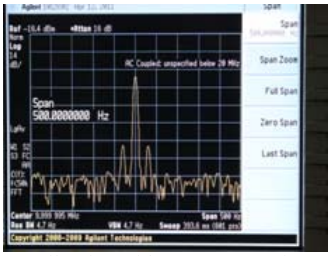
(b) Compensation voltage



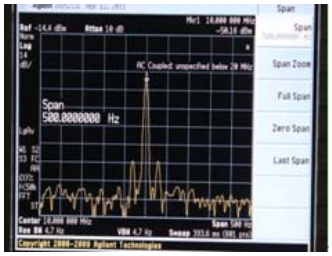
(c)Vibration without compensation Gmax=4g, f=15Hz



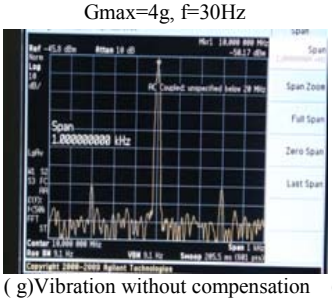
(d) Vibration with compensation kp=4 Gmax=4g, f=15Hz



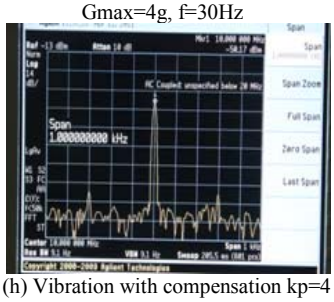
(e)Vibration without compensation Gmax=4g, f=30Hz



(f) Vibration with compensation kp=4 Gmax=4g, f=30Hz



(g)Vibration without compensation Gmax=6g, f=300Hz



(h) Vibration with compensation kp=4 Gmax=6g, f=300Hz

Figure 8. Experiment result

TABLE I. COMPARISON AFTER COMPENSATION

$g_{\max}$ *(g)*	f_v *(Hz)*	$L_v^1(\beta)$ *(dB)*	$L_v^1(\beta)$ *(dB) With compensation*
4g	15	-47.6	-59.5
4g	30	-54.6	-60.2
6g	300	-66.2	-71.5

V CONCLUSION

A new acceleration compensation technique which can substantially reduce the acceleration sensitivity with low cost is proposed. The total cost of digital circuit is less than $20 and the compensation makes very little change to OCXO. It is easy to realize digital compensation for VCOCXO. The method has good future for high accuracy OCXO.

REFERENCES

[1] R.L. Filler, "The Acceleration Sensitivity of Quartz Crystal Oscillators: A Review," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control VOL. 35, No. 3, pp.297-305，May 1988,

[2] J.-J. Gagnepain and F. L. Walls, "Quartz crystal oscillators with low acceleration sensitivity," NBSIR 77-855, National Bureau of Standards, 1977

[3] StevenSteven J. Fry, Gregory A. Burnett, Reducing the acceleration sensitivity of AT-strip quartz crystal oscillators, 2010 IEEE frequency control symposium, page:25-30

[4] J. M. Przyjemski, "Improvement in system performance using a crystal oscillator compensated for acceleration sensitivity," 32nd Annu. Symp. Freq. Contr., 1978, pp. 426-431.

[5]D. A. Emmons, "Acceleration sensitivity compensation in high performance crystal oscillators," 10th PTTI Con$, 1978.

[6] Marvin E.Frerking, "Vibration compensated crystal oscillator",U.S.Patent No 4,891,611,1990

Electrically coupled MEMS Oscillators

Deepak K. Agrawal, Pradyumna Thiruvenkatanathan, Jize Yan and Ashwin A. Seshia
Department of Engineering,
University of Cambridge,
Cambridge CB2 1PZ, UK
Email: dka23@cam.ac.uk

***Abstract*—In this paper, we demonstrate synchronization of two electrically coupled MEMS oscillators incorporating nearly identical silicon tuning fork microresonators. It is seen that as the output of the oscillators are coupled, they exhibit a synchronized response wherein the output amplitudes and signal-to-noise ratios of the two oscillators are improved relative to the case where the two oscillators are uncoupled. The observed output frequency of each oscillator before coupling is 219402.4 Hz and 219403.6 Hz respectively. In contrast, when the oscillators are driven simultaneously, they lock at a common output frequency of 219401.3 Hz and their outputs are found to be out-of-phase with respect to each other. A 6 dBm gain in output power and a reduction in the phase fluctuations of the output signal are observed for the coupled oscillators compared to the case when the oscillators are uncoupled.**

I. INTRODUCTION

Over the last two decades, advancements in design, fabrication and packaging of silicon MEMS oscillators have resulted in their commercialization with the potential to replace quartz crystal oscillators from certain market sectors due to their smaller sizes, scalability to VHF-UHF frequencies and the potential for integration with CMOS electronics. However, MEMS oscillators are currently limited in their noise performance and output power capabilities, in part due to their smaller dimensions and the associated onset of non-linear self-limiting behaviour at relatively modest oscillation amplitudes.

Several methods have been previously employed to address these limitations. Of these, perhaps the most notable is the effort to enhance the noise performance of silicon MEMS oscillators by increasing the output power using array-coupled silicon microresonators [1]. In this work, an array of nine identically designed mechanically coupled wine-glass disk resonators operating in a combined resonant mode were embedded in an oscillator feedback loop to enhance the output motional current, thereby improving the phase noise performance. While it can be inferred from their results that the use of multiple resonators, enhances the oscillator response relative to the typical use of a single resonator, the use of multiple resonators also changes the modal dynamics, as each resonator adds an additional degree of freedom, making the system more susceptible to excitation of spurious modes. Moreover, electrostatic transduction of the resonator array in order to excite a particular combined vibration mode for an array coupled system can impose constraints on transducer geometry and routing interconnect as well as increase the effects of capacitive feedthrough parasitics.

Various noise sources with electrical, mechanical and environmental origins result in amplitude and phase fluctuations in the output of the MEMS oscillator [2]. Phase fluctuations are accumulated over time and are of particular interest in timing and frequency control applications [3]. In an attempt to address the phase stability of MEMS oscillators, we propose the utilization of the phenomenon of synchronization wherein synchronized oscillators lock at a common output frequency and a constant relative phase difference.

One of the first observations of synchronization was reported by the Dutch researcher Christiaan Huygens in the 17^{th} century. He observed that when two symmetric pendulum clocks were suspended from a common frame, they began to oscillate at a common frequency and out-of-phase with respect to each other due to the weak coupling that caused by the imperceptible motion of the frame [4].

More recently, Bennett *et al.* re-examined the phenomenon using two symmetric pendulum clocks and showed that the onset of synchronization was related to the extent of mismatch of their uncoupled oscillation frequencies and the strength of coupling between them [5]. The oscillation frequency for a MEMS oscillator is determined by the resonating element which is a MEMS resonator in our case. Tolerances in fabrication processes often result in variations in the dimensions of the resonator and in their material properties. Due to these variations, the nominal resonant frequencies across identically designed resonators are usually different. MEMS resonators can be coupled by mechanical or electrical methods [6]. As opposed to mechanical resonators in more conventional technologies, electrical coupling of MEMS resonators can be implemented by parallel-plate capacitive coupling elements that can be readily implemented in a standard micromachining process. In this work, the synchronization of two electrically coupled silicon MEMS oscillators is demonstrated.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

II. Device Description And Operation

A. Device Configuration

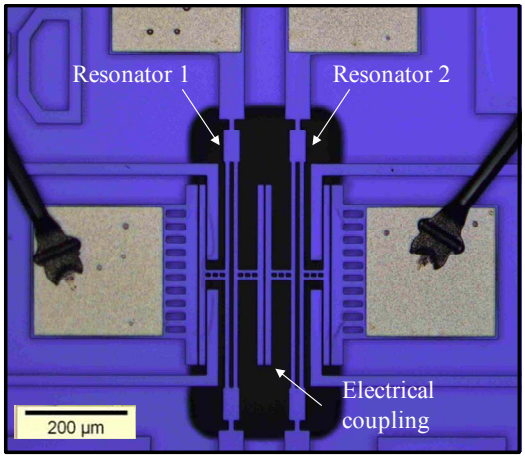


Figure 1. Optical micrograph of electrically coupled DETF resonators.

Fig. 1 shows an optical micrograph of two electrically coupled identically designed silicon double-ended-turning-fork (DETF) resonators [6]. The device is fabricated in a commercial foundry process using the silicon-on-insulator MEMS process through MEMSCAP Inc., USA. The two resonators are separated by a coupling gap as can be seen in Fig. 1. In this device, each tine is 350 µm long, 7 µm wide and 25 µm thick. Additional parallel plates are added at the point of maximum displacement to ensure uniform electrostatic actuation of the fundamental vibration mode. Each added electrode is 280 µm long, 10 µm wide and 25 µm thick. All the actuation gaps in both the resonators are 2 µm including the coupling gap. The out-of-phase normal mode of the resonator is of interest, as it offers a higher quality factor compared to the in-phase normal mode, due to the cancellation of stress at the anchor points. It is to be noted that any potential difference across the coupling gap generates an attractive electrostatic force which is displacement dependent and results in a negative mechanical spring like behaviour. The electrostatic coupling force (F_c) between two resonators can be determined by:

$$F_c = -\frac{\varepsilon_o A}{d^2}(V_{DC1} - V_{DC2})^2 \qquad (1)$$

Here V_{DC1} and V_{DC2} are the *DC* bias applied to each of the resonators. A is the actuation area, d is actuation gap and ε_o is permittivity of free space. It can be deduced from equation (1) that the change in potential difference across these resonators results in a modified coupling force.

B. Device Characterization

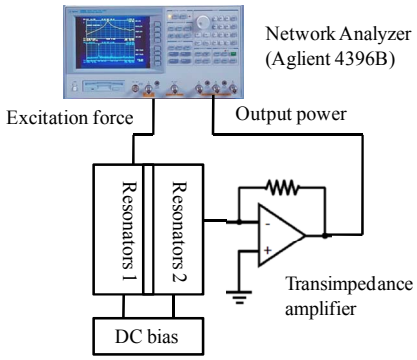


Figure 2. Schematic representation of open loop measurement setup for silicon DETF MEMS resonators.

A schematic representation of the measurement setup for open-loop experiments is shown in Fig. 2. The device is tested in a custom vacuum chamber at a pressure of ≈50 mtorr. Two-port electrostatic transduction is employed to detect the motion of the resonators. An *AC* excitation force is applied to the network analyzer to drive the resonators while *DC* bias is used to detect the motional current. The output motional current is amplified and converted to a voltage by a transimpedance amplifier and then sent to the network analyzer.

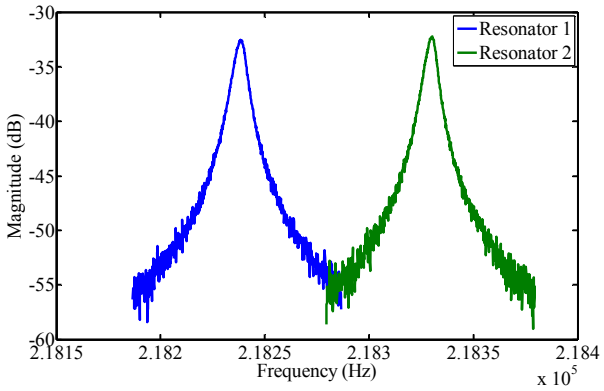


Figure 3. Open loop response of uncoupled DETF MEMS resonators.

The result of the measurements for the uncoupled response after feedthrough cancellation is shown in Figure 3. The feedthrough component is obtained by applying 0 V potential difference across the actuation gap. The quality factor and resonance frequency for each resonator are 31177, 32956 and 218238.3 Hz, 218330.0 Hz respectively. These values are determined by a nyquist circle fit to the measurements of the uncoupled response. Although, the resonators are designed to be identical, fabrication tolerances result in variations in dimensions and material properties, consequently leading to deviations in the resonance frequencies for identically

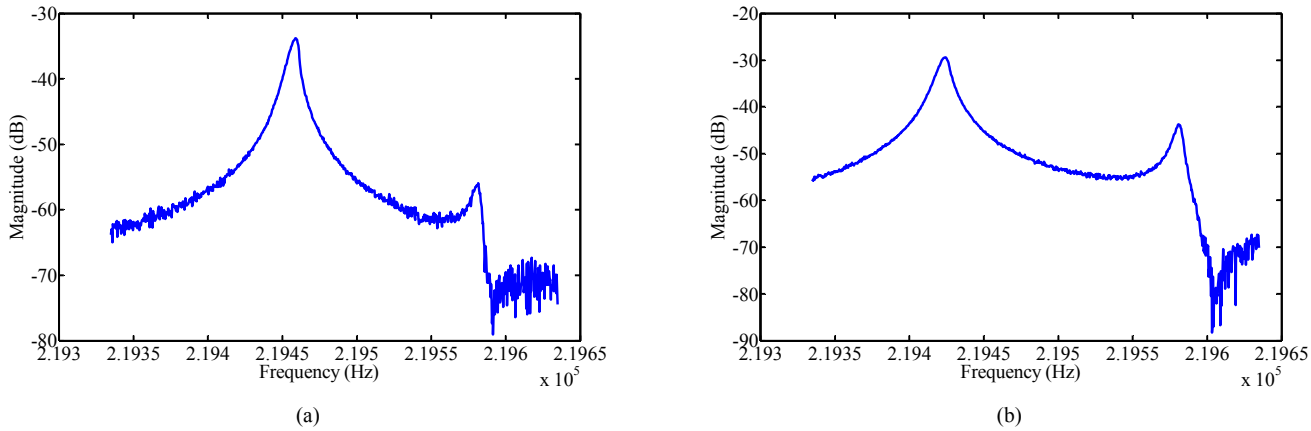


Figure 4. Open loop response of coupled DETF MEMS resonators observed at resonator 1 under two different conditions of electrical coupling (a) 4 V potential difference across the coupling gap (b) 8 V potential difference across the coupling gap.

TABLE I. SUMMARY OF EMPIRICAL RESULTS

Oscillator	**Frequency (Hz)**	**Output power (dBm)**	**Phase ratio (degrees)**	
			Mean	***Standard deviation***
Uncoupled Oscillator 1	219402.4	-55.6	169.4	41.4
Uncoupled Oscillator 2	219403.6	-57.6	156.5	54.9
Coupled Oscillator 1	219401.3	-50.8	180.4	18.2
Coupled Oscillator 2	219401.3	-50.9	178.8	20.4

designed structures.

The coupled response of the resonators is shown in Fig. 4(a) and (b) for two different coupling forces. In first case, the potential difference across the coupling gap is 4 V while it is 8 V in the second case. The coupled response is observed at resonator 1. The lower frequency mode is found to be out-of-phase normal mode which occurs first, due to the negative value of the electrical coupling spring.

III. IMPLEMENTATION OF ELECTRICALLY COUPLED OSCILLATORS

Fig. 5 illustrates the block diagram representation for the coupled MEMS oscillators. The two DETF resonators are embedded within the feedback loop of two distinct oscillator circuits. Each oscillator comprises a transimpedance amplifier for converting the motional current to a voltage and to meet the loop gain requirement for sustained oscillation. This block is followed by a band pass filter which removes unwanted normal modes. The filter output is then fed to a comparator. The comparator acts as a hard voltage limiter and limits the effect of amplitude fluctuations in the feedback signal. For both oscillators, output signals are observed at the output of the transimpedance amplifier as shown in Fig. 5. The response of the above configuration is compared for four different cases and the results are tabulated in Table I.

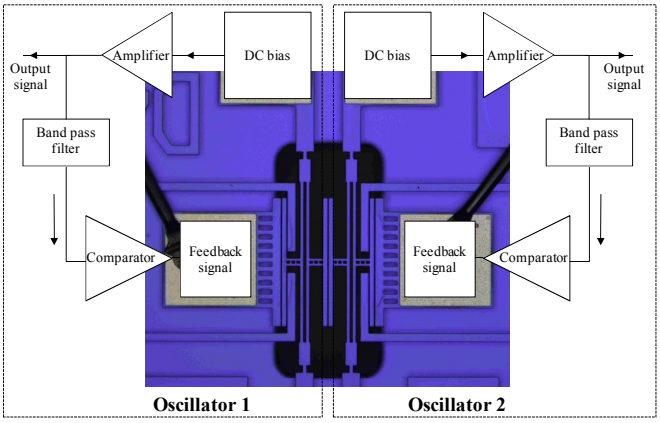


Figure 5. Schematic description of electrically coupled MEMS oscillators.

The first two cases are when both the oscillators are uncoupled (operated independently) and the next two cases when both are coupled (operated simultaneously). The observed output frequencies for oscillators 1 and 2 when operated independently, are 219402.4 Hz and 219403.6 Hz. When operated simultaneously, the oscillators lock at a common oscillating frequency of 219401.3 Hz as shown in Fig. 6(a) and (b) respectively.

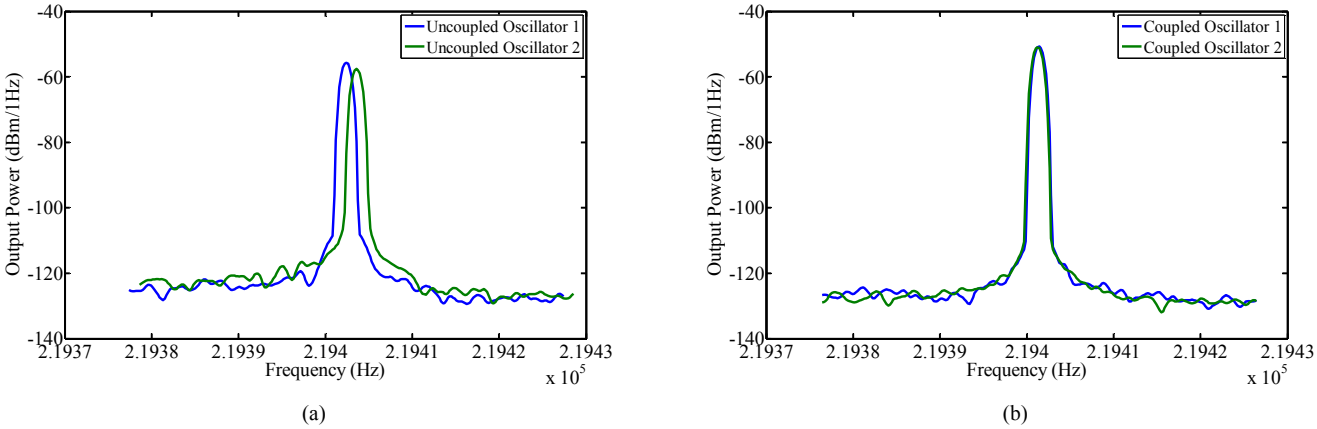


Figure 6. Output response of the MEMS oscillators compared before and after capacitive coupling (a) oscillators are uncoupled (b) oscillators are coupled.

In order to compare the oscillator response for uncoupled and coupled cases, a frequency normalized plot for oscillator 2 is shown in Fig. 7. A 6 dBm gain in the output power and a clear reduction in noise floor can be observed for the case when the oscillators are coupled relative to the case when the two oscillators are uncoupled. Both oscillators describe the same general behaviour.

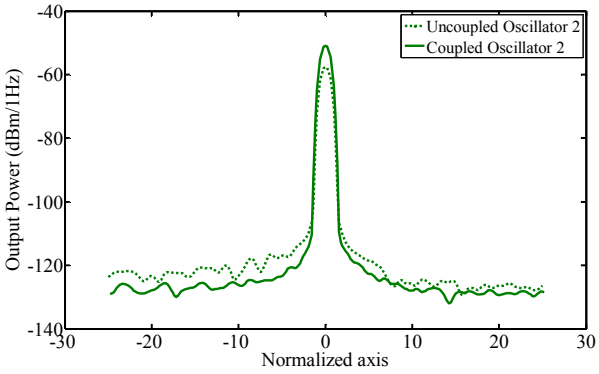


Figure 7. Frequency normalized response of oscillator 2 for uncoupled and coupled cases.

Table I summarizes the empirical results for the uncoupled and coupled MEMS oscillators. The phase ratio describes the relative phase of oscillator 1 with respect to oscillator 2. These measurements are done using a frequency counter (Agilent 53230A) at a sampling rate of 0.5 sec for a total measurement time of 250 sec. For each case, the mean and standard deviation of the data are provided. From Table I, it can be seen that the coupled MEMS oscillators are oscillating with the same frequency. Moreover, the phase difference between the coupled oscillators is found to be approximately 180^{o} demonstrating locking to a common oscillation frequency with a constant phase difference. Moreover, the relative phase fluctuation is clearly reduced when the oscillators are coupled as compared to the case when the oscillators are uncoupled.

Measurements of phase noise and long-term oscillator stability before and after coupling are ongoing.

IV. Conclusion

We propose the use of synchronization to enhance output signal power and enhance the signal-to-noise performance for coupled MEMS oscillators. The synchronized response of two coupled MEMS oscillators is demonstrated wherein the output amplitudes and signal-to-noise ratios of the coupled oscillators are clearly enhanced relative to the case before coupling. In addition, a reduction in relative phase fluctuations is observed for two coupled MEMS oscillators with the two oscillators locking 180^{o} out of phase. These initial results are promising and demonstrate that the phenomenon of synchronization can be extended to coupled MEMS oscillators resulting in improvements in oscillator noise performance. Future work includes more detailed investigations of noise and long-term stability in synchronized MEMS oscillators.

Acknowledgment

We thank Prof Jim Woodhouse for useful discussions. This work was funded in part by the UK-India Education and Research Initiative (UKIERI) grant SA06-250.

References

[1] Y. W. Lin, S. S. Li, Z. Ren, and C. T.-C. Nguyen, "Low phase noise array-composite micromechanical wine-glass disk oscillator", in Tech. Dig., IEEE, IEDM, Washington DC, pp. 287-290, Dec 2005.

[2] J. R. Vig and Y. Kim, "Noise in microelectromechanical system resonators," IEEE Trans. Ultrason. Ferroelect. Freq. Contr., vol. 46, no. 6, pp. 1558-1565, Nov. 1999.

[3] B. Razavi, "A study of phase noise in CMOS oscillators," IEEE J. Solid-State Circuits, vol. 31, no. 3, pp. 331-343, Mar 1996.

[4] A. Pikovsky, M. Rosenblum, J. Kurths and R.C. Hilborn, Synchronization: a universal concept in nonlinear sciences. Cambridge University Press, 2001, pp. 1-3.

[5] M. Bennett, MF. Schatz and H. Rockwood and K. Wiesenfeld, "Huygens's clocks," Proc. R. Society A, pp. 563-579, Aug 2002.

[6] P. Thiruvenkatanathan, J. Yan, J. Woodhouse, and A. A. Seshia, "Enhancing parametric sensitivity in electrically coupled MEMS resonators," J. Microlectromech. Syst., vol. 18, pp. 1077-1086, Oct 2009.

Monitoring of Opto-Electronic Oscillator Fiber Delay by a Probe-Signal Method

Wen-Hung Tseng[1&2], and Kai-Ming Feng[2]
[1]Telecommunication Laboratories, Chunghwa Telecom
Taoyuan, Taiwan
[2]Institute of Photonics Technologies, National Tsing Hua University
Hsinchu, Taiwan
whtseng@cht.com.tw

***Abstract*—Opto-electronic oscillators (OEOs) are based on a long optical fiber that acts as a high *Q* cavity. However, a long fiber is sensitive to the temperature fluctuation. The aim of this paper is to study the possibility of stabilizing the OEO cavity by employing the technology of radio frequency transfer over fiber links. We propose a scheme of a dual-loop OEO with a probe signal and experimentally demonstrate that the probe signal is useful for monitoring the fiber delay without observable interference. This paper will simulate the adjusting of the OEO optical cavity by a variable optical delay module and discuss its best long-term stability.**

I. INTRODUCTION

Ultra-stable frequency transfer over an optical fiber has been a state of art technology in the recent years [1]-[3]. By compensating the delay variation according to the phase information of the round-trip signal, a very stable optical delay line could be obtained. A radio frequency (RF) transfer using amplitude modulation of the laser carrier has demonstrated a frequency stability of the 10^{-15} level at 1-s averaging time and below 10^{-18} at 1 day [4].

An ultra-stable optical delay line is vey important component for many kinds of optical systems, such as the optoelectronic oscillator (OEO). This oscillator, based on optical fibers to act as a high quality factor (*Q*) cavity, is capable of generating radio-frequencies with very high spectral purity [5]. Very low phase noise of variant OEO architectures have been demonstrated, such as the dual-loop OEO [6][7], coupled OEO [8], and OEOs with photonic resonator-based filters [9]. However, the long-term instability of OEO is still a challenge due to its sensitive to the surrounding temperature variations. It is still a challenge for OEO to reach a good long-term frequency stability without a thermally stabilized chamber [10].

The aim of the paper is to study whether the technologies of RF frequency transfer can be adequate for stabilizing the optical fiber loop of OEO. We experimentally demonstrate the use of a probe signal for monitoring the fiber delay and evaluate the performance of OEO with stabilizing its delay.

II. BACKGROUND

There are two fiber paths with different time delays in the dual-loop OEO as illustrated in [6]. If we assume the delay of the long fiber loop is τ_1, and that of the short fiber loop is τ_2, the oscillation frequency is given by

$$f_o = \frac{m}{\tau_1} = \frac{k}{\tau_2}, \tag{1}$$

where *m* and *k* are integers. The mode spacing of the OEO is then determined by the short loop delay τ_2, and the phase noise is determined by the long loop delay τ_1. This leads to the oscillation signal having narrow spectral linewidth with the suppressed unwanted modes. Because of the frequency pulling effect between a narrow spectral width of the long loop and a broader one of the short loop, a small tuning of oscillation frequency is chiefly determined by the long loop delay variations unless a frequency hopping happens [6][11].

If the long loop delay varies $\Delta\tau_1$, the oscillator frequency is then tuned by an amount Δf_o, and the relation is expressed as

$$\frac{\Delta f_o}{f_o} = \frac{-\Delta\tau_1}{\tau_1}. \tag{2}$$

The frequency stability is then negatively related to the stability of long loop delay. If there is a possible to monitor the loop delay, one can stabilize the fiber delay by the techniques of RF frequency transfer. An approach to precisely monitoring the long fiber delay is demonstrated as follows.

III. Experimental Setup

Fig. 1 shows the scheme of a dual-loop OEO with a probe signal. These devices are commonly available in time keeping institutes for the purpose of time and frequency transfer. In order to add a probe signal, a RF coupler is used in front of the optical modulation. Our light source is a 1310 nm single-mode laser. After the direct modulation, the optical signal is divided into two loops by a 50/50 optical coupler. The longer one is a 10 km spool of standard single-mode fiber (SMF), and the short one is a 30 m SMF pigtail. An additional RF power splitter, followed after the photodetector of the long fiber, provides the output port of the probe signal. Since the loss of the short fiber is much small than the long one, the additional attenuator is essential to balance the loop gains. After combining the signals from these two loops, two cascaded RF amplifiers provide a high gain to compensate the loss of components for reaching the oscillation condition. A RF band-pass filter, having a 4 MHz bandwidth centered at 70 MHz, selects an oscillation around the intermediate frequency. Finally, a half of the RF power is taken off by a power splitter and used as an OEO output. The other part of RF signal is combined with the probe tone and sent back to form the feedback loops. The output is measured by a spectrum analyzer and a frequency counter (SR 620).

A pseudo-random noise (PRN) encoded signal is employed to monitor the long fiber loop delay of the OEO. The PRN signal is based on the spread-spectrum techniques, resulting in a signal with a wider bandwidth and a lower spectral power density. The probe signal is generated by a modem (TimeTech SATRE), usually used for two-way satellite time and frequency transfer (TWSTFT) [12]. The modem acts as a time interval counter to measure the propagation time when the probe signal passing through the long fiber and returning back to the modem.

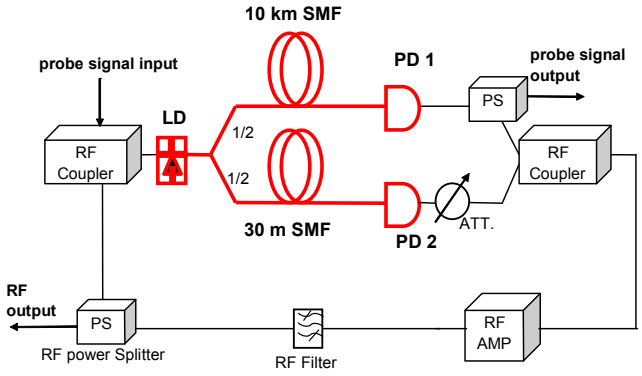


Figure 1. Experimental setup. LD: direct modulation laser diode; PD: photo detector; ATT: variable power attenuator; SMF: single mode fiber.

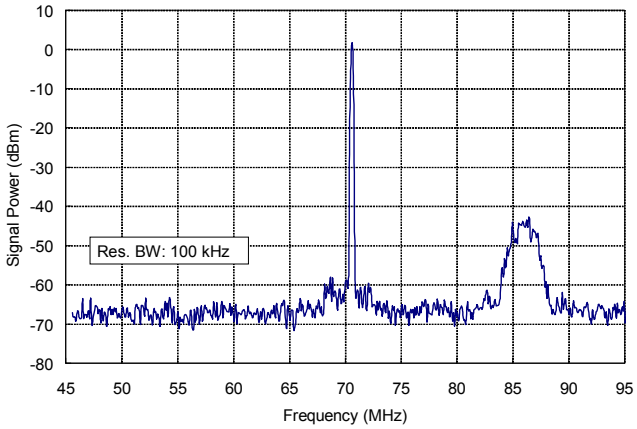


Figure 2. RF spectrum of OEO tone at 70.6 MHz and probe signal centered at 86 MHz.

To avoid the interference between the oscillation and the probe signal, the frequency of the probe signal should lie outside of the bandwidth of the band-pass filter, and far away the oscillation frequency. Moreover, its power intensity must be much lower than that of the oscillation to avoid the beats between them. However, the probe signal should maintain its carrier-to-noise ratio for the monitoring precision. Fig. 2 shows the spectrum of the RF output. The oscillation frequency is at 70.6 MHz with a peak power of 1.7 dBm. The center frequency of the probe signal is placed at 86 MHz, which is 15.4 MHz far from the oscillation. While the chip rate of the probe signal is 2.5 MHz, the actual bandwidth larger than -60 dBm is 4 MHz. The peak power occurred near the 86 MHz is -42.3 dBm, which is about 44 dB lower than the oscillation signal. The carrier-to-noise ratio of the received probe signal reaches 69 dB/Hz. In the observation of the frequency counter, the existence of this probe does not affect the stability for the period larger than 1 s. In addition, the probe signal can be easily eliminated by an additional filter following the RF output.

IV. Results

For focusing on the long term behavior, the frequency measurements were recorded every second by counter, and there was no frequency hopping during our observation. Meanwhile, the delay of the long fiber was monitored every second by the probe signal from a modem. To reduce the short-term measurement noise of the monitored data, we used the 600-s moving averages of the monitored data to represent the delay of the fiber in this study. Fig. 3 shows the measured frequency variation and the negative value of the monitored delay variation. It is clear that the frequency varied with the delay variation.

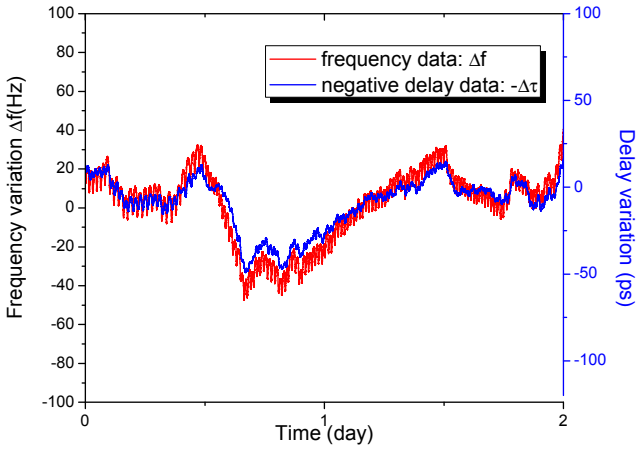


Figure 3. Frequency variation of the OEO versus the negative delay data of the long fiber loop.

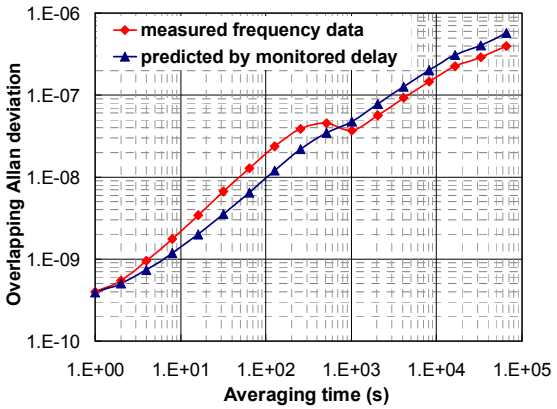


Figure 4. Measured and predicted Allan deviation of the OEO.

Fig. 4 shows the Allan deviation results. The measured frequency fluctuation is represented by red line. The frequency stability is 3.9×10^{-10} at a 1-s averaging time. Due to the thermal drift, the curve has no descending or flat part, and ascends to 10^{-7} after 4000-s averaging times. The other blue curve is the predicted results according to the delay data. As the relationship described by (2), we treated the negative fractional delay fluctuation (i.e. $-\Delta\tau/\tau$) as the fractional frequency data, and calculated their frequency stability. These two curves have the same tendency and were tied to a similar level. This means that the predicted results are in certain agreement with the measured frequency data. Nevertheless, the measured data in the short term were a little worse than the predicted one because the delay of the non-monitored filter rapidly responded to the temperature fluctuation. However, for the long term (τ > 700 s), the stability of the measured frequency is lower than the predicted value.

V. Performance Evaluation of Active Tuning

According to the monitored fiber delay data, we are able to stabilize the fiber delay by compensating its variation. To carry out the active optical compensation, the long fiber delay would be controlled by inserting an available fiber device, such like a motorized variable optical delay line and a piezoelectric fiber stretcher. Based on the specification of commercial devices, we simulated the delay of active tuning fiber loop of an OEO. The delay varying speeds are classified as none, ±0.01 ps/sec, and ±0.05 ps/sec; and the control dynamic range is 600 ps. We steered the delay line according to the difference between the typical value and the value of the past 600-s moving average of the monitored data. Finally, we got a simulated result of a stabilized loop delay.

Fig. 5 shows the Allan deviations of the monitored delay and the simulated stable delay after the tuning. Since the delay data were actually the time interval readings of the phase differences, the results were calculated as the phase data. The long-term stability of the delay of the free running fiber loop was almost limited to the level of 10^{-15}. Through the active compensation, the stability of the fiber delay could be below 10^{-18} after 3×10^{4} s averaging.

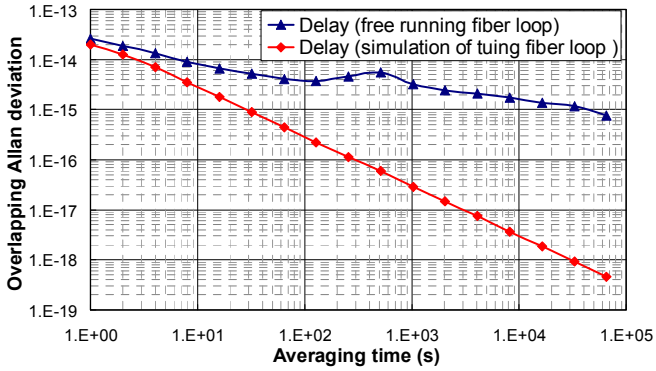


Figure 5. Comparison between delay stabilities of the measured free-running fiber loop and simulated tuning fiber loop.

By reading the stabilized delay data as frequency data, we could evaluate the frequency stability for a cavity stabilized OEO. Fig. 6 shows the evaluated results of the simulated delay data. The long-term frequency stability of the OEO could be kept within 10^{-10} after 600-s averaging times. Compared with the free running OEO, the results show that the frequency stability was reduced from 10^{-7} to 3×10^{-11} at a 10^{4}-s averaging time. This would be a great improvement.

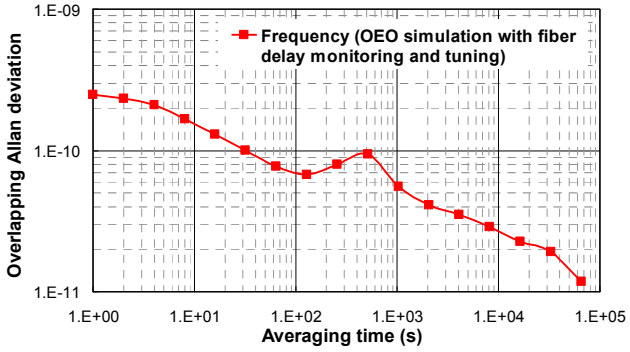


Figure 6. Frequency stability of an OEO from the simulation of tuning fiber loop delay.

VI. CONCLUSION

We have demonstrated a probe-signal method for monitoring the fiber delay of a dual-loop OEO. Since the probe signal can keep a good carrier-to-noise ratio with a much lower power level than that of the oscillated signal, there is no observable interference in the oscillation loop.

The measurement noise of the monitored data is not precise enough that we have to use their 600-s moving averages to represent the fiber delay. In the future work, we plan to carry out the active stabilizing of a fiber delay loop according to the monitored data. Then, more precise monitoring is necessary for real-time phase correction. Through the better selection of components, we may employ a higher chip-rate PRN signal, which requires a broader bandwidth, to increase the precision of monitored data.

REFERENCES

[1] C. Daussy et al., "Long-distance frequency dissemination with a resolution of 10^{-17}," Phys. Rev. Lett. vol. 94, no. 20, 203904, 2005.

[2] H. Jiang et al., "Long-distance frequency transfer over an urban fiber link using optical phase stabilization," J. Opt. Soc. Am. B, vol. 25, pp. 2029–2035, 2008.

[3] M. Kumagai, M. Fujieda, S. Nagano, M. Hosokawa, "Stable radio frequency transfer in 114 km urban optical fiber link," Opt. Lett. vol. 34, no. 19, pp. 2949–2951, 2009.

[4] O. Lopez, A. Amy-Klein, M. Lours, C. Chardonnet and G. Santarelli, "High-resolution microwave frequency dissemination on an 86-km urban optical link," Appl. Phys. B, vol. 98, no. 4, pp. 723–727, 2010.

[5] X. S Yao and L. Maleki, "Optoelectronic microwave oscillator," J. Opt. Soc. B, vol. 13, pp. 1725–1735, 1996.

[6] X. S. Yao, L. Maleki, Y. Ji, G. Lutes, M. Tu., "A Dual-Loop Opto-Electronic Oscillator," TMO Progress Report, 42-135, 1998, http://ipnpr.jpl.nasa.gov/progress_report/42-135/135C.pdf.

[7] W. Zhou and G. Blasche, "Injection-locked dual opto-electronic oscillator with ultra-low phase noise and ultra-low spurious level," IEEE Microwave Trasn. Theory Tech., vol. 53, pp. 929–933, 2005.

[8] N. Yu, E. Salik, L. Maleki, "Ultralow-noise mode-locked laser with coupled optoelectronic oscillator configuration," Opt. Lett., vol. 30, pp. 1231–1233, 2005.

[9] A. A. Savchenkov, V. S. Ilchenko, J. Byrd, W. Liang, D. Eliyahu, A. B. Matsko, D. Seidel, L. Maleki, "Whispering-gallery mode based opto-electronic oscillators," in *Proceedings of 2010 IEEE International Frequency Control Symposium*, pp. 554–557, 2010.

[10] M. Kaba et al., "Improving thermal stability of opto-electronic oscillators," IEEE Microw. Mag., vol. 7, no. 4, pp. 38–47, 2006.

[11] S. Romisch, J. Kitching, E. Ferre-Pikal, L. Hollberg, F. L. Walls, "Performance evaluation of an optoelectronic oscillator," IEEE Trans. Ultrason. Ferroelectr. Freq. Control., vol. 47, pp. 1159–1165, 2000.

[12] A. Bauch, D. Piester, M. Fujieda, W. Lewandowski, "Directive for operational use and data handling in two-way satellite time and frequency transfer (TWSTFT)," Rapport BIPM-2011/01, http://www.bipm.org/utils/common/pdf/rapportBIPM/2011/01.pdf.

Some considerations on acoustic resonator phase noise modeling and recent short-term stability experimental results

F. Sthal[#], J. Imbaud[#], M. Devel[*],
P. Salzenstein[#], R. Bourquin[#]
[#]Frequency and Time Department
[*]Micro-Nano Sciences and Systems Department
FEMTO-ST Institute
ENSMM, CNRS, UFC, UTBM
Besançon, France
fsthal@ens2m.fr

G. Cibiel
Microwave and Time-Frequency Department
CNES (Centre National d'Etudes Spatiales – French National Space Study Center)
Toulouse, France

***Abstract*— This paper presents a critical review of different acoustic resonator models exposed in the past, in the light of recent experimental results. Among the few models attempting to explain the 1/*f* behavior of the quartz crystal resonator, the *1/f* quantum noise theory of P. Handel and the thermal boson theory of M. Planat, are recalled. The latest published experimental results are analyzed taking into account these models. In the case of Handel's model, the noise depends on the resonator volume under the electrodes and the unloaded quality factor of the resonator. Handel's model is supposed to give an order of magnitude of the lowest possible noise level, but now some experimental results are under its prediction. An approach using the vibrating volume of the resonator is proposed to renew the lowest noise level estimation and found to be compatible with the latest experimental results.**

I. INTRODUCTION

High-precision quartz oscillators are mainly used in telecommunications, metrology and space applications such as local oscillators in satellites. The short-term stability of such oscillators is essential.

In 1994, J. Norton [1] has measured the best result of short-term stability of a quartz crystal resonator. It was a 5 MHz SC cut mounted in a BVA structure. The result in term of Allan standard deviation was $3.74 \cdot 10^{-14}$ for τ equals to 10 s. In 1999, at 10 MHz, R.J. Besson [2] measured two BVA quartz crystal oscillators with $5 \cdot 10^{-14}$ Allan standard deviation. Then, for ten years, there had been no significant improvement of quartz oscillators in the short term stability domain and no real solution had been found to improve the phase noise of BAW resonators. J. Norton's result represented the best acoustic resonator noise performance that could be found in the literature, pretty close to the lowest theoretical estimates. However, recently, better experimental results [3-4] have shown that quartz crystal oscillators may still have not said their last word in the short-term stability domain.

The oscillator limitation is theoretically imposed by the resonator. However, few models attempt to explain the *1/f* behavior of the quartz crystal resonator. In this paper, we will briefly recall the *1/f* quantum noise theory of P. Handel and the thermal boson theory of M. Planat in order to compare the latest published experimental results to the lower limitations computed with these two models. In the case of Handel's model, the noise depends on the resonator volume under the electrodes and the unloaded quality factor (Q) of the resonator. It gives an order of magnitude of the noise, but now some experimental results are below its prediction. In this paper, we will show that a lower theoretical estimate can nonetheless be extracted from Handel's model provided one uses the vibrating volume of the resonator instead of the volume under the electrodes.

II. MODELS

The following part presents a summary of the two models trying to elucidate the physical origin of the *1/f* flicker noise of a quartz crystal resonator with a focus on the origin of the *1/f* formula. We note, that both models can be seen as the two opposite sides of the fluctuation-dissipation theorem [5], since Planat studies the energy fluctuations around statistical equilibrium of a population of massless degenerate thermal bosons, whereas Handel proposes an origin for the dissipation occurring when the oscillator comes back to its mode after a small perturbation.

A. Planat's Model

In 2003, M. Planat proposed a statistical thermodynamics based model to explain the *1/f* noise in quartz crystal resonators [6]. In this paper, Planat considered the case of a population of massless bosons with a degenerate quantum harmonic oscillator energy spectrum which predict a hyperbolic fluctuation spectrum at low frequencies. According to the equipartition law of statistical mechanics, the partition function is introduced using a degeneracy parameter of the

978-1-61284-110-6/11/$26.00 ©2011 IEEE

energy level nhf with n the number of bosons, h the Planck constant and f the frequency. The free energy F may be approximated in the low frequency part of the spectrum by:

$$\frac{F}{kT} \approx -\frac{\pi^2}{6x} - \frac{1}{2} ln\left(\frac{x}{2\pi}\right) + \frac{x}{24} \qquad (1)$$

with $x = hf / kT$

Subsequently, the exposed model keeps only the first term of (1).

According to the Einstein's conventional approach, the energy fluctuations of a system in equilibrium within a larger system of temperature T are:

$$\varepsilon^2 = \left\langle \left(E - \langle E \rangle\right)^2 \right\rangle = kT^2 \frac{\partial \langle E \rangle}{\partial T} \qquad (2)$$

One can reformulate this relation for the fluctuations of the energy dE = u(f,T)df in the bandwidth df:

$$d\varepsilon^2 = kT^2 \frac{\partial u}{\partial T} df = S_u(f) df \qquad (3)$$

with u the energy spectral density and $S_u(f)$ the power spectral density of the fluctuations of u.

Thus, energy fluctuations in the bandwidth df are given by:

$$\left[\frac{S_u(f)}{u^2}\right]_{LF} = \frac{9hc_{ph}^3}{4\pi^3 V} \frac{1}{kTf} = \frac{A_{ph}}{Vf} \qquad (4)$$

With $3 / c_{ph}^3 = 1 / c_C^3 + 1 / c_B^3 + 1 / c_A^3$, c_A, c_B and c_C the velocities of the longitudinal mode A and the shear modes B and C, respectively.

Considering [7-8], the fractional frequency fluctuations in a quartz crystal resonator are:

$$S_y(f) = \frac{S_\omega(f)}{\omega^2} = \frac{1}{Q^4} \frac{A_{ph}}{Vf} = \frac{h_{-1}}{f} \qquad (5)$$

(5) has been used to give an order of magnitude of the power spectral density of the fractional frequency fluctuations of a 5 MHz 5th overtone (OT 5), AT-cut quartz crystal resonator.

B. Handel's Model

P. H. Handel proposes in [9] a summary of his *1/f* quantum noise theory. According to him, the general origin of the intrinsic *1/f* noise is due to a quantum self-interference process caused by very small spontaneous bremsstrahlung energy losses. The rate Γ of phonon interactions which remove phonons from the main quartz resonator mode is modulated by the quantum *1/f* effect. The time-derivative of the polarization vector P of the quartz crystal is affected. According to bremsstrahlung, when the phonon is removed, a constant energy is radiated away per unit frequency interval at any frequency, its expression is:

$$E = \frac{1}{(4\pi\varepsilon_0)} \frac{\left(\Delta\dot{P}\right)^2}{3c^3} \qquad (6)$$

With $\Delta\dot{P}$, the variation of current density of the quartz crystal.

Then the number of photon emitted as bremsstrahlung of frequency f, per Hertz is given by dividing (6) by the energy of photon hf:

$$Z = \frac{2\alpha\left(\Delta\dot{P}\right)^2}{3\pi f e^2 c^2} \qquad (7)$$

With $\alpha = (1 / 4\pi\varepsilon_0)(2\pi e^2 / hc) = 1 / 137$, the Sommerfeld's fine structure constant.

Thus the one side spectrum is inversely proportional to f:

$$S_{\delta\Gamma / \Gamma}(f) = \frac{4\alpha\left(\Delta\dot{P}\right)^2}{3\pi f e^2 c^2} \qquad (8)$$

If the main resonant mode is described as a forced harmonic oscillator with losses then the quantum *1/f* fluctuations are present in the loss coefficient Γ of the following equation:

$$\frac{d^2x}{dt^2} + 2\Gamma \frac{dx}{dt} + \omega_0^2 x = F(t) \qquad (9)$$

The quality factor of the resonator is defined as:

$$Q = \omega_r / 2\Gamma \qquad (10)$$

The resonant frequency is given by:

$$\omega_r^2 = \omega_0^2 - 2\Gamma^2 \qquad (11)$$

Thus

$$S_y(f) = \frac{1}{4Q^4} S_{\delta\Gamma / \Gamma}(f) \qquad (12)$$

Where $y = \Delta f / f_0$ is the relative frequency fluctuation.

The vibrational energy of the crystal can be written as:

$$W = n\hbar\langle\omega\rangle = 2\left(\frac{Nm}{2}\right)\left(\frac{dx}{dt}\right)^2 = \left(\frac{m}{Ne^2}\right)\chi^2\left(\dot{P}\right)^2 \qquad (13)$$

With m the reduced mass of the elementary oscillating dipoles, e their charge, N their number in the crystal, χ a polarization constant, <ω> the average circular frequency of a thermal phonon interacting with phonons in the main resonator mode and n the average number of phonons in that typical thermal phonon mode.

Thus with:

$$\left(\Delta\dot{P}\right)^2 = \frac{N\hbar\langle\omega\rangle}{n} \frac{e^2}{4\chi^2} \qquad (14)$$

We can write:

$$S_y(f) = \frac{\beta}{f}\left(\frac{V}{Q^4}\right) \qquad (15)$$

With: $\beta = n_0 \frac{\alpha\hbar\langle\omega\rangle}{12n\pi mc^2\chi^2} = 1/cm^3$, n_0 the number of dipoles per volume unit and V in cm^3.

III. ACOUSTIC VOLUME ESTIMATION

The acoustic volume estimation is carried out using Tiersten's definitions as *e. g.* in [10].

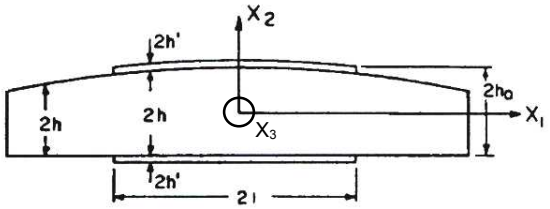


Figure 1. Geomeric definition of a plano-convex resonator

Due to the plano-convex shape, the thickness of the resonator $2h$ is given by the following expression (Fig. 1):

$$2h \cong 2h_0\left[1 - \frac{(x_1^2 + x_3^2)}{4Rh_0}\right] \qquad (17)$$

The amplitude of the mechanical displacement in the resonator is:

$$u_{nmp} = A \sin\frac{n\pi x_2}{2h} H_m\left(\sqrt{\alpha_n}x_1\right)H_p\left(\sqrt{\beta_n}x_3\right) \cdot e^{-\alpha_n(x_1^2/2)}e^{-\beta_n(x_3^2/2)}e^{j\omega_{nmp}t} \qquad (18)$$

with:

$$\alpha_n^2 = \frac{n^2\pi^2\hat{c}}{8Rh_0^3M_n} \text{ and } \beta_n^2 = \frac{n^2\pi^2\hat{c}}{8Rh_0^3P_n} \qquad (19)$$

Where n = 1, 3, 5, … the overtone number, m, p = 0, 2, 4, …, label the different mode shapes in the plane of the resonator, H_m *and* H_p are Hermit polynomials, R is the curvature radius, $2h_0$ is the thickness of the resonator center, M_n and P_n are the dispersion constants and $\hat{c}$ is the effective elastic constant associated to the propagation of the (C) mode.

For the 3rd overtone (OT 3), we have:

$$u_{300} = A \sin\frac{3\pi x_2}{2h} e^{-\alpha_3(x_1^2/2)}e^{-\beta_3(x_3^2/2)}e^{j\omega_{300}t} \qquad (20)$$

$$\alpha_3^2 = \frac{9\pi^2\hat{c}}{8Rh_0^3M_3} \text{ and } \beta_3^2 = \frac{9\pi^2\hat{c}}{8Rh_0^3P_3} \qquad (21)$$

In the case of SC-cut quartz resonator [11], the cut is defined by a doubly rotation (φ = 22°45', θ = 34°) and in N/m^2:

$$\hat{c} \cong 34.6 \cdot 10^9, M_3 \cong 57 \cdot 10^9, P_3 \cong 67 \cdot 10^9$$

The energy of the acoustic wave is given by:

$$W_1 \cong \frac{1}{2}\hat{c}\left(\frac{3\pi}{2h}\right)^2 \iiint_V \left[\cos\left(\frac{3\pi}{2h}x_2\right)e^{-\alpha_3\frac{x_1^2}{2}}e^{-\beta_3\frac{x_3^2}{2}}\right]^2 dV \qquad (22)$$

To define the acoustic volume, we can consider the vibration amplitude which remains constant over the surface S_{eq}, then we compute a cylindrical volume that has the same energy than the acoustic wave.

$$W_2 \cong \frac{1}{2}\hat{c}\left(\frac{3\pi}{2h}\right)^2 \int\left[\cos\left(\frac{n\pi}{2h}x_2\right)dx_2\right]^2 \times S_{eq} \qquad (23)$$

Thus:

$$W_1 = W_2 \Rightarrow S_{eq} = \iint_S \left[e^{-\alpha_3\frac{x_1^2}{2}}e^{-\beta_3\frac{x_3^2}{2}}\right]^2 dx_1dx_3 \qquad (24)$$

Consequently, the acoustic volume is defined by:

$$V_{ac} = S_{eq}2h_0 \qquad (25)$$

IV. APPLICATION OF THE MODELS TO EXPERIMENTAL MEASUREMENTS

As seen in [4], the best measured resonators are the 5 MHz BVA SC-cut, OT 3, resonator from Oscilloquartz (Neuchatel, Switzerland). In [12], a succinct description of the resonator is given which allows us to compute an order of magnitude of our approach. The diameter of the resonant part of the resonator is around 20 mm, the electrodes diameter D is about 11 mm. The thickness of the resonator is chosen to vibrate on the OT 3 at 5 MHz. The acoustic velocity for the thickness shear C-mode of a SC-cut in quartz crystal is about 3610 m/s. Thus, the thickness at the center of the resonator $2h_0$ is given by 3λ/2 formula and is approximately equal to 1.15 mm.

The energy trapping is quasi-independent of the electrodes dimension, in contoured resonators. It influences the motional resistance and the static capacitor. The displacement can be computed from the center of the resonator to the external edge of the resonator. The radius of curvature is one of the most important parameter. Usually, sufficient energy trapping is obtained when the lateral boundary has no effect on the vibration. Thus, we will consider an attenuation of the mechanical displacement of about 10^{-7} between the center and the boundary of the resonator. This trapping corresponds to a curvature radius R of about 290 mm (fig. 2).

The Q factor of this kind of resonator can be found from the dissipative matrix η_{ij} of the quartz crystal [13]. These constants must be expressed in the SC-cut orientation. Then, the values are considered as the imaginary part of the effective elastic coefficients. Resolution of the Christoffel equations [11] gives the complex eigenvalue $\hat{c} + i\hat{\eta}$. Then, an order of magnitude of the Q factor can be obtained using the following equation:

$$Q = \frac{\hat{c}}{\hat{\eta} 2\pi f} \qquad (26)$$

In our case, $\hat{\eta} \cong 3.95 \cdot 10^{-4}\, Ns/m^2$. A SC-cut quartz crystal resonator, at 5 MHz, OT 3, gives a maximum of quality factor of about $2.8 \cdot 10^6$. This value corresponds with the one given in [3].

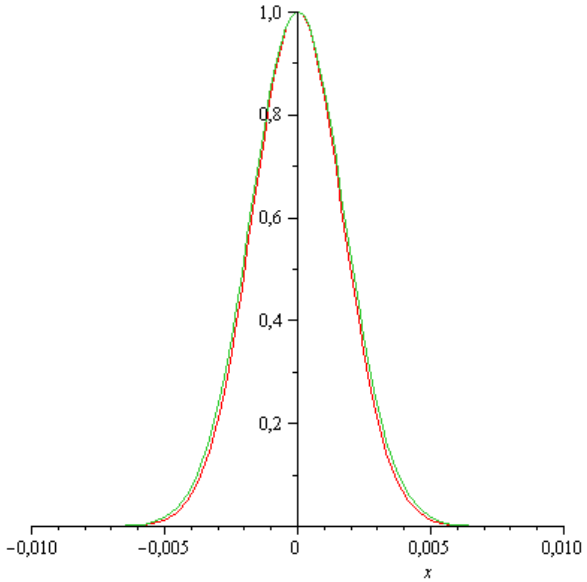


Figure 2. Energy trapping of a plano-convex resonator, SC-cut 5 MHz, OT3, R = 290 mm. (red curve: along x_3 axis, green curve: along x_1 axis).

Consequently, we can compute the motional resistor R_{mot} from the motional capacitance C_{mot} using:

$$R_{mot} = \frac{1}{QC_{mot} 2\pi f} = \frac{L_{mot} 2\pi f}{Q} \qquad (27)$$

In the case of OT 3, C_{mot} is given by [10]:

$$C_{mot} = \frac{64\hat{e}_{26}^2 \left(\int_0^{\frac{D}{2}} e^{-\alpha_3 \frac{x_1^2}{2}} dx_1 \int_0^{\frac{D}{2}} e^{-\beta_3 \frac{x_3^2}{2}} dx_3 \right)^2 \sqrt{\alpha_3}\sqrt{\beta_3}}{9\pi^3 h_0 \hat{c}} \qquad (28)$$

With D the electrode diameter, $\hat{e}_{26} \cong -0.0576\ C/m^2$ the effective piezoelectric constant.

The computing value of C_{mot} is around 280 fF. We obtain R_{mot} equals to 42 Ω, L_{mot} is about 3.6 H. These results are closed to the measured value in [3].

The static capacitor C_0 can be also computed with the following formula [9]:

$$C_0 = \frac{\pi D^2}{4}\left(1 + \frac{(D/2)^2}{8Rh_0}\right)\frac{\varepsilon_{22} + \frac{\hat{e}_{26}^2}{\hat{c}}}{2h_0} \qquad (29)$$

With $\varepsilon_{22} = 39.78 \cdot 10^{-12} F/m$ the permittivity constant in SC-cut.

We have C_0 around 3.4 pF.

Short-term stability is usually given in the time domain with an Allan variance. The Allan variance is computed from the Fourier frequency and from the power spectral density of the frequency fluctuations [14]:

$$\sigma_y^2(\tau) = \int_0^\infty S_y(f) \frac{2\sin^4 \pi f\tau}{(\pi f\tau)^2} df \qquad (30)$$

In the case of the flicker frequency noise (which is the resonator noise), relation between the Allan standard deviation and the power spectral density (PSD) of relative frequency fluctuations $S_y(f)$, is given by [15]:

$$\sigma_{y\_floor} = \sqrt{2\ln(2) S_y(1Hz)} \qquad (31)$$

Table 1 presents the comparison between measurements and the results proposed by both models. For Planat's model, only the velocity of the C-mode of the acoustic wave is used to the computation. The result is really higher than the experiment. Using the acoustic volume in this model is not relevant because of the invert proportionality.

TABLE I. COMPARISON OF THE SHORT-TERM STABILITY IN TERM OF ALLAN STANDARD DEVIATION σ_{Y_FLOOR} OF A SC-CUT, 5 MHZ, OT 3 RESONATOR

Frequency (MHz)	5
Quality factor Q (10^6)	2.8
R_{mot} (Ω)	42
C_{mot} (fF)	280
L_{mot} (H)	3.6
C_0 (pF)	3.4
Volume under electrodes V_{elec} (cm^3)	$1.093 \cdot 10^{-1}$
Acoustic volume V_{ac} (cm^3)	$1.076 \cdot 10^{-2}$
Volume ratio	10
$\sigma_{y\_floor\_planat}$ (vol. under electrodes)	$5.2 \cdot 10^{-12}$
$\sigma_{y\_floor\_Handel}$ (vol. under electrodes)	$5.4 \cdot 10^{-14}$
$\sigma_{y\_floor\_Handel}$ (Acoustic vol.)	$1.67 \cdot 10^{-14}$
$\sigma_{y\_floor\_exp\ [1]}$ (oscillator measurement)	$3.75 \cdot 10^{-14}$
$\sigma_{y\_floor\_exp\ [4]}$ (oscillator measurement)	$2.5 \cdot 10^{-14}$

The flicker noise seems to be the limit of the short-term stability of a quartz crystal resonator. An order of magnitude of this limit is given by the previous calculus around $1.6\ 10^{-14}$ with the Handel's model using the definition of the acoustic volume.

If it was possible to eliminate this noise, the noise limit should be given by the thermal noise of the resonator. In term of Allan standard deviation, its expression is usually [16]:

$$\sigma_{y\_therm}(\tau) = \frac{1}{Q_L}\sqrt{\frac{kT}{2P\tau}} \qquad (32)$$

With a loaded Q-factor Q_L around $1.6 \cdot 10^6$ for the previous resonator and a excitation power of 60 µW, we get around T = 80°C:

$$\sigma_{y\_therm}(\tau) \approx 4 \cdot 10^{-15} \cdot \tau^{-1/2} \qquad (33)$$

At 1s, without *1/f* noise, the short-term stability limit should be $4 \cdot 10^{-15}$.

V. Conclusion

Regarding the experimental results, the *1/f* story of the quartz crystal is not over. The considered models should be the fundamental limit of the flicker noise of a resonator. Values of Handel's model are under the experimental measurements if the volume is considered as the acoustic volume. The experimental results seem to be now very close to the fundamental limit. The model validity could be proved by varying the acoustic volume while keeping the Q factor of the resonator constant.

Acknowledgment

The authors thank M. Planat for his helpful discussion and the CNES for financial support of the work.

References

[1] J. R. Norton, "Performance of ultrastable quartz oscillators using BVA resonators," Proc. European Frequency and Time Forum, Weihenstephan, Germany, March 1994, pp. 457-465.

[2] R. J. Besson, M. Mourey, S. Galliou, F. Marionnet, F. Gonzalez, P. Guillemot, R. Tjoelker, W. Diener, A. Kirk, "10 MHz Hyperstable quartz oscillators," Proc. Joint Meeting IEEE Ann. Freq. Cont. Symp. and European Frequency and Time Forum, Besançon, France, april 2009, pp. 326-330.

[3] J. Chauvin, P. Weber, J.P. Aubry, F. Lefebvre, F. Sthal, S. Galliou, E. Rubiola and X. Vacheret, "A new generation of very high stability BVA oscillators," Proc. Joint Meeting IEEE Ann. Freq. Cont. Symp. and European Frequency and Time Forum, Genova, Switzerland, June 2007, pp. 1261-1268.

[4] P. Salzenstein, A. Kuna, L. Sojdr and J. Chauvin, "Significant step in ultra high stability quartz crystal oscillators," Electron. Lett., vol. 46, no. 21, pp. 1433-1434, (2010).

[5] R. Kubo, "The fluctuation-dissipation theorem," Rep. Prog. Phys., vol. 29, pp. 255-284, (1966).

[6] M. Planat, "Thermal *1/f* noise from the theory of partitions: application to a quartz resontor, " Physica A, vol. 318, pp. 371-386, (2003).

[7] J. J. Gagnepain, J. Uebersfeld, G. Goujon, P. Handel, "Relation between 1/f noise and Q-factor in quartz resontors," Proc. IEEE Ann. Freq. Cont. Symp., Ft. Monmouth, New Jersey, May 1981, pp. 476-483.

[8] P. H. Handel, "Nature of *1/f* frequency fluctuations in quartz resonators," Solid State Electron., vol. 22, pp. 875-876, (1979).

[9] P. H. Handel, "Quantum *1/f* effect in resonant biochemical piezoelectric and MEMS sensors," IEEE Trans. on Ultrason., Ferroelec. and Freq. Cont., vol. 52, no. 9, Sept., pp. 1461-1467, (2005).

[10] D. S. Stevens, H. F. Tiersten, "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation," J. Acous. Soc. Am., vol. 79, n° 6, pp. 1811-1826, (1986).

[11] R. Bourquin, B. Dulmet, J. J. Boy, "SC-cut resonator operating in anharmonic modes with B-mode reduction," Proc. European Frequency and Time Forum, Brighton, United Kingdom, March 1996, pp. 239-243.

[12] D. Schneuwly, Y. Schwab, "Quartz clocks," 3rd Int. Telec. Sync. Forum, London, UK, Oct. 2005.

[13] J. Lamb, J. Richter, "Anisotropic acoustic attenuation with new measurements for quartz at room temperatures," Proc., Roy. Soc. Series A, vol. 293, pp. 479-492, (1966).

[14] L.S. Cutler, C. L. Searle, "Some aspects of the theory and measurment of frequency fluctuations in frequency standards," Proc. IEEE, vol. 54, no. 2, Feb., pp. 136-154, (1966).

[15] F. Sthal, M. Mourey, F. Marionnet, W. F. Walls, "Phase noise measurements of 10 MHz BVA quartz crystal resonators," IEEE Trans. on Ultrason., Ferroelec. and Freq. Cont., vol. 47, no. 2, March, pp. 369-373, (2000).

[16] F. L. Walls, J. Vig, Phase Noise "Fundamental limits on the frequency stabilities of crystal oscillators," IEEE Trans. on Ultrason., Ferroelec. and Freq. Cont., vol. 42, no. 4, July, pp. 576-588, (1995).

Effects of Volume and Frequency Scaling in AlN Contour Mode NEMS Resonators on Oscillator Phase Noise

Matteo Rinaldi and Gianluca Piazza
Department of Electrical and Systems Engineering
University of Pennsylvania
Philadelphia, PA, USA
{rinaldim, piazza}@seas.upenn.edu

***Abstract*—This paper presents, for the first time, the analytical modeling and the experimental verification of the effects of volume and frequency scaling in AlN contour mode NEMS resonators (CMRs) on oscillator phase noise and associated limit of detection (LOD) when the device is used for sensing mass-induced frequency shifts. Both the analytical model and the experiments show that the LOD is improved by scaling the thickness of the AlN plate and by increasing the operating frequency of the device (*i.e.* implementing a high frequency resonant nano-plate). LOD values in the order of *10s of zg/μm²* are experimentally extracted.**

I. INTRODUCTION

The need to detect extremely small concentrations of analytes requires reliably measuring extremely small variations in the sensor output signal. In this perspective, gravimetric sensors, such as Quartz Crystal Microbalances (QCMs) [1-3], have a significant advantage over other sensor technologies (conductance-based sensors, Chem-FET, or optical sensors [4-6]) since they use frequency as the output variable, which is one of the physical quantities that can be monitored with the highest accuracy [7]. Nevertheless, smart, compact, portable and low cost electronic noses for multiple gas detection cannot be implemented with bulky and unintegrable QCMs. Even if QCMs have been successfully employed as gravimetric sensors thanks to their extremely high quality factors, which permit to obtain limits of mass detection in the order of few nanograms, their relatively large volume and their inability to be directly integrated on silicon render them unattractive for the fabrication of sensor arrays composed of a large number of mechanical elements.

In response to this growing demand for miniaturization, the use of micro and nano electromechanical (MEMS/NEMS) resonators for sensing applications has been extensively explored. The scaling of the device dimensions enables the fabrication of nanomechanical resonators whose mechanical resonance frequency is extremely sensitive to mass loading. Nevertheless, the greatly reduced dimensions of these devices render their transduction extremely difficult and require the use of cumbersome, complex and power inefficient read-out techniques. In addition, the enhancement in sensitivity associated to the scaling of the device dimension is in most cases offset by the reduced stability and power handling of the nanomechanical device, which result in a limited gain in limit of detection (LOD) [8, 9].

The LOD, or resolution [9], is defined as the minimum value of analyte concentration that can be detected. It depends on both device sensitivity as well as signal to noise ratio and it determines the ultimate performance of a gas sensor. For gravimetric gas sensors the LOD is strictly related to both the sensitivity of the device resonance frequency to mass adsorbed per unit area (not absolute mass) and the minimum frequency shift (induced by gas molecule adsorption) that can be resolved by the sensor readout. Although the device sensitivity is influenced exclusively by the mass and frequency of operation of the mechanical element (hence it is enhanced by scaling the device dimensions), the minimum measurable frequency shift is a function of the noise of the acoustic oscillator (*i.e.* resonant element and electronic circuitry), hence the power handling and the dimensions of the mechanical device.

Therefore, optimal sensor performance is attained by synthesizing an ultra-low mass resonant device (hence with ultra-high frequency sensitivity to mass loading per unit area) that can be efficiently excited into mechanical vibration (hence large surface area of the transducer is desirable since it facilitates efficient transduction) and that can handle input power levels high enough to guarantee very low noise performance. Suspended membranes with thickness in the nanometer range are therefore desirable. Flexural plates (instead of beams) are a good compromise [10, 11], but higher frequencies and power handling are achieved in bulk mode acoustic resonators [12-14].

In this perspective, the recently demonstrated Nanoscale AlN Contour Mode Resonant Sensor (nano-CMR-S)

This work was supported by NCMR/NSF grant no. IIS-07-15024.

technology [15] represents the best candidate for the implementation of high performance gravimetric gas sensors. In fact, this technology enables the fabrication of ultra-sensitive nanoscaled devices that can be efficiently actuated and sensed piezoelectrically directly on chip [15, 16]. The capability of this technology to maintain high Q and high coupling, k_t^2, despite the device volume reduction has been demonstrated [15, 16] and it represented the enabling feature for the demonstration of individual sensors [15, 16] and sensor arrays [17, 18], connected to compact, low power and low noise CMOS oscillator circuits for direct frequency readout.

In this work, for the first time, the noise performance of the AlN nano-CMR-S technology is studied from both analytical and experimental points of view. In particular the effects of volume and frequency scaling of the NEMS resonator on the sensor limit of detection are investigated. Both the analytical model and the experiments show that the LOD is improved by scaling the thickness of the AlN plate and by increasing the operating frequency of the device (*i.e.* implementing a high frequency resonant nano-plate). LOD values in the order of *10s of* $zg/\mu m^2$ are experimentally extracted.

II. Analytical Model

An AlN CMR-S is composed of an AlN film sandwiched between two patterned metal electrodes (Figure 1). When an AC voltage is applied across the thickness (T) of the device a contour-extensional mode of vibration is excited through the equivalent d_{31} piezoelectric coefficient of AlN.

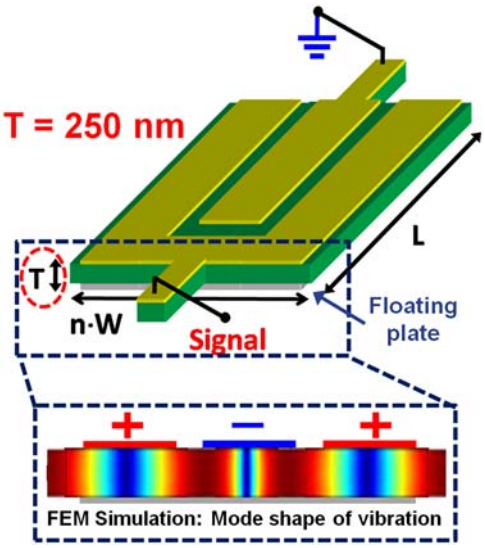


Figure 1 *Schematic representation of an AlN NEMS CMR-S: a number, n, of sub-resonators each of width, W, is mechanically coupled in order to excite a higher mode of vibration in the AlN nano-plate (f_0 is set by W) The inset shows a FEM simulation of the device mode of vibration.*

Given the equivalent mass density, ρ_{eq}, and Young's modulus, E_{eq}, of the material stack (AlN and metal electrodes) that forms the resonator, the center frequency, f_0, of this laterally vibrating mechanical structure, is univocally set by the periodic pattern of width W formed by the metal electrodes on the AlN plate [19]. The other two geometrical dimensions, thickness, T, and length, L, set the equivalent electrical impedance of the resonator [19] and can be designed independently of the desired resonance frequency.

The sensitivity, S, to mass per unit area of a CMR-S loaded on its top surface and its associated Limit of Detection, *LOD,* can be expressed as in Eq. (1) [20, 21].

$$S = -\frac{f_0}{2\rho_{eq}T} \qquad LOD = \frac{\Delta f_{\min}}{S_{CMR-S}} \tag{1}$$

Where Δf_{min} indicates the minimum frequency shift detectable by the sensor read-out. When the resonant sensor is connected to a self-sustained oscillator loop for direct frequency readout, the minimum detectable frequency shift, Δf_{min}, is limited by the noise induced frequency fluctuations of the oscillator output signal. Different types of noise (whose physical meaning and origin go beyond the focus of this paper) such as, white phase modulation (WPM), white frequency modulation (WFM), flicker frequency modulation (FFM) and random walk frequency modulation (RWFM) affect the oscillator signal and contribute to the generation of frequency fluctuations [22]. Such noise induced time domain frequency fluctuations are typically described by a statistical tool called Allan Deviation (square root of Allan Variance) [23]. The Allan Deviation (units of Hz) is commonly used to measure the time domain frequency fluctuation of clocks and oscillators for a given integration time, τ (measurement time).

For oscillator circuits composed of an acoustic resonator (such as the NEMS resonator of this work) connected in the positive feedback loop of a sustaining amplifier the Allan Deviation (hence the noise induced frequency fluctuations) can be expressed according to Eq. (2) [22].

$$\Delta f_{\min} = \sigma_y(\tau) = \left[\frac{f_0}{2\cdot Q\cdot\sqrt{P_C}}\right]\cdot\left[\alpha_{WP}\cdot\frac{1}{\tau}+\alpha_{WF}\cdot\frac{1}{\sqrt{\tau}}+\alpha_{FF}+\alpha_{RW}\sqrt{\tau}\right] \tag{2}$$

Where f_0 is the oscillator signal frequency; Q is the quality factor of the acoustic resonator; P_C is the power level at which the resonator is driven; α_{WP}, α_{WF}, α_{FF}, α_{RW} are respectively coefficients for white phase modulation, white frequency modulation, flicker frequency modulation and random walk frequency modulation; τ is the measurement time (inverse of the measurement bandwidth).

It is important to note that the value of Allan Deviation, $\sigma_Y(\tau)$, for a given measurement time, τ, can be lowered by increasing the driving power P_C. Therefore, in order to minimize the noise of the oscillator it is possible to drive the acoustic resonator with the maximum power level that it can handle. A mechanical resonator is characterized by a critical value of vibration amplitude (greatest displacement) after which non-linear phenomena occurs inducing a distorted admittance vs. frequency curve, which exhibits an infinite slope right at the point of bifurcation [24]. This critical value of vibration amplitude is associated to a critical value of drive power, P_{max}, (maximum power that the resonator can handle before bifurcation occurs) which is a function of the resonator geometry, quality factor, Q, and material properties.

For AlN Contour-Mode Resonators a simple analytical expression for the critical power, P_{max}, due to elastic non-linearity has been derived in [25]. Such expression is reported in Eq. (3) and can be employed to describe the maximum power level that the AlN nano-CMR-S can handle.

$$P_{\max} = K \cdot \frac{A}{\sqrt{\rho Q^2}} \cdot \frac{T}{W} \qquad K = \frac{2\pi^3 \sqrt{Y_0^3}}{\sqrt{3}\left|9Y_2 - 10Y_1^2\right|} \tag{3}$$

Where A is the device surface area ($n \cdot W \cdot L$); W the device frequency setting dimension; T the thickness of the material stack forming the AlN nano-CMR-S (Figure 1); ρ the device equivalent mass density; Q, the linear quality factor of the resonator; Y_0 the device equivalent Young's modulus; Y_1 and Y_2, the first-order, and second-order corrections of the strain dependent Young's modulus.

According to Eq. (3), the power handling, P_{max}, of an AlN nano-CMR-S (for a given value of device surface area, A) is directly proportional to the device resonance frequency, f_0, and inversely proportional to the square of the device quality factor, Q^2. Substituting Eq. (3) in Eq. (2) and dividing the result by the sensitivity (Eq. (1)) of the AlN nano-CMR-S it is possible to find an expression for the LOD of the AlN CMR-S technology as function of the device material and design parameters:

$$LOD(\tau) = \rho^{5/4} \cdot \sqrt{\frac{T \cdot W}{A \cdot K}} \cdot \left[\alpha_{WP} \cdot \frac{1}{\tau} + \alpha_{WF} \cdot \frac{1}{\sqrt{\tau}} + \alpha_{FF} + \alpha_{RW} \sqrt{\tau} \right] \tag{4}$$

Eq. (4) shows that the LOD of a NEMS AlN CMR-S connected to a self-sustained oscillator circuit and driven at critical power, P_{max}, improves by scaling the resonator thickness, T, and by increasing its surface area, A, and frequency of operation, f_0, ($W \propto 1/f_0$) independently of the device quality factor, Q.

Therefore, the analytical model suggests that optimum sensor performance can be achieved by designing an AlN nano-plate (with greatly reduced thickness, T, and relatively large area, A) and by exciting it into a high frequency bulk acoustic contour-extensional mode of vibration.

III. Experimental Verificaiton

In order to experimentally verify the proposed analytical model three nanoscale AlN CMR-S operating at 224 MHz, 634 MHz and 1 GHz (Table 1) were fabricated on the same chip and simultaneously connected to a single multiplexed CMOS oscillator taped-out in the ON Semiconductor 0.5 µm process.

The AlN nano-CMR-S were fabricated with a 3 mask fabrication process analogue to the one reported in [21] while the design of CMOS multiplexed oscillator circuit is described in [17, 26]. Briefly, the CMOS multiplexed oscillator is implemented by means of an inverter amplifier biased in its active region (Figure 2). Transistors *M1* and *M2* form the CMOS inverting amplifier while transistor *M3* acts as a large resistor to provide biasing of *M1* and *M2* in the active region. The single CMOS Pierce-like oscillator circuit can be connected to up to 8 nano-CMR-S by means of an equivalent number of CMOS switches operating in a time multiplexed mode and addressed by a 3 to 8 digital decoder integrated on chip.

Table 1: Resonator Geometrical Dimensions

Resonator	Width	Fingers Number	Length	Surface Area	AlN Thickness
	W (µm)	n	L (µm)	A (µm²)	T (nm)
1	15	3	200	9000	250
2	9	5	200	9000	250
3	3	15	200	9000	250

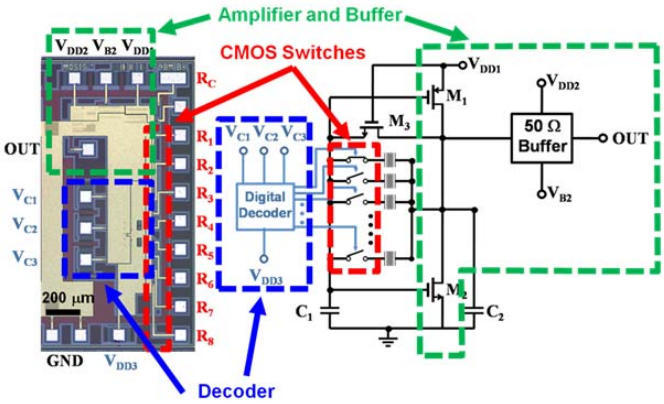


Figure 2: *Micrograph and circuit schematic of the multiplexed CMOS oscillator chip (1.05 mm²).*

The electrical responses of the three fabricated devices were first characterized in open-loop configuration using an RF probe station and the admittance curves were measured by an Agilent® N5230A Network Analyzer after performing a short-open-load (SOL) calibration on a reference substrate. The measured electrical responses of the devices were fitted to the Modified Butterworth van Dyke (MBVD) equivalent electrical circuit [27] and the device performances are reported in Table 2.

In order to characterize the multiplexed oscillator signal for each of the 3 different operating frequencies (corresponding to each of the nano-CMR-S), the 3 combinations of the 2 bit address were provided to the decoder by a Data Acquisition (DAQ) system so as to sequentially turn on each resonator. The Allan Deviation, $\sigma_Y(\tau)$, was measured by a Fluke PM6690 Frequency Counter. The bias of the multiplexed oscillator circuit was adjusted for each of the 3 operating frequencies in order to achieve an output power level, P_{out}, as close as possible to the critical power level, P_{max}, of the corresponding resonator. In order to extract the LOD, each measured value of Allan Deviation was divided by the mass sensitivity value, S, of the corresponding resonator according to Eq. (1). The experimentally extracted LOD values (at the 3 different operating frequencies) are plotted in Figure 3 for different values of measurement time, τ.

TABLE 2: RESONATOR PERFORMANCE: P_{MAX} IS ANALYTICALLY CALCULATED USING THE VALUE OF NON-LINEAR COEFFICIENT, K, EXPERIMENTALLY EXTRACTED IN [25] FOR A 250 NM THICK ALN FILM.

Resonator	Frequency	Q	k_t^2	C_0	R_m	P_{max}
	MHz		%	fF	Ω	dBm
1	224	1109	1.7	554	79	-8
2	634	852	1.67	699	26	-1.1
3	1000	650	1.5	874	13	3.5

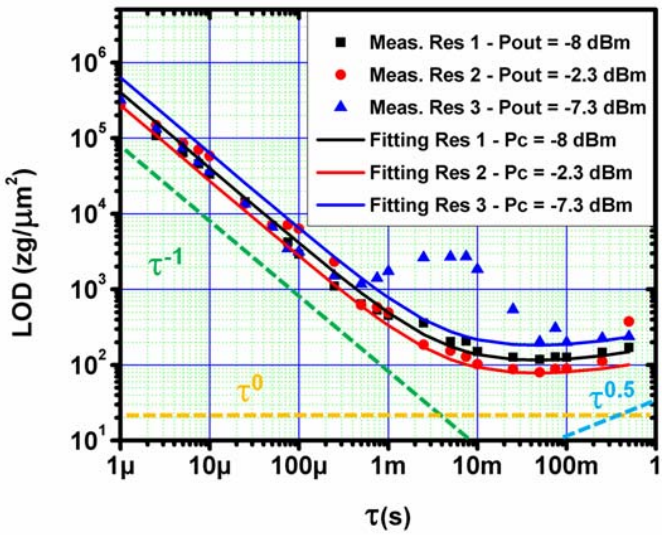


Figure 3: Experimentally extracted LOD values and equivalent analytical model fitting. The experimentally extracted LOD at 1 GHz for measurement time values, τ, comprised between 1 and 20 ms cannot be fitted by the proposed analytical model: this is repeatable, but not yet explained by the authors. It is worth noting that for very fast measurement times, below 1 ms, the relative measurement error of the instrument is larger than 10^{-7}.

The experimentally extracted LOD was fitted by the analytical model discussed in the previous section of this paper and described by Eq. (5) (note that Eq.(5) = Eq. (4) for $P_c = P_{max}$).

$$LOD = \left[\frac{\rho \cdot T}{Q \cdot \sqrt{P_C}}\right] \cdot \left[\alpha_{WP} \cdot \frac{1}{\tau} + \alpha_{WF} \cdot \frac{1}{\sqrt{\tau}} + \alpha_{FF} + \alpha_{RW}\sqrt{\tau}\right] \quad (5)$$

Where: $\alpha_{WP} = 2.2 \cdot 10^{-9}$ [$W^{0.5} \cdot s$]; $\alpha_{WF} = 9.2 \cdot 10^{-11}$ [$W^{0.5} \cdot s^{0.5}$]; $\alpha_{FF} = 4.9 \cdot 10^{-7}$ [$W^{0.5}$]; $\alpha_{RW} = 4.6 \cdot 10^{-7}$ [$W^{0.5}/s^{0.5}$]. The same set of noise coefficient values was used to fit the experimentally derived LOD for all of the 3 different operating frequencies showing the independence of the noise coefficients from the operating frequency, f_0, quality factor, Q, and drive power level of the 3 NEMS resonators employed.

It is important to note that not enough power can be provided by the currently designed circuit to reach the critical power level, P_{max}, at all the operating frequencies. The maximum output power at high operating frequency is limited by the CMOS technology currently employed. This is particularly significant for the 1 GHz case whose output power level (Figure 3) is ~11 dB smaller than the value of maximum power, P_{max} (Table 2), that the corresponding resonator can handle. Consequently the LOD at 1 GHz is higher than the one achieved at lower frequencies. Figure 4 shows a simulated LOD as function of frequency and for a fixed optimum measurement time, τ, of 50 ms. The LOD is simulated employing the same noise coefficients that fit the experimental data, but power levels, P_C, equal to critical power values, P_{max}, at all the operating frequencies. The simulation confirms that the LOD at 1 GHz would be smaller than the one achieved at lower frequencies if the resonator was driven at its critical power, P_{max}.

Therefore, the experiment shows that, as predicted by the design equation, the LOD of an AlN nano-CMR-S connected to a self-sustained oscillator circuit and driven at its critical power level, P_{max}, improves by increasing the operating frequency of the device independently of the quality factor, Q, of the resonator (the device with highest Q shows the worst LOD). A higher Q, though, would still be beneficial to lower the power consumption of the overall circuit, in agreement with the fact that Q is a concept related mainly to energy dissipation rather than noise.

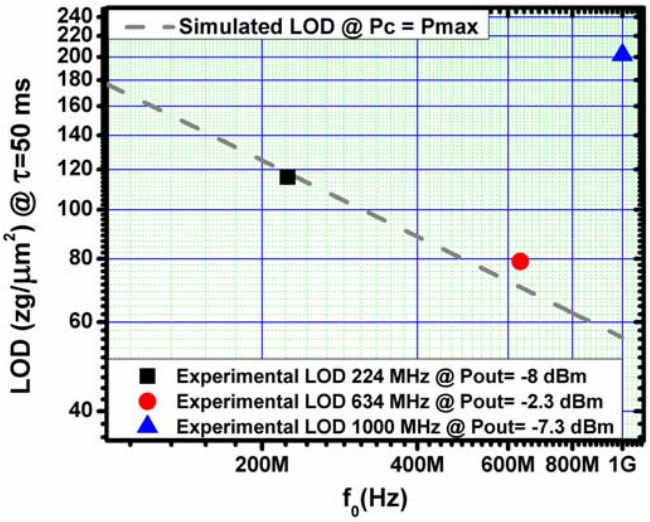


Figure 4: Experimental and simulated LOD as function of frequency and for a fixed measurement time, τ, of 50 ms. The LOD is simulated employing the same noise coefficients that fit the experimental data, but power levels, P_C, equal to critical power values, P_{max}, for each of the operating frequencies.

IV. CONCLUSION

In this paper the analytical modeling and the experimental verification of the effects of volume and frequency scaling on the LOD of AlN Contour-Mode NEMS Resonant Sensors was presented for the first time. An analytical model that relates resonator design parameters, such as resonance frequency and thickness, to the limit of detection of the resonant sensor was developed. In order to experimentally verify the proposed analytical model three nanoscale AlN CMR-S operating at 224 MHz, 634 MHz and 1 GHz were fabricated on the same chip and simultaneously connected to a single multiplexed CMOS oscillator taped-out in the ON Semiconductor 0.5 μm process. The LOD at the three different frequencies was experimentally extracted through Allan Deviation measurements and the

experimental data was fitted by the proposed analytical model. Both the analytical model and the experiments showed that the LOD is improved by scaling the thickness of the AlN plate and by increasing the operating frequency of the device (*i.e.* implementing a high frequency resonant nano-plate). LOD values in the order of *10s of* $zg/\mu m^2$ were experimentally extracted.

ACKNOWLEDGMENT

The authors wish to thank Chiara Zuniga and Dr. Chengjie Zuo for the precious discussions and the staff of the Wolf Nanofabrication Facility at The University of Pennsylvania.

REFERENCES

[1] D. S. Ballantine, R. M. White, S. J. Martin, A. J. Ricco, E. T. Zellers, G. C. Frye, and H. Wohltjen, *Acoustic Wave Sensor Theory, Design and Physico-Chemical Applicatipons*: Academic Press, 1997.

[2] C. Di Natale, A. Macagnano, F. Davide, A. D'Amico, R. Paolesse, T. Boschi, M. Faccio, and G. Ferri, "An electronic nose for food analysis," *Sensors and Actuators B, Chem.,* pp. 521-526, 1997.

[3] E. Zampetti, S. Pantalei, A. Macagnano, E. Proietti, C. Di Natale, and A. D'Amico, "Use of a multiplexed oscillator in a miniaturized electronic nose based on multichannel quartz crystal microbalance," *Sensor and Actuators B, Chemical,* vol. 131, pp. 159-166, 2008.

[4] C. Staii, A. T. Johnson, M. Chen, and A. Gelperin, "DNA-Decorated Carbon Nanotubes for Chemical Sensing," *Nano Letters,* vol. 5, pp. 1774-1778, 2005.

[5] M. Andersson, M. Holmberg, I. Lundström, A. Lloyd-Spetz, P. Mårtensson, R. Paolesse, C. Falconi, E. Proietti, C. Di Natale, and A. D'Amico, "Development of a ChemFET sensor with molecular films of porphyrins as sensitive layer," *Sensors and Actuators B: Chemical,* vol. 77, pp. 567-571, 2001.

[6] K. Matsubara, S. Kawata, and S. Minami, "Optical chemical sensor based on surface plasmon measurement," *Appl. Opt.,* vol. 27, pp. 1160-1163, 1988.

[7] J. R. Vig, R. L. Filler, and Y. Kim, "Uncooled IR imaging array based on quartz microresonators," *Journal of Microelectromechanical Systems,* vol. 5, pp. 131-137, 1996.

[8] K. L. Ekinci, Y. T. Yang, and M. L. Roukes, "Ultimate limits to inertial mass sensing based upon nanoelectromechanical systems," *Journal of Applied Physics,* vol. 95, pp. 2682-9, 2004.

[9] A. D'Amico and C. Di Natale, "A contribution on some basic definitions of sensors properties," *IEEE Sensors Journal,* vol. 1, pp. 183-190, 2001.

[10] H. J. Lee, K. K. Park, O. Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "CMUT as chemical sensor for DMMP detection," in *2008 IEEE International Frequency Control Symposium*, 2008, pp. 434-439.

[11] H. J. Lee, K. K. Park, P. Cristman, O. Oralkan, M. Kupnik, and B. T. Khuri-Yakub, "A Low-Noise Oscillator based on a Multi-Membrane CMUT for High Sensitivity Resonant Chemical Sensors," in *IEEE 22nd International Conference on Micro Electro Mechanical Systems, 2009 (MEMS 2009)* 2009, pp. 761-764.

[12] R. Abdolvand, G. K. Ho, J. Butler, and F. Ayazi, "ZNO-on-nanocrystalline diamond lateral bulk acoustic resonators," in *IEEE 20th International Conference on Micro Electro Mechanical Systems, 2007 (MEMS 2007)*, 2007, pp. 795-798.

[13] R. Ruby and P. Merchant, "Micromachined thin film bulk acoustic resonators," in *Proceedings of the 48th IEEE International Frequency Control Symposium*, 1994, pp. 135-138.

[14] R. C. Ruby, Y. Bradley, Y. Oshmyansky, A. Chien, and J. D. Larson, "Thin film bulk wave acoustic resonators (FBAR) for wireless applications," in *Proceedings of the 2001 IEEE International Ultrasonics Symposium*, 2001, pp. 813-821.

[15] M. Rinaldi, B. Duick, C. Zuniga, C. Zuo, and G. Piazza, "SS-DNA functionalized ultra-thin-film ALN Contour-mode Resonators with self-sustained oscillator for volatile organic chemical detection," in *IEEE 23rd International Conference on Micro Electro Mechanical Systems (MEMS 2010)* 2010, pp. 132-135.

[16] M. Rinaldi, C. Zuniga, C. Zuo, and G. Piazza, "GHz Range Nanoscaled AlN Contour-Mode Resonant Sensors (CMR-S) with Self-Sustained CMOS Oscillator," in *Proceedings of Solid-State Sensors, Actuators and Microsystems Workshop (Hilton Head 2010)*, Hilton Head, USA, 2010.

[17] M. Rinaldi, C. Zuniga, B. Duick, and G. Piazza, "Use of a Single Multiplexed CMOS Oscillator as Direct Frequency Read-Out for an Array of Eight AlN Contour-Mode NEMS Resonant Sensors," in *Proceedings of IEEE Sensors 2010*, 2010, pp. 2666-2670.

[18] M. Rinaldi, C. Zuniga, and G. Piazza, "ss-DNA functionalized array of ALN Contour-Mode NEMS Resonant Sensors with single CMOS multiplexed oscillator for sub-ppb detection of volatile organic chemicals," in *IEEE 24th International Conference on Micro Electro Mechanical Systems (MEMS 2011)*, 2011, pp. 976-979.

[19] G. Piazza, P. J. Stephanou, and A. P. Pisano, "Piezoelectric Aluminum Nitride Vibrating Contour-Mode MEMS Resonators," *Journal of Microelectromechanical Systems,* vol. 15, pp. 1406-1418, 2006.

[20] C. Zuniga, M. Rinaldi, S. M. Khamis, A. T. Johnson, and G. Piazza, "Nanoenabled microelectromechanical sensor for volatile organic chemical detection," *Applied Physics Letters,* vol. 94, pp. 223122-223122-3, 2009.

[21] M. Rinaldi, C. Zuniga, and G. Piazza, "Ultra-thin-film AlN contour-mode resonators for sensing applications," in *Ultrasonics Symposium (IUS), 2009 IEEE International*, 2009, pp. 714-717.

[22] E. Rubiola, *Phase noise and frequency stability in oscillators*: Cambridge University Press, 2008.

[23] D. W. Allan, "Statistics of atomic frequency standards," *Proceedings of the IEEE,* vol. 54, pp. 221-230, 1966.

[24] V. Kaajakari, T. Mattila, A. Lipsanen, and A. Oja, "Nonlinear mechanical effects in silicon longitudinal mode beam resonators," *Sensors and Actuators A: Physical,* vol. 120, pp. 64-70, 2004.

[25] C. Zuo, M. Rinaldi, and G. Piazza, "Power handling and related frequency scaling advantages in piezoelectric AlN contour-mode MEMS resonators," in *2009 IEEE International Ultrasonics Symposium (IUS09)*, 2009, pp. 1187-1190.

[26] M. Rinaldi, C. Zuo, J. Van der Spiegel, and G. Piazza, "Reconfigurable CMOS Oscillator Based on Multifrequency AlN Contour-Mode MEMS Resonators," *IEEE Transactions on Electron Devices,* vol. 58, pp. 1281-1286, 2011.

[27] J. D. Larson III, P. D. Bradley, S. Wartenberg, and R. C. Ruby, "Modified Butterworth-Van Dyke Circuit for FBAR Resonators and Automated Measurement System," in *2000 IEEE International Ultrasonics Symposium (IUS 2000)*, 2000, pp. 863-868.

Low Noise Oscillator Based on 2D Superconducting Resonator

Jean-Claude Mage, Bruno Marcilhac, Maurice Poulain,
Yves Lemaitre, Julien Kermorvant,
Thales R&T
91767 Palaiseau, France
jean-claude.mage@thalesgroup.com

Jean-Marc Lesage
Maîtrise de l'information
DGA, Bruz
BP7 – 35998 Rennes Armées, France
jean-marc.lesage@dga.defense.gouv.fr

***Abstract*—High Tc superconducting materials provide very high Q microwave resonators, which are supposed to make low phase noise oscillators. The expected phase noise has never been observed on operational oscillators. Our goal is to outperform the specifications of acoustic resonators and whispering gallery modes, with 2D resonators operated at 1 GHz: Q > 100,000 low 1/f noise Si transistors, -150 dBc/Hz at 100 Hz from the carrier.**

I. INTRODUCTION

High critical Temperature Superconducting (HTS) materials have been studied, and are commercialized for microwave filtering applications. They provide very selective, very low insertion filters thanks to high quality factor resonators [1-3]. These very high Q resonators are also expected to provide low phase noise oscillator. Nevertheless this has never been clearly demonstrated on operational oscillators. 3D resonators [4] exhibit very high Q > 10^6, but vibration sensitivity creates parasitic lines in the spectrum. 2D resonators [5] provide intermediate Q that decreases versus frequency. The choice of the reference frequency **f** is a very important point for the success of the demonstration.

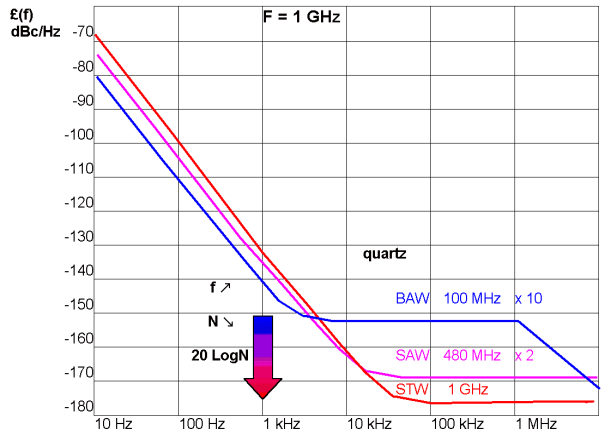


Figure 1. Phase noise £(f) at 1 GHz versus carrier offset frequency: Increasing the reference frequency lowers the multiplication rank and the noise floor level far from the carrier

State of the art reference oscillators are usually made with acoustic wave resonators (Bulk or Surface Acoustic Waves: BAW or SAW) in the megahertz range, and then multiplied by **N** to reach the operating frequency **F** of the system (or **F** is divided by **N** in synthesizers). The floor level is increased by 20 Log **N** (Fig.1). Thus the trend is to increase the reference frequency up to 1000 MHz. The theory of acoustic attenuation [6,7] shows that, for a given material and a given cut, the product of the quality factor times the frequency is a constant. According to the Leeson's theory [8], the phase noise rises from the floor level k_BT (-174 dBm/Hz at 300 K), at the offset frequency **f/2Q**. This results in a shift away from the carrier as the square of the frequency **f**, and in a detrimental increase of the colored noise close to the carrier (Fig.2).

A compromise can be found for many systems, or combining for example a 10 MHz BAW and a 1 GHz SAW can help for more stringent requirements. Nevertheless higher quality factors are highly desirable for improved oscillators, especially for radar applications.

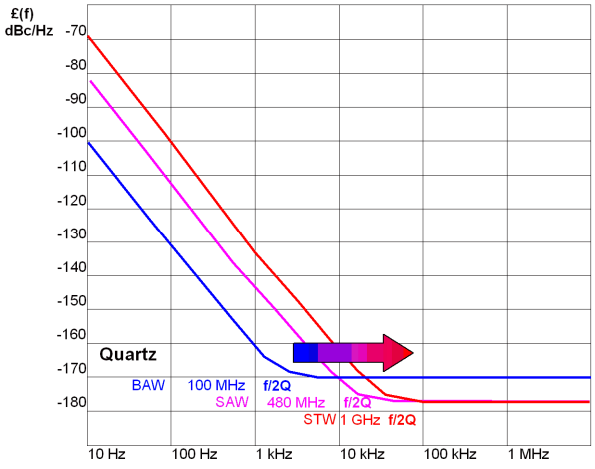


Figure 2. Increasing the reference frequency lowers the quality factor (Q.F is constant), shifts the corner away from the carrier as F^2 and increases the colored noise close to the carrier.

This research is partially funded by DGA REI 06.34.028.

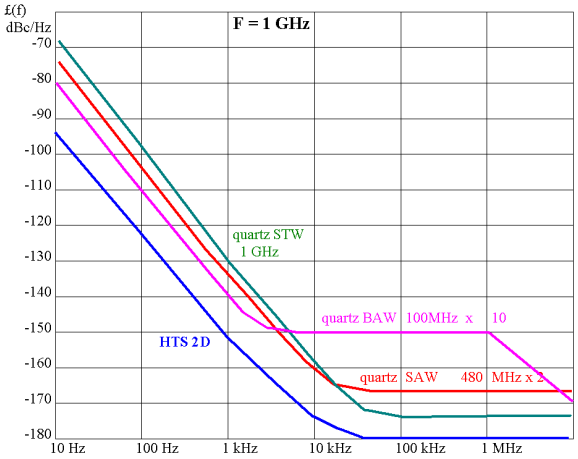


Figure 3. Comparison at 1 GHz between state of the art acoustic wave oscillators and expected phase noise of a HTS oscillator according to Leeson's theory for an unloaded Q of 100,000 and a loaded Q of 50,000 [5].

The goal of the study is to go beyond the performances of acoustic resonators (and whispering gallery modes). We know that 2D HTS resonators can provide a Q higher than 100,000 at 1 GHz (one order of magnitude higher than acoustic resonators), and only 10,000 or less at 10 GHz (a similar law Q.F ~ constant is found for acoustic, dielectric and HTS resonators, through different theoretical approaches). The choice of a 1 GHz 2D HTS resonator seems the most pertinent. Figure 3 confirms the validity of this approach with a phase noise lower than all acoustic wave oscillators for all offset frequencies. Moreover the advantages of 1 GHz are:

- upper frequency range of acoustic oscillators
- lower multiplication rank
- higher Q
- available low flicker noise silicon bipolar transistors
- light cryocooler compatible size (15x15 mm)
- noise floor level –180 dBm/Hz at 70 K

II. SUPERCONDUCTING RESONATOR

A. *Resonator Patterning*

$YBa_2Cu_3O_7$ films are deposited on both sides of a magnesium oxide single crystal substrate by sputtering [9,10].

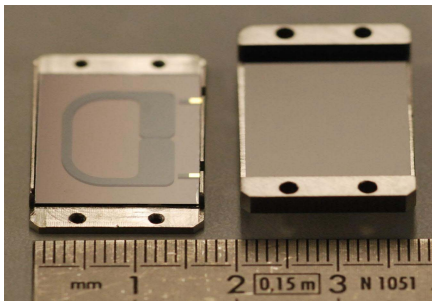


Figure 4. Superconducting resonator: $YBa_2Cu_3O_7$ film (700 nm thick) deposited on a magnesium oxide substrate (500 µm thick)

The surface resistance Rs of the superconducting films is measured at 10 GHz using the dielectric resonator method [10]. The values are between 200 and 300 µΩ at 77 K, two orders of magnitude lower than pure copper. Since the loss of superconductors evolves as the square of the frequency, the Rs at 1 GHz is 2 to 3 µΩ, more than three orders of magnitude lower than normal metals, which vary as the square root of the frequency.

The resonator (Fig. 4) is designed as a folded half wavelength line. Instead of the classical hairpin shape, we optimized the geometry in a nearly closed C in order to: minimize the size (large pads in the capacitive part of the line), improve the power handling capability (large half circle for the inductive part of the line where the current flows), and minimize the radiated energy (very short electric dipole corresponding to the gap).

Both the line and the ground plane are superconducting. The top and the walls of the enclosure are also coated with superconducting wafers in order to minimize eddy current losses.

Accurate integrated input and output couplings modeled with Sonnet e.m. software [12] are patterned simultaneously and overlaid by gold contacts.

B. *Resonator Microwave testing*

The very low surface resistance of $YBa_2Cu_3O_7$ films and the optimized design of the resonator give quality factors much higher than 100,000, typically 150,000 at 77 K (Fig.5). 77 K is often used as a reference temperature since it corresponds to liquid nitrogen, but standard Stirling cryocoolers (utilized for infrared cameras) can easily reach lower temperatures.

It is quite realistic to operate between 60 and 70 K where a Q higher than 200,000 is obtained. Moreover the power handling capability improves by one order of magnitude between 77 and 60 K. This is a most important point since high Q resonators have to carry high reactive power. High Q is not sufficient to get a low phase noise; the reactive power really makes the spectral purity of the oscillators.

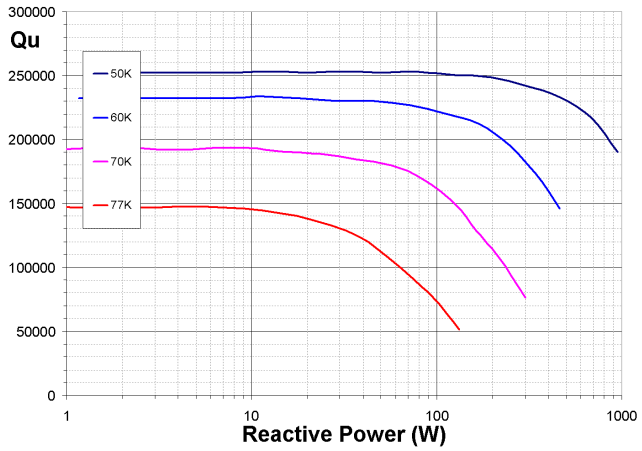


Figure 5. 1 GHz HTS planar resonator unloaded quality factor Q_U versus reactive power and temperature

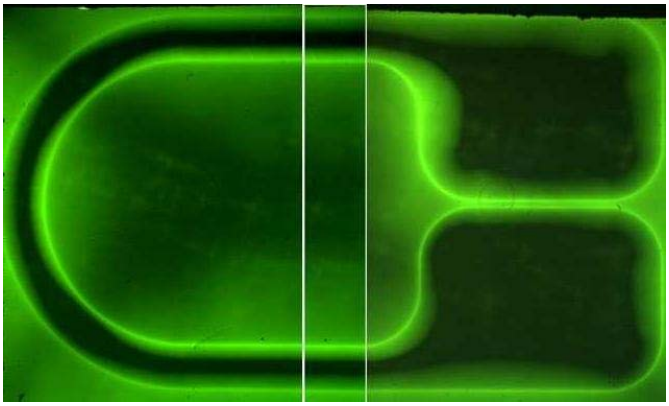


Figure 6. Magneto-optical control of the superconducting properties of the $YBa_2Cu_3O_7$ resonator

The power handling capability of the resonator is very dependent on the quality of the film and of the patterning, which can be controlled by a magneto-optical microscope (Fig.6). A perfect superconductor expels the magnetic field (Meissner effect). By exploiting the Faraday effect in a garnet film, it is possible to visualize the magnetic field map and to control the quality of the superconducting pattern.

III. Oscillator Circuit

The feedback oscillator circuit (Fig. 7) is designed using the AWR-MWO® software [13] for setting the oscillation conditions. This circuit includes:

- NEC-CEL SC 5801 silicon Bipolar Junction Transistor (BJT SPICE model with documented AF and KF flicker noise coefficients).
- Sonnet modeled resonator including integrated input and output couplings to set the gain close to unity
- Sonnet adjusted delay line to set the phase equal to 2π along the closed loop
- Infineon BBY55 Varactor for fine phase trimming
- Bias circuits

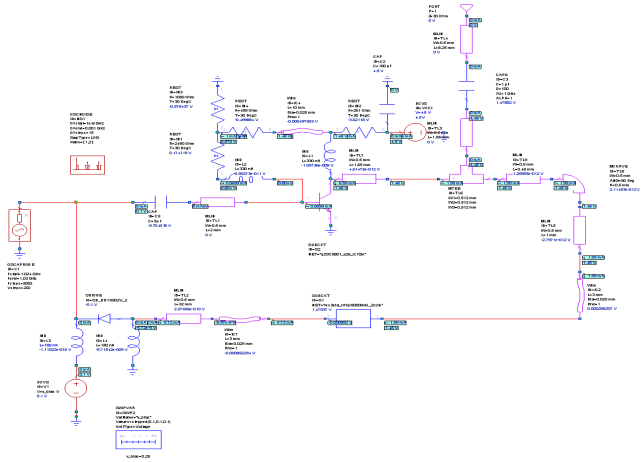


Figure 7. MWO-AWR 1 GHz oscillator modeling including SONNET modeled HTS resonator

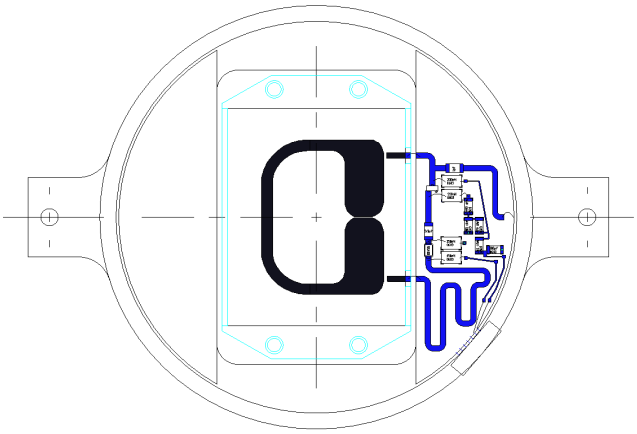


Figure 8. Feedback loop oscillator lay-out with:
60 K cooled 2D HTS resonator (black and light blue, central part),
room temperature low flicker noise Si BJT (deep blue, right part)

It was first intended to use a cooled SiGe transistor, but the flicker noise of SiGe is poorer than that of Si, and increases at low temperature. We chose a Si transistor with a very low flicker noise (1/f corner at 350 Hz), but Si transistors do not operate at 60 K. Thus the active part of the circuit is kept at room temperature (Fig. 8).

The length of the delay line is carefully adjusted in order to get the right phase when closing the loop. The Sonnet modeling takes into account the meander shape of this line and all the connecting lines of the circuit. The permittivity of the alumina substrate is controlled by the split cavity TE_{01n} mode method [14]. And the S parameters are re-injected in AWR-MWO. In spite of all these precautions, a varactor is inserted to provide a fine-tuning for technological dispersion compensation. The setting of the phase strictly equal to 2π is essential for the optimal operation and the purity of the oscillator. It may require an automated control (not utilized here)[15].

IV. Phase Noise

The phase noise of the feedback oscillator circuit is analyzed by the harmonic balance method in AWR-MWO.

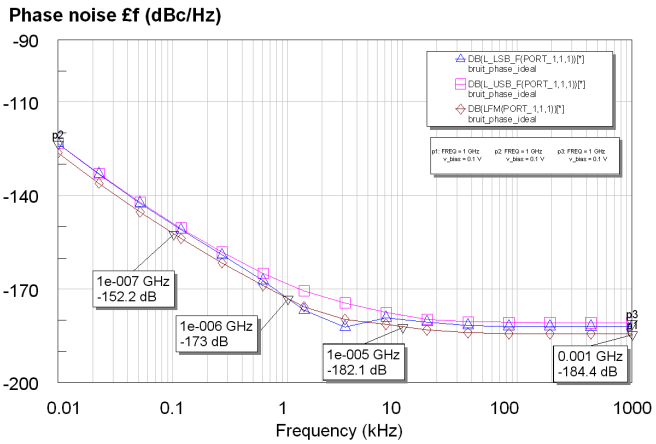


Figure 9. MWO-AWR phase noise modeling £(f_m) = LFM @ 1 GHz with an unloaded quality factor Q_U = 230 000

This modeling gives a theoretical noise, with a floor level at -184 dBc/Hz, which is coherent with a 10 mW output, and a colored noise frequency corner slightly below 10 kHz, which is coherent with a loaded Q_L higher than 50,000. Q_L was set at 50,000 in the open loop configuration, but it cannot be measured when the loop is closed and can only be estimated from the corner frequency.

The slope of the colored noise is -20dB/decade down to 100 Hz and clearly turns up to -30 dB/decade only below 100 Hz, which is better than expected from the transistor flicker noise specifications (1/f corner at 350 Hz).

These results are not incompatible with the straightforward application of the Leeson's theory. Nevertheless, they may appear over-optimistic and make the validity of the software questionable. We decided with the French defense administration DGA to benchmark it versus the Agilent ADS® software [16] (fig. 10).

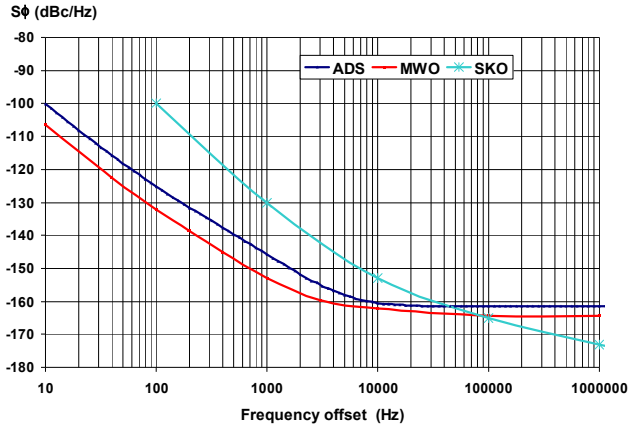


Figure 10. Comparison between phase noise Sϕ = 2 £(f) of multiplied HTS oscillator modeled by MWO-AWR and Agilent ADS, and measured PSI Poseidon® SKO @ 10 GHz

The phase noise Sϕ = 2 £(f) modeled by ADS, and transposed from 1 to 10 GHz for comparison, is about 6 dB higher than AWR-MWO, but remains quite low compared to acoustic oscillators and a SKO sapphire whispering gallery mode oscillator, manufactured by PSI Poseidon [17], which is the X band reference. The floor level is of the sapphire oscillator is lower; but the unloaded Q of a sapphire resonator is typically 200,000 at 10 GHz and 300 K, which gives a frequency corner around 50 kHz. The flicker noise of X band transistors is also higher, and the slope is -30 dB/decade up to 2000 Hz. The colored noise of the 10 GHz 3D oscillator is 10 to 20 dB higher than the multiplied 1 GHz HTS 2D oscillator.

V. CONCLUSION

We designed a superconducting planar resonator and designed a feedback oscillator. The results obtained by modeling with AWR-MWO and Agilent ADS are in good adequacy with the Leeson's theory and deserve the further effort to build a demonstrator. Practical interest of this oscillator lies in the capability of a compact cryogenic integration, which is our actual purpose.

ACKNOWLEDGMENT

The authors wish to acknowledge Olivier Llopis (LAAS) for transistor choice, Jesus Obregon (XLIM), Jean Chazelas and Alain Raymond Poezevara (Thales) and for useful discussions on oscillators and phase noise.

REFERENCES

[1] STI, http://www.suptech.com/home.htm

[2] M. J. Lancaster, University of Birmingham "Passive Microwave Device Applications of High-Temperature Superconductors", Cambridge University Press, ISBN: 9780521034173

[3] M. Reppel, J.C. Mage, "Superconducting microstrip bandpass filter on $LaAlO_3$ with high out-of-band rejection", IEEE Microwave and Guided Wave Letters,Volume: 10 Issue: 5 , May 2000, pp.180 -182

[4] L.Hao, J.C. Gallop, N. Klein, M Winter, "Low phase noise temperature compensated cryogenic whispering gallery mode", Conference on Precision Electromagnetic Measurements Digest, 2000

[5] S. A.. Vitusevich, N. Klein, A. G. Zaitsev, and J. Geerk "Phase Noise of an HTS Resonator Operated in the Nonlinear Regime" IEEE Trans. on Applied Superconductivity, Vol. 13, No. 2, pp.324, June 2003

[6] A. Akhiezer, J. Phys. (Moscow) 1, 277 (1939).

[7] L. Landau and G. Rumer, Phys. Z. Sowjetunion 11, 8 (1937).

[8] D. B. Leeson, "A simple model of feedback oscillator noise spectrum," Proceedings of the IEEE, 54, pp.329-330, 1966

[9] Y. Lemaître, M. Mercandalli, B. Dessertenne, D. Mansart, B. Marcilhac, J.C.Mage, "Large-area high-quality YBCO thin films by inverted cylindrical magnetron technique" , Physica C, Volume 235, pp. 643-644, 1994

[10] Y. Lemaitre, D.Mansart, B. Marcilhac, J.C. Mage, J. Siejka, "Double-sided sputtering deposition of YBCO thin films on 2" $LaAlO_3$ wafers for microwave applications", IEEE Transactions on Applied Superconductivity, Vol. 9 Issue 2 Part 2 , June 1999, pp.2363 -2366

[11] J.C. Mage, D. Dieumegard, "High Tc superconductors for microwave filters", AGARD, Applications of Superconductivity to Avionics 8 p (SEE N91-14030 05-76), 1990

[12] Sonnet, http://www.sonnetsoftware.com/

[13] AWR MicroWave Office ®, http://web.awrcorp.com/Usa/Products/Microwave-Office/

[14] M.D. Janezic, J. Baker-Jarvis, "Full-Wave Analysis of a Split-Cylinder Resonator for Nondestructive Permittivity Measurements" IEEE Trans. MTT, Vol. 47, N° 10, Oct. 1999, pp. 2014-2020

[15] Z. Galani, M. J. Bianchini, R. C. Waterman, R. Dibiase, R. W. Laton and J. B. Cole, "Analysis & design of a single resonator GaAs FET oscillator with noise degeneration", in IEEE Trans.MTT Vol. 32, pp.1556-64, 1984.

[16] Agilent ADS® software, http://www.home.agilent.com/agilent/product.jspx?nid=-34360.0.00&lc=eng&cc=US

[17] PSI, Poseidon Scientific Instruments ®, http://www.psi.com.au/

Phase Noise Improvement in Balanced SAW Oscillators

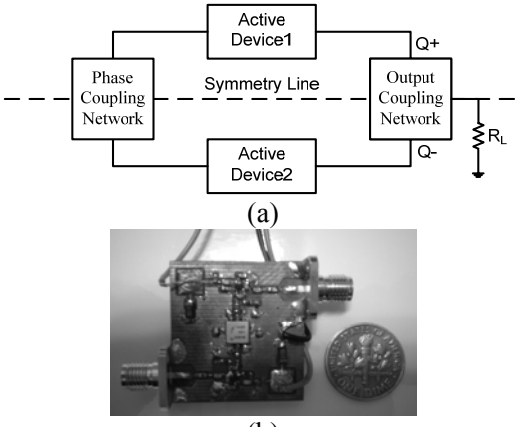
Yao Huang Kao, Jon Hung Lin, and I-Jhih Wu

Department of Communication Engineering, Chung Hua University, Hsinchu Taiwan

Email: yhkao@chu.edu.tw

SAW oscillators as stable clock sources are received a number of applications in information and communication industry. As the frequency is raised up several hundred MHz and GHz range, the output with differential waveform such as LVDS has many advantages for power supply noise rejection and direct match to the preceded stages of mixer and phase detector as in the clock and data recovery applications. However, the configuration of the oscillator normally adopts the Pierce or Colpitts oscillators, which needs a single to differential converter. Here, a balanced SAW oscillator is first presented. The differential outputs possess inherently 180^o out of phase in odd harmonics and in phase in even harmonics. By using these features, the powers of combined signal and noise are raised 6dB and 3dB, respectively. Thus, the phase noise of the combined signal can be improved at least 3dB, which is useful in low phase noise applications.

The balanced oscillator is constructed by using two identical Colpitts oscillators. The functional block is revealed in Fig. 1a. The one-port SAW resonator acts as the tank and is also acted as the phase coupling network between the identical oscillators. Under fundamental vibration mode, the RF current of the opposite plates of the resonator is out of phase, which drives the corresponding outputs of two identical oscillators 180^o out of phase. The balanced SAW oscillator has been fabricated on FR4 substrate and its photograph is shown in Fig.1b. The size of the pc board is 25mm x 25mm for the convenience of measurement. The DC supply voltage is 3.0V. Current consumption is 17mA with each output -2.5dBm. The differential outputs are fed to an output coupling network (0^o or 180^o power combiner) and a push-push oscillator will be available. Output power is increased to +3dBm with 6 dB improvement. Thus, a significant reduction in phase noise was also expected. The measured phase noise is shown in Fig.2. At 100 kHz offset from the carrier, the phase noise is about -158 dBc/Hz for balanced outputs and -164 dBc/Hz for the output combined with 180^o power combiners. Thus, 3dB, at least, improvement in phase noise is verified.



(a)

(b)

Fig. 1 (a)Functional Block and (b) PCB layout of the Balanced Oscillator with One-port SAW Resonator at 433MHz

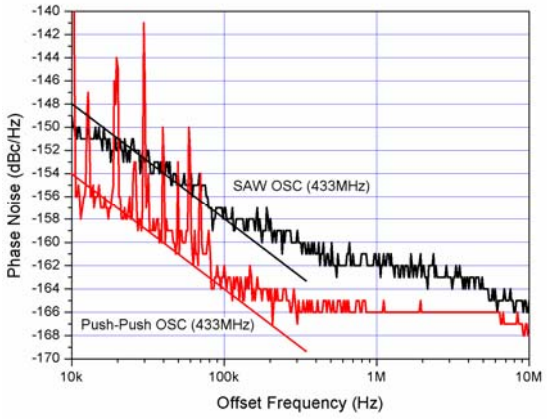


Fig. 2 Comparison of measured phase noises with and without power combination

Generation of 100 GHz with Parts in 10^{16} Frequency Stability Using Cryogenic Sapphire Oscillators

Romain Bara, Jean-Michel Le Floch, Michael E. Tobar *Fellow IEEE*, Paul L. Stanwix, Stephen Parker, John G. Hartnett, Eugene N. Ivanov

University of Western Australia School of Physics M013, 35 Stirling Hwy., Crawley, 6009 Western Australia, Australia

Rbara@cyllene.uwa.edu.au, lefloch@cyllene.uwa.edu.au, mike@physics.uwa.edu.au

***Abstract*—We report on the generation of millimetre wave signals with frequency instability of a few parts in 10^{16}. Such performance was achieved at 103.750 GHz by multiplying frequencies of two almost identical 12.969 GHz oscillators based on liquid helium cooled sapphire dielectric resonators. The multiplication stages were shown to only add a small amount of noise at averaging times less than 10 seconds, resulting in a minimum of frequency instability of $5x10^{-16}$ at 20 seconds averaging time. Such ultra-stable signal sources operating at frequencies of 100 GHz are very important for many applications including high frequency radio astronomy.**

INTRODUCTION

Precise frequency generation is important for a variety of applications including radar, telecommunications, positioning and navigation, geodesy and time keeping, tests of fundamental physics and radio astronomy [1-8]. So far, signals with the short-term fractional frequency instability of a few parts in 10^{16} have been generated at microwave frequencies only. This was achieved either with classical oscillators based on sapphire loaded cavity resonators cooled to liquid helium temperature, termed as "sapphire clocks" or "cryogenic sapphire oscillators" (CSO) [9-11], or via the use of mode-locked laser technology by extracting high order harmonics of pulse repetition rate from a femtosecond pulse train referenced to an ultra-stable laser [8]. In this letter we demonstrate that the excellent fractional frequency stability of a typical microwave CSO can be transferred to frequencies above 100 GHz by means of direct frequency multiplication.

In the Extremely High Frequency (EHF) band (30-300 GHz) low noise frequency generation at parts in 10^{16} frequency instability below 1000 seconds averaging time has never been achieved. Such precise frequency generation at EHF has application for radio astronomy, high-speed data links, point-to-point high bandwidth communication and space-borne radio communications. In particular the next generation of high frequency VLBI will require a reference oscillator better than the best Hydrogen Masers (which typically have stabilities of parts in 10^{13} to 10^{14} over averaging times of 1 to 30 second) [12].

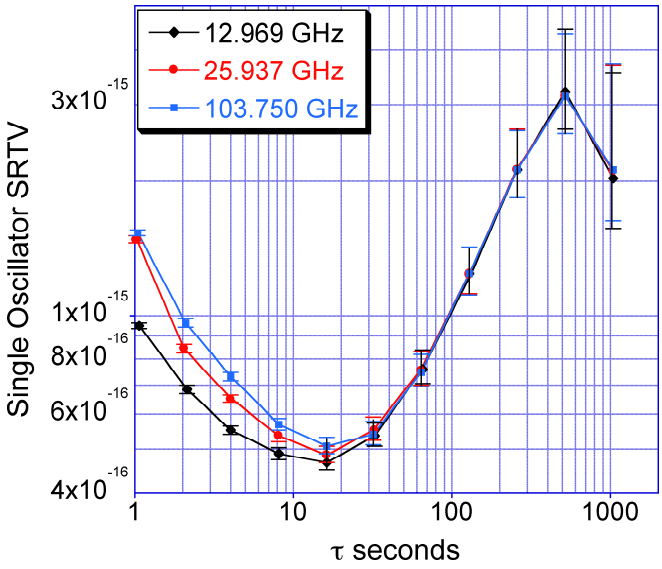


Fig. 2. Frequency instability as measured by the Square Root Triangle Variance with respect to a single oscillator.

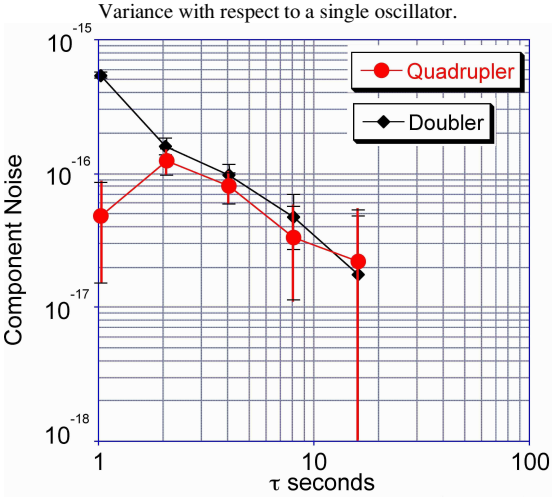


Fig. 3. Frequency instability added by the doubling (bold diamond) and quadrupling (bold circle) stages as calculated by the quadratic difference between the 12.969, 25.937 GHz and 103.750 GHz stages.

CONCLUSION

To our knowledge this is the first time fractional frequency instability below 10^{-15} in the range of integration times 2 to 100 s was demonstrated for mm-wave signals at frequencies above 100 GHz. This was achieved by multiplying frequency of a 12.969 GHz cryogenic sapphire oscillator. The excess noise of frequency multiplication stages was estimated as contributing to oscillator frequency instability at the level of a few parts in 10^{17} for integration times above 4 s. The noise in the multiplication stages was low enough to enable a very low level of frequency instability to be measured at the high frequency of 103.75 GHz. This has important implications for many precision measurement and metrological applications, in particular for the next generation of high frequency precision radio astronomy.

This work was support by the Australian Research Council Grants No. DP0878108, DP0878108 and FL0992016.

REFERENCES

[1] A. E. E. Rogers, S. Doeleman, M. Wright, G. C. Bower, D. C. Backer, S. Padin, J. A. Phillips, D. T. Emerson, L. Greenhill, J. M. Moran, and K. I. Kellermann, "Small-Scale Structure and Position of Sagittarius a-Asterisk from VLBI at 3 Millimeter Wavelength," *Astrophys. Journ.*, vol. 434, pp. L59-L62, 1994.

[2] S. S. Doeleman, J. Weintroub, A. E. E. Rogers, R. Plambeck, R. Freund, R. P. J. Tilanus, P. Friberg, L. M. Ziurys, J. M. Moran, B. Corey, K. H. Young, D. L. Smythe, M. Titus, D. P. Marrone, R. J. Cappallo, D. C. J. Bock, G. C. Bower, R. Chamberlin, G. R. Davis, T. P. Krichbaum, J. Lamb, H. Maness, A. E. Niell, A. Roy, P. Strittmatter, D. Werthimer, A. R. Whitney, and D. Woody, "Event-horizon-scale structure in the supermassive black hole candidate at the Galactic Centre," *Nature*, vol. 455, pp. 78-80, 2008.

[3] M. E. Costa, D. G. Blair, M. J. Buckingham, A. J. Giles, S. K. Jones, A. N. Luiten, P. J. Turner, A. C. Young, P. Hong, and A. G. Mann, "A Sapphire Oscillator for VLBI Radio Astronomy," *Meas. Sci. Technol.*, vol. 3, pp. 718-22, 1992.

[4] M. E. Costa, D. G. Blair, M. J. Buckingham, M. W. Sinclair, R. H. Ferris, D. L. Jauncey, J. E. Reynolds, A. K. Tzioumis, and R. A. Preston, "Application of Commercial Antennas to VLBI Radio Astronomy," *Rev. Sci. Instrum.*, vol. 66, pp. 995-999, 1995.

[5] M. E. Tobar, P. L. Stanwix, M. Susli, P. Wolf, C. R. Locke, E. N. Ivanov, "Rotating resonator-oscillator experiments to test Lorentz invariance in electrodynamics," *Lect. Notes Phys*, vol. 702, pp. 416-50, 2006.

[6] P. L. Stanwix, M. E. Tobar, P. Wolf, M. Susli, C. Locke, E. Ivanov, J Winterflood, F van Kann, "Test of Lorentz invariance in electrodynamics using rotating cryogenic sapphire microwave oscillators," *Phys. Rev. Lett.*, vol. 95, 040404, 2005.

[7] P. Wolf, S. Bize, A. Clairon, A. N. Luiten, G. Santarelli, and M. E. Tobar, "Tests of Lorentz invariance using a microwave resonator," *Phys. Rev. Lett.*, vol. 90, no. 6, 060402, 2003.

[8] J. Millo, M. Abgrall, M. Lours, E. M. L. English, H. Jiang, J. Guena, A. Clairon, M. E. Tobar, S. Bize, Y. Le Coq, and G. Santarelli, "Ultralow noise microwave generation with fiber-based optical frequency comb and application to atomic fountain clock," *Appl. Phys. Lett.*, vol. 94, 141105, 2009.

[9] J. G. Hartnett, C. R. Locke, E. N. Ivanov, M. E. Tobar, and P. L. Stanwix, "Cryogenic sapphire oscillator with exceptionally high long-term frequency stability," *Appl. Phys. Lett.*, vol. 89, 203513, 2006.

[10] M. E. Tobar, E. N. Ivanov, C. R. Locke, P. L. Stanwix, J. G. Hartnett, A. N. Luiten, R. B. Warrington, P. T. H. Fisk, M. A. Lawn, M. J. Wouters, S. Bize, G. Santarelli, P. Wolf, A. Clairon, P. Guillemot, "Long-term operation and performance of cryogenic sapphire oscillators," *IEEE Trans. on Ultrason. Ferroelec. Freq. Contr.*, vol. 53, pp. 2386-93, 2006.

[11] C. R. Locke, E. N. Ivanov, J. G. Hartnett, P. L. Stanwix, and M. E. Tobar, "Invited article: Design techniques and noise properties of ultrastable cryogenically cooled sapphire-dielectric resonator oscillators," *Rev. Sci. Instrum.*, vol. 79, 051301, 2008.

[12] S. Doeleman, "High Frequency Very Long Baseline Interferometry: Frequency Standards and Imaging an Event Horizon," in *Proceedings of the 7th Symposium Frequency Standards and Metrology*, pp. 175-183, 2009.

DFB Laser Diode Pumped Rubidium Atomic Frequency Standard

J. Q. Huang, Y. Gu, X. H. Qi, Y. H. Wang, S. Q. Liu, T. Q. Dong
School of Electronics Engineering and Computer Science
Peking University
Beijing, China
wangyanhui@pku.edu.cn

J. Zhang, Z. H. Lu
School of Physics
Huazhong University of Science and Technology
Wuhan, China

Abstract—**We present our recently experimental results on a DFB laser diode pumped Rubidium atomic frequency standard. Our experiment was based on a homemade lamp-pumped Rubidium frequency standard with the discharge lamp assembly removed. In the experiment, we first locked a DFB laser diode to the ^{87}Rb cycling transition ($5S_{1/2}$, F=2→$5P_{3/2}$, F=3) by means of saturation absorption locking technique. After that, we optimized the parameters carefully related to the clock performance, such as temperature of the physical package, amplitude of the modulation signal for the laser locking.**

The preliminary short-term stability of our clock is measured $5.25\times10^{-12}\tau^{-1/2}$ (1s<τ<100s)**.**

I. INTRODUCTION

In a traditional Rubidium frequency standard, a discharge Rb lamp is used as the light source. And after the discharge lamp, an ^{85}Rb filter gas cell is typically employed so that only "b" line light of ^{87}Rb is left to produce the pumping light. However, since the filter process is always imperfect, it results in broadening the microwave resonance line, generally 1-2 kHz and additional light noise. These make the signal contrast of a traditional Rb clock keep below 1% and damp the further improvement in the short-term stability.

To eliminate the problem resulted from the discharge lamp and the filter process, and improve the short-term stability of Rubidium frequency standard remarkably, the narrow linewidth laser diode was proposed to be implemented as the pumping light source. The narrow linewidth laser diode, which can offer excellent control of the light spectra, can greatly improve the pumping rate and avoid the additional light noise. As a result, the short-term stability can be greatly improved in this condition. For the desired improvement, interest and attempts in this field have never diminished, since the preliminary experiment result reported by L. L. Lewis and M. Feldman [2]. On one hand, theoretical analysis and calculation were carried out to make certain of the influence of different noises and shot noise limit of the laser pumped Rb clock [3-9]; on the other hand, experiments were designed to improve the signal-to-noise ratio (SNR) and reduce the light shift effect to advance the performance of laser-pumped Rb clock [10-14]. And during last few years, progress in this field was mainly made on the compact laser-pumped Rb clock [15-16] and the pulsed Rb clock [17-20].

In this paper, we report our recently preliminary experiments on a DFB laser diode pumped Rubidium frequency standard, which is based on a homemade traditional Rb clock with the discharge lamp assembly removed. The similar work was also carried out by C. Affolderbach, *et al*. on a commercial Rb clock [1]. Our paper is organized in the following four sections. Section one introduces the experimental setup. Section two is about the parameter optimizing. The third and fourth sections are respectively the preliminary experimental result and the conclusion.

II. EXPERIMENTAL SETUP

Our experimental setup can be divided into two main parts: the laser head and the clock module. The laser head consists of a DFB laser diode and the frequency stabilization devices. And the clock module is our one homemade traditional Rubidium frequency standard with the discharge lamp removed as shown in Fig. 1.

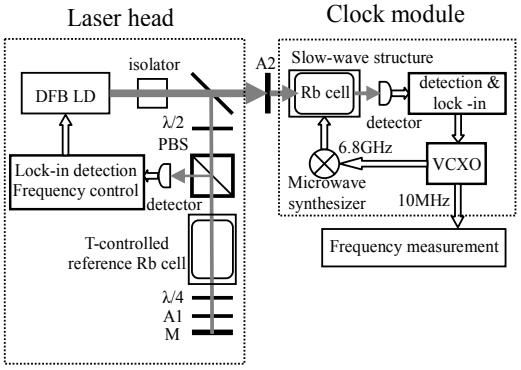


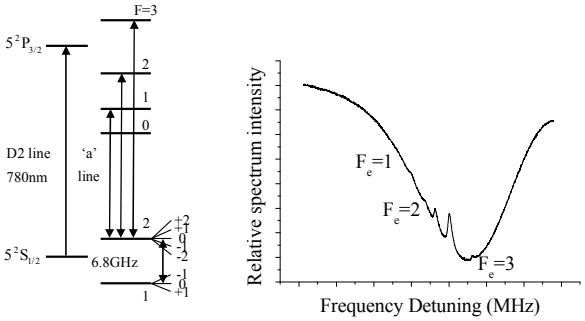
Fig.1. Experimental setup. λ/2 and λ/4, wave plates; A1 and A2, attenuators; M, mirror; PBS, polarizing beam splitter.

In the experiment, we lock the DFB laser diode to the cycling transition ($5S_{1/2}$, F=2→$5P_{3/2}$, F=3) of ^{87}Rb by means of saturated absorption locking technique. In this condition, all the ^{87}Rb atoms on the ground state ($5S_{1/2}$, F=2)

This work is supported by "the Fundamental Research Funds for the Central Universities".

are pumped to the excited state ($5P_{3/2}$, F=3) and fall back to the ground state ($5S_{1/2}$, F=1) through nonradiative transitions. Because of this pumping process, around 33% of all the ^{87}Rb atoms can be utilized for magnetic resonance, with the SNR evidently improved. And to suppress the spectrum drift caused by temperature fluctuation, another temperature controlled reference Rb cell is employed. The energy level structure and saturated absorption spectra for frequency stabilization of ^{87}Rb are presented in Fig. 2.

In the clock module, the microwave resonance signal is generated in the Rb gas cell with the microwave slow-wave structure. The VCXO is locked to the clock transition of ^{87}Rb (F_g=1, m_F=0→F_g=2, m_F=0) and provides 10MHz. signal.



(a) **(b)**

Fig.2. (a) Hyperfine level structure of ^{87}Rb (D_2, 'a') line. (b) The saturated absorption spectra of ^{87}Rb (D_2, 'a') line.

III. PARAMETER OPTIMIZATION

After finishing the basic setup of the clock system, we optimize the parameters carefully related to the clock performance. In this experiment, the environment temperature drift and the light shift effect make great influences on the long-term stability of the clock. The temperature point of the physical package also affects the short-term stability. For the parameters optimizing, we lay emphasis on searching the best temperature point of the physical package, and the amplitude of the modulation signal for laser frequency stabilization.

A. Temperature of the physical package

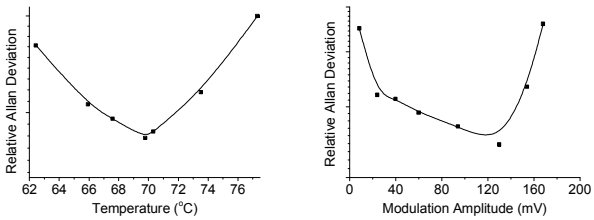
In the Rb gas cell, the temperature factor affects seriously the saturated vapor pressure of Rb atoms. Consequently, the microwave resonance signal intensity will change when altering the physical package temperature, with the laser light intensity unchanged. In this way, the short term stability $\sigma(\tau)$ can be seriously influenced.

On this step, we search for the temperature point where the clock works best. And to find out this key point, we carried out measurement on the temperature dependences of the clock short-term stability (Fig. 3a). As a result, we found that short-term stability appears best when the temperature controlled around 70°C.

B. Amplitude of the modulation signal

The amplitude of the modulation signal for the laser frequency stabilization has a strong influence on the SNR of the error signal. And the error signal affects apparently the frequency stabilization of the laser diode. Since the AM and FM noise from the laser diode greatly limit the short-term stability of the clock, we changed the amplitude of the modulation signal to improve the laser-locking effect and reduce these noises. And the frequency of the modulation signal was maintained at 8 kHz on this step.

To search for the best parameter, we measured the short-term stability of the clock while regulating the amplitude (Fig. 3b). From Fig. 3b, we can read that neither too big nor too small amplitude the clock performs best. The most appropriate amplitude of the modulation signal for laser locking was found to be 130mV in the end.



(a) **(b)**

Fig.3. (a) The relevance between the physical package temperature and the clock short-term stability. (b) The relevance between the modulation signal amplitude and the clock short-term stability.

IV. PRELIMINARY RESULT

After optimizing the parameters above, we measured a preliminary short-term stability of the DFB laser diode pumped Rubidium frequency standard, $5.25\times10^{-12}\tau^{-1/2}$ ($1s<\tau<100s$). And this is 20% better than the lamp-pumped Rubidium clock on the same physical package (Fig. 4).

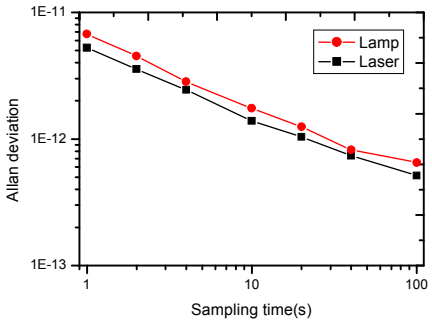


Fig.4. Preliminary short-term frequency stability of DFB laser diode pumped Rubidium frequency standard.

V. CONCLUSION

In this experiment, we built a DFB laser diode pumped Rubidium frequency standard. The laser diode was locked to the cycling transition ($5S_{1/2}$, F=2→$5P_{3/2}$, F=3) of ^{87}Rb. After

relevant parameters optimized, the preliminary measured short-term stability of the clock was $5.25\times10^{-12}\tau^{-1/2}$ ($1s<\tau<100s$). And our recent experiments indicate that the laser noises limit greatly the performance of the clock. When considering the long-term frequency stability, the long term environmental temperature drift and the light shift effects are the major restrictions. We are also trying to improve these factors to get better results for long terms.

REFERENCES

[1] C. Affolderbach, G. Mileti, and F. Droz, "A laser optically-pumped rubidium vapor-cell frequency standard using a DFB laser diode," CLEOE-IQEC, pp. 1, 2007.

[2] L. L. Lewis and M. Feldman, "Optical pumping by lasers in atomic frequency standards," 35th Ann. Freq. Contr. Symp. , pp. 612-624,1981.

[3] J. C. Camparo and R. P. Frueholz, "Fundamental stability limits for the diode-laser-pumped rubidium atomic frequency," J. Appl. Phys. Vol. 59, No. 10, pp. 3313-3317, 1986.

[4] G. Mileti, J. Q. Deng, F. L. Walls, J. P. Lower, R. E. Drullinger, "Recent progress in laser-pumped rubidium gas-cell frequency stability," 50th Proc. of IEEE Int. Freq. Contr. Symp., pp. 1066-1072, 1996.

[5] J. Q. Deng, G. Mileti, D. A. Jennings, R. E. Drullinger, and F. L. Walls, "Improving the short-term stability of laser pumped Rb clocks by reducing the effects of the interrogation oscillator," 51st Proc. of IEEE Int. Freq. Contr. Symp., pp. 438-445, 1997.

[6] G. Mileti, J. Q. Deng, F. L. Walls, D. A. Jennings and R. E. Drullinger, "Laser-pumped rubidium frequency standard: new analysis and progress," IEEE J. Quant. Elect., Vol. 34, No. 2, pp. 233-237, 1998.

[7] J. Vanier, and C. Mandache, "The passive optically pumped Rb frequency standard: the laser approach," IEEE IFCS-EFTF Joint Conference, pp. 1346-1351, 2007.

[8] V. Zholnerov, A. Besedina, G. Kazakov, "Comparison of achievable performances as regards rubidium gas cell frequency standards with continuous and pulse laser pumping and with coherent population trapping," IEEE IFCS-EFTF Joint Conference, pp. 587-593, 2009.

[9] M. Hashimoto and M. Ohtsu, "Experiments on a Semiconductor Laser Pumped Rubidium Atomic Clocks," IEEE J. Quant. Elect., Vol. 23, NO. 4, pp. 446-451, 1987.

[10] N. D. Bhaskar, "Potential for improving the rubidium frequency standard with a novel optical pumping scheme using diode lasers," IEEE Trans. Ultra. Ferro. and Freq. Contr., Vol. 42, NO. 1, pp. 15-20, 1995.

[11] Y. Saburi, Y. Koga, S. Kinugawa, T. Imamura, H. Suga and Y. Ohuchi, "Short-term stability of laser-pumped rubidium gas cell frequency standard," IEEE Elect. Lett., vol. 30, No. 8, pp. 633-635, 1994.

[12] X. R. Huang, X. H. Tu, G. H. Mei, and X. W. Zhu, "Experiment of diode laser pumped rubidium frequency standard", Chinese Journal of Lasers, Vol. A29, No. 6, 2002.

[13] C. Affolderbach, C. Andereeva, S. Cartaleva, T. Karaulanov, G. Mileti, and D. Slavov, "Light-shift suppression in laser optically pumped vapor-cell atomic frequency standards," Appl. Phys. B Vol. 80, No. 7, pp. 841-848, 2005.

[14] A. Besedina, O. Berezovskaya, A. Gevorkyan, and V. Zholnerov, "Study of a 87Rb vapor discriminator with laser pumping for a high stability onboard frequency standard," IEEE IFCS-EFTF Joint Conference, pp. 607-612, 2007.

[15] C. Affolderbach, G. Mileti and F. Droz, "A compact, high-performance laser-pumped rubidium atomic frequency standard," 18th EFTF, pp. 228-232, 2004

[16] C. Affolderbach, F. Droz, and G. Mileti, "Experimental demonstration of a compact and high-performance laser-pumped rubidium gas cell atomic frequency standard," IEEE Trans. Instr. and Meas., Vol. 55, No. 2, 2006.

[17] A. Godone, S. Micalizio, F. Levi, and C. E. Calosso, "Physics characterization and frequency stability of the pulsed rubidium maser", Physics Review A, Vol.74, No. 4, pp. 43401-1-12, 2006.

[18] A. Godone, S.e Micalizio, C. E. Calosso, and Filippo Levi, "The pulsed rubidium clock," IEEE Trans. Ultra. Ferro. and Freq. Contr. , Vol. 53, No. 3, pp. 525-529, 2006.

[19] A. Godone, F. Levi, S. Micalizio, E. K. Bertacco, and C. E. Calosso, "Frequency-stability performance of the pulsed optically pumped rubidium clock: recent results and future perspective," IEEE Trans. Instr. and Meas., Vol. 56, No. 2, pp. 378-382, 2007.

[20] S. Micalizio, A. Godone, F. Levi, and C. Calosso, "Pulsed optically pumped ^{87}Rb vapor cell frequency standard: a multilevel approach," Physical Review A, Vol. 79, No. 1, pp. 013403-1-19, 2009.

Rotating Dual Cryogenic Sapphire Oscillators with 10^{-16} Fractional Frequency Stability for Tests of Lorentz Invariance

Stephen Parker, Paul Stanwix, Michael Tobar,
John Hartnett and Eugene Ivanov
School of Physics
The University of Western Australia
Crawley 6009, Australia
Email: Stephen.Parker@physics.uwa.edu.au

Moritz Nagel and Achim Peters
Institut für Physik
Humboldt-Universität zu Berlin
Berlin 12489, Germany

***Abstract*— We describe the latest improvements and current status of our cryogenic sapphire oscillator Michelson-Morley experiment that has recently been moved from The University of Western Australia to Humboldt University of Berlin.**

I. Introduction

The original Michelson-Morley experiment [1] used an optical interferometer to search for an orientation dependent change in the speed of light due to the presence of a luminiferous aether. In contrast, this experiment features two orthogonally aligned cryogenic sapphire oscillators that are mounted in the same vacuum space and actively rotated in the laboratory. The beat frequency is recorded and analysed for orientation dependent frequency changes which could be caused by orientation dependent variations in the speed of light. These possible variations are expressed as bounds placed on coefficients in the Standard Model Extension framework [2], [3] and describe violations of Lorentz invariance. Our first cryogenic sapphire oscillator Michelson-Morley experiment ran at The University of Western Australia from 2004 - 2006 [4] and was able to constrain nine photon sector coefficients from the Standard Model [5]. In order to place better bounds on these coefficients we need to improve the frequency stability of the experiment.

II. Cryogenic Sapphire Oscillators

A 5.1 cm diameter cylindrical sapphire crystal is mounted in a silver plated copper cylindrical cavity and excited at 12.968 GHz via SMA antenna probes. This sapphire loaded cavity sustains a standing wave whispering gallery mode ($WGE_{16,0,0}$) with a dominant radial electric field. WGE modes are twice as sensitive to Lorentz Violations compared to WGH modes (dominant crystal axis electric field) which the previous experiment operated in. At cryogenic temperatures the sapphire crystals exhibit quality factors on the order of one billion which is an order of magnitude improvement compared to the previous experiment. These improvements are achievable due to the larger sized diameter of the new crystals. The beat frequency of the two cryogenic sapphire oscillators exhibits a temperature/frequency turning point at 5.5 K (figure 1).

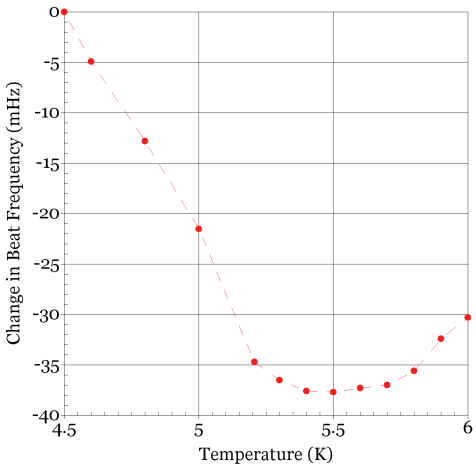


Fig. 1. Beat frequency as a function of sapphire loaded cavity temperature.

Power, frequency and temperature control loops are implemented using voltage controlled attenuators, phase shifters and heaters respectively. These control loops maintain oscillation at the desired mode and increase frequency stability.

III. Experiment Design and Rotation

Two sapphires are mounted orthogonally in the same copper block and are sealed within two vacuum cans to reduce thermal fluctuations. The cans are submerged in a 200 L liquid helium bath that allows for three weeks continuous operation. Three copper heatsinks provide a thermal path to the liquid helium (figure 2). The majority of the control electronics are housed at room temperature above the dewar. Ambient temperature fluctuations in the lab cause phase shifts in the microwave control electronics and may limit the beat frequency stability, the full impact of this effect is currently being analysed. The dewar is rotated with a period of 44 seconds on an air-bearing table that records angular position to within 350 μRads. Power

978-1-61284-110-6/11/$26.00 ©2011 IEEE

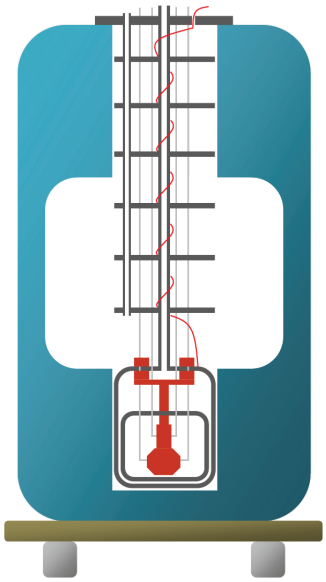


Fig. 2. Cross section of dewar containing liquid helium bath and two sapphire loaded cavities.

and electrical signals enter and leave the table via a rotating connector. A rotating mercury connector is used to bring the 10 MHz reference on to the table for the frequency counter and helium gas returns via a custom made rotating connector.

The table is mounted on three aluminium cylinders which in turn are each mounted on a pressure sensor. These are used to align the centre of mass with the rotation axis. A set of orthogonal tilt sensors are mounted in the middle of the table. Two of the aluminium cylinders are temperature controlled and thus thermally expanded to compensate for long term tilt variations. The experiment tilts up to 6 μRads during one rotation (figure 3) which is believed to be the underlying cause of the dominant systematic as the beat frequency oscillates in a similar manner. Stationary tests have found a beat frequency tilt sensitivity of 17 mHz / Degree which is not enough to limit the frequency stability of the beat. The previous experiment was rotated using a belt and stepper motor system and had no tilt compensation system.

IV. Frequency Stability and Future Work

For short integration times the frequency stability of the beat is limited by rotation induced systematics as can be seen by the hump at twice the rotation frequency in the Allan Deviation plot (figure 4). For longer integration times the frequency stability becomes similar for rotating and stationary results. The frequency stability at half the rotation period is lowered by over an order of magnitude compared to the previous experiment. We are continuing to investigate ways to increase the frequency stability.

The experiment will run for a full year and we anticipate being able to improve our bounds placed on Standard Model Extension coefficients by a factor of 75 compared to our

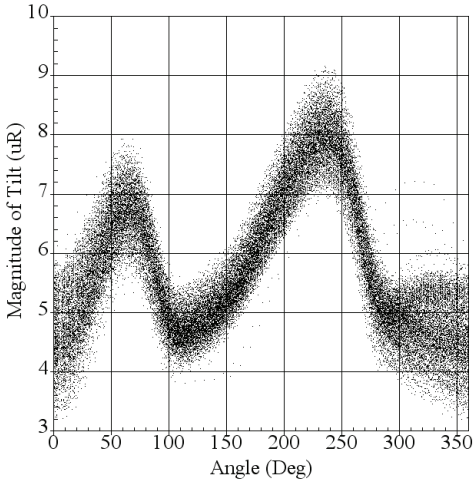


Fig. 3. Magnitude of tilt of experiment as a function of angular position. Data was collected over 1100 rotations.

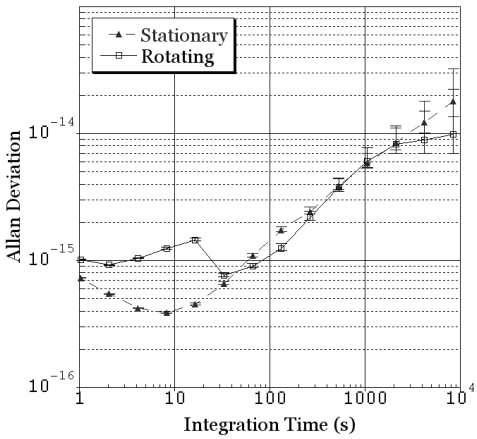


Fig. 4. Beat frequency stabilities with linear drift and $\sqrt{2}$ removed.

previous setup. After this run the cryogenic sapphire oscillators will be combined with a set of optical resonators for a joint experiment that will allow us to place bounds on a wider range of Lorentz violating coefficients [6].

Acknowledgment

This work was funded by Australian Research Council grants FL0992016, DP0878108 and DP1092690. The authors are grateful for assistance provided by The University of Western Australia Convocation Postgraduate Research Travel Award and the IFCS / EFTF Student Travel Awards.

References

[1] A.A. Michelson and E.W. Morley, Am. J. Sci. **34**, 333 (1887); Phil. Mag. **24**, 449 (1887).
[2] D. Colladay and V.A. Kostelecký, Phys. Rev. D **55**, 6760 (1997); **58**, 116002 (1998).
[3] V.A. Kostelecký, Phys. Rev. D **69**, 105009 (2004).
[4] P.L. Stanwix *et al.*, Phys. Rev. D **74**, 081101 (2006).
[5] M.A. Hohensee *et al.*, Phys. Rev. D **82**, 076001 (2010).
[6] V.A. Kostelecký and M. Mewes, Phys. Rev. D **80**, 015020 (2009).

Light Shift in CPT Based Cs Miniature Atomic Clocks

D. Miletic, C. Affolderbach, G. Mileti
Laboratoire Temps-Fréquence (LTF)
Université de Neuchâtel
Neuchâtel, Switzerland
gaetano.mileti@unine.ch

M. Hasegawa, C. Gorecki
University of Franche-Comté
FEMTO–ST , UMR 6174 CNRS
Besançon Cedex, France

Abstract— **We report on our studies of the light shift in a miniature atomic clock based on Coherent Population Trapping (CPT) in Cesium. There is no data published in the literature on the frequency and intensity-related light shift coefficients for CPT atomic clock on the Cesium D_1 line, only for the D_2 line. With our data we fill this gap and report on the measurements of light shift for the Cs D_1 line in a MEMS buffer-gas cell. The behavior of the light shift due to different laser intensities, laser frequencies and due to different cell temperatures is discussed.**

I. INTRODUCTION

Coherent Population Trapping (CPT) can occur in a Λ system in which the two hyperfine components of the $S_{1/2}$ ground state of an alkali atom are coupled to a common excited state by two laser fields that are nearly resonant with the two atomic transitions [1]. If the frequency difference of the two laser fields is close to the atomic hyperfine splitting of the two ground states, quantum coherence between the two hyperfine components will be effectively generated. This quantum coherence reduces the total absorption and fluorescence and, under certain conditions, atoms can be completely trapped in a so-called coherent dark state, where they do not interact with the light any more. CPT resonances can be very narrow, which makes them good candidates for the realization of atomic clock. For this application however, systematic shifts of the CPT resonance have to be understood and controlled.

In CPT, the presence of two resonant light fields not only couples the two atomic states but may also shift the relative position of the two energy states through the AC Stark effect. This "light-shift" depends on the frequency and intensity of the light and may be expressed as a function of the two coefficients α and β, often referred to as the intensity light-shift coefficient (α) and the frequency light-shift coefficient (β) [2]. In the case of CPT in an ideal three-level system with no decay of the ground-state coherence, significant light-shift suppression is expected. When a single-mode laser, frequency modulated in the GHz range is used to create the CPT signal, each component of its spectrum will contribute to the ground-states shifts. In principle, the two first order sidebands (which are used in this case for coupling the Cs ground states) will create frequency light-shifts of opposite signs and should therefore cancel out. However, second order effects and the shifts due to the carrier and to all the other laser sidebands create additional shifts resulting in non negligible light-shift. In some cases, all these contribution may cancel each other and result in a suppressed light-shift [3]. Here we show that this suppression can also be achieved with carefully selected cell temperature.

II. EXPERIMENTAL SETUP

The experimental scheme is shown in the Figure 1.

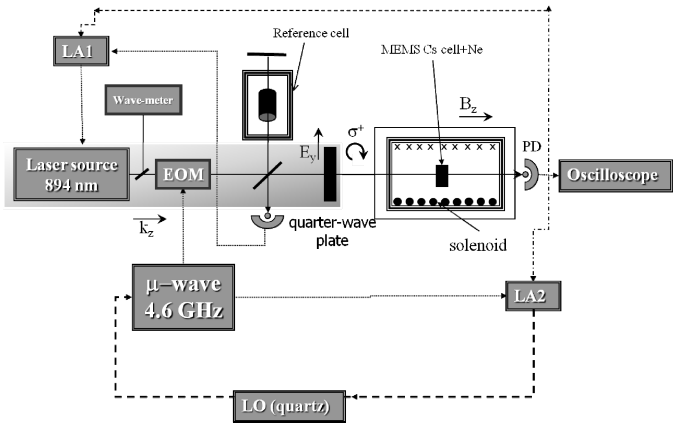


Figure 1. Experimental scheme

A DFB laser at 894.6 nm, resonant with the Cs D1 line is used as a light source [5]. Optical beam is modulated at ν ~4.6GHz by using a fiber-coupled Electro-Optical Modulator (EOM) and two phase-coherent electromagnetic fields, separated by 2ν ~ 9.2 GHz, are generated. In this case the two

This work was funded by EU FP7 (MAC-TFC project, no. 224132) and by the Swiss National Science Foundation (project 200020-118062).

first order sidebands (+ 1 and -1) are used to create the Λ-scheme. A beam splitter is placed directly after the EOM in order to split the laser beam in two different paths. One part of the beam is sent trough an evacuated macroscopic cell and the resulting absorption signal can be used to stabilize the laser frequency using a standard lock-in scheme. The second part of the laser beam is circularly polarized by a quartz-plate before passing through a MEMS Cesium vapor cell placed in the clock's Cs cell package. The Cs MEMS cell contains a Neon buffer gas and its external dimensions are 2.4 x 4 x 6 mm [4]. A photodiode is placed directly after the Cs cell package to detect the CPT signal. This signal can also be used to stabilize the laser frequency in alternative to the separate absorption setup. In this case, the laser is locked to a broadened line (Voigt profile resulting from Doppler and homogeneous broadening due to buffer gas). In the Cs cell package, the MEMS cell is temperature controlled and is surrounded by a solenoid generates a static magnetic field parallel to the light propagation vector to lift the degeneracy of the Zeeman sublevels. The ensemble is surrounded by 2 μ-metal magnetic shields to reduce perturbations from external magnetic fields. The cell temperature can be stabilized in a relatively wide temperature range (45 to 90˚C). A commercial frequency synthesizer (referenced to a Hydrogen Maser or to quartz a VCO – for spectroscopy or closed-loop clock mode, respectively) produces the 4.6 GHz microwave radiation.

III. EXPERIMENTAL RESULTS

We measured the light-shift as a function of the laser power at different MEMS cell temperature and the results are reported in Figure 2. In this case the laser frequency was locked using the buffer-gas MEMS cell and the measurements have been performed for two different situations of laser detuning: coupling the ground states through the F'=4 excited state or coupling through the F'=3 state. These two couplings are achieved by tuning the laser frequency to the optical transition involving the respective excited atomic state. The light-shift coefficient α is extracted as the value of the slopes of the line fits in Figure 2a and 2b.

In the case when the laser is locked to the F'=3 state, for lower temperatures the light shift coefficient α is negative. For higher temperatures, the light shift becomes positive. For cell temperatures around 82˚C, the light shift is significantly suppressed. In the case of F'=4 laser locking, light shift is always positive for any cell temperature and we did not find a temperature with suppressed light shift. Similar results were obtained in the case when the laser is locked to the separate evacuated macroscopic cell.

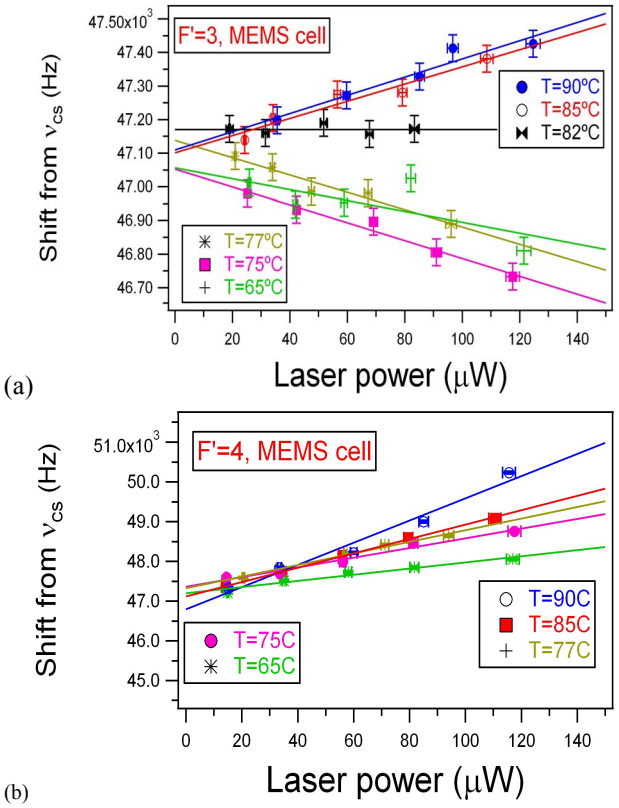


Figure 2. (a): Shift of the CPT line as a function of the laser power for laser frequency tuned to the F'=3 state. The measurements were repeated for different cell temperature (b): Same measurements like in Figure 2 (a) for laser frequency tuned to the F'=4 state.

Our results show that the intensity light-shift changes its sign with changing cell temperature (Fig 2a). With increasing temperature, the first order sidebands will be absorbed more strongly by the Cs atoms, reducing their average light intensity acting on the atoms in the cell. Thus the overall balance between positive and negative light-shift contributions from the different laser sidebands will change, as it is seen on our measurements.

We have demonstrated the impact of the reported light-shift reduction by recording the frequency stability of our clock. At time scales around ~ 1000 s we measured an improvement of clock stability by factor of 10, reaching the 8×10^{-11} level, by choosing the proper temperature of the buffer-gas cell [6].

IV. CONCLUSION

We presented the first experimental results on the light shift in a micro-fabricated cell filled with Cs and pure Ne buffer gas, for CPT spectroscopy on the Cs D1 line (at 894.6 nm). The light-shift is found to change the sign as a function of cell temperatures. A cell temperature was found for which

light shift is significantly suppressed. We previously demonstrated that a mixture of two buffer gases is not needed, because a turnover temperature around 80˚C can also be obtained with pure Ne as buffer gas [7]. This temperature coincides with the one suitable for the low intensity light-shift reported above, in the case when the laser is locked to F'=3 state. Furthermore, this temperature is also in the correct temperature range for operating a miniature atomic clock.

ACKNOWLEGMENTS

We thank our colleagues from LTF-UniNe for their contributions. We also thank V. Giordano and the team from FEMTO-ST CNRS-UFC, Besançon, France for helpful discussions.

REFERENCES

[1] G. Alzetta, A. Gozzini, L. Moi, and G. Orriols, "An experimental method for the observation of R.F. transitions and laser beat resonances in oriented Na vapor", Nuovo Cimento, Vol. 36, pp. 5–20, August 1976.

[2] Cohen-Tannoudji C. and Dupont-Roc J., "Experimental Study of Zeeman Light Shifts in Weak Magnetic Fields", Phys. Rev. A, vol. 5, pp. 968 – 984, February 1972.

[3] Zhu M. and Cutler L. S., "Theoretical and experimental study of light shift in CPT-based Rb vapor cell frequency standard", Proc. Recise Time and Time Interval Meeting (PTTI), Reston (VA), USA, November 28-30, 2000, pp. 311 – 324.

[4] M. Hasegawa, R.K. Chutani, C. Gorecki, R. Boudot, P. Dziuban, V. Giordano, S. Clatot, L. Mauri, "Microfabrication of cesium vapor cells with buffer gas for MEMS atomic clocks", Sensors and Actuators A: Phys., vol. 167, pp. 594 – 601, March 2011.

[5] F. Gruet, D. Miletic, C. Affolderbach, G. Mileti, V. Vilokkinen, P. Melanen, "Spectral characterization of aged and non-aged 894 nm DFB for their application in Cs atomic clocks", Proceedings of the International Symposium on Reliability of Optoelectronics for Space, Cagliari, Italy, May 11 – 15, 2009, pp. 295 – 299.

[6] D. Miletic, C. Affolderbach, M. Hasegawa, R. Boudot,C. Gorecki and G. Mileti, "AC Stark shift in CPT based Cs miniature atomic clocks", submmited for publication (2011).

[7] D. Miletic, P. Dziuban, R. Boudot, M. Hasegawa, R. K. Chutani, G. Mileti, V. Giordano and C. Gorecki, "Quadratic Dependence on Temperature of the Cs 0-0 Hyperfine Resonance Frequency in a Single Ne Buffer Gas Microfabricated Vapor Cell", Electron. Lett., vol. 46, pp. 1069 - 1071, July 2010.

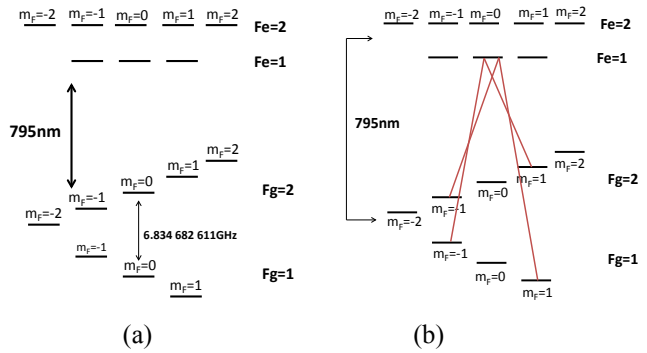
Light shift of double resonance and coherent population trapping in wall-coated cells for compact Rb clocks

D. Miletic, T. Bandi, C. Affolderbach and G. Mileti
Laboratoire Temps – Fréquence (LTF)
Institut de Physique, Université de Neuchâtel
Neuchâtel, Switzerland
Email: gaetano.mileti@unine.ch

***Abstract*— We report on a comparative study on the light shift in two possible clock schemes for vapor-cell standards, Double Resonance (DR) and Coherent Population Trapping (CPT), by using the same wall coated cell and compact atomic resonator. Our experimental setup allowed us to measure and study the AC Stark shift, resulting from a mono-chromatic or a non monochromatic optical excitation. The light-shift was investigated as function of the laser intensity and the laser frequency.**

I. INTRODUCTION

The light-shift (or AC Stark shift) of the atomic clock transition is often one of the main sources of instability in vapor-cell atomic frequency standards ("Rb clocks"). Here we report on our investigations of the light-shift in two possible clock schemes shown in Fig. 1: Double Resonance (DR) interrogation (Fig. 1a) [1] and Coherent Population Trapping (CPT, Fig. 1b). The CPT resonances are created using the so-called lin//lin scheme [2, 3]. For the DR interrogation, the optical and microwave radiations are applied separately applied to the atoms, but simultaneously and in continuous mode, while for CPT the microwave frequency is imprinted onto the light field by frequency modulation. In both cases, the same laser wavelength was used (Rb D1 transition, 795 nm).



(a) (b)

Figure 1. (a): Double resonance scheme. (b) CPT scheme.

This work was funded by the Swiss National Science Foundation (project 200020-118062) and the European Space Agency (ESA). We also acknowledge the support from the ASRH (Association Suisse pour la Recherche Horlogère).

II. EXPERIMENTAL SETUP

The block diagram of the clock setup is shown in Fig. 2. Light from a compact laser head [4] is frequency-stabilized to the Rb D1 transition (795 nm) and passes through an Electro-Optical Modulator (EOM). Before irradiating the resonance cell, the laser beam is linearly polarized, and has a diameter of 5 mm. The vapor cell is 14 mm long and has 14mm diameter, contains pure 87-Rb, and its inner walls are coated with paraffin. The cell is placed inside a magnetron-type cavity in which the resonant microwave field has a TE011-like geometry, as described in [5]. A solenoid generates a static magnetic field parallel to the light propagation vector and the whole physic package is surrounded by 2 mu-metal magnetic shields. DR (resp. CPT) is obtained by injecting the microwave signal into the magnetron cavity (resp. the EOM).

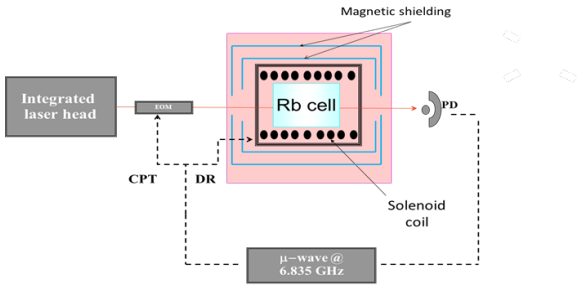


Figure 2. Block diagram of the experimental setup.

III. LIGHT SHIFT MEASUREMENTS

The clock transition may be shifted due to the AC Stark effect [6]. This shift depends on the spectrum of the optical radiation and may be expressed - in first and local approximation - as a function of the two coefficients α and β, often referred to as the intensity light-shift coefficient (α) and the frequency light-shift coefficient (β) [7]. These coefficients and the achievable laser frequency and intensity stabilities ultimately determine the eventual medium and long-term performance of the clock.

A. *Intensity light-shift coefficient (α)*

We measured the intensity light-shift as a function of the laser power at different laser detuning in both the DR and CPT clock schemes and the results are reported in Fig. 3 and 4, respectively. The measurements have been performed when locking the laser frequency to the four available sub-Doppler references (F_{ground} = 1, 2 to $F_{excited}$ = 1, 2). The light-shift coefficient α is given by the slopes of the line fits in Fig. 3 and 4. For instance, when locking the laser to the F_{ground} = 2 to $F_{excited}$ = 1 sub-Doppler transition, and the coefficients are $-7.5 \cdot 10^{-11}$/μW for the DR scheme and $-3.2 \cdot 10^{-11}$/μW for the CPT scheme, respectively.

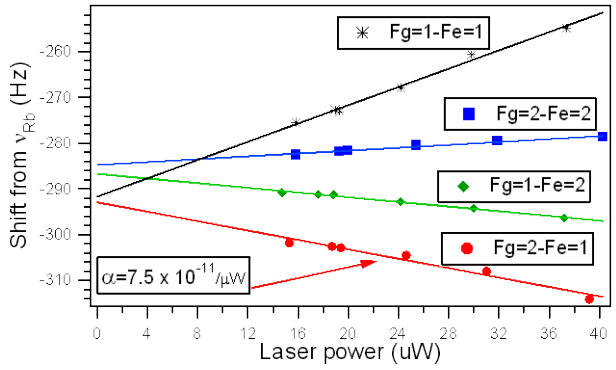


Figure 3. Shift of the DR line as a function of the laser power for laser frequency tuned to the 4 different laser detunings. Shifts are given relative to the unperturbed Rb hyperfine splitting frequency.

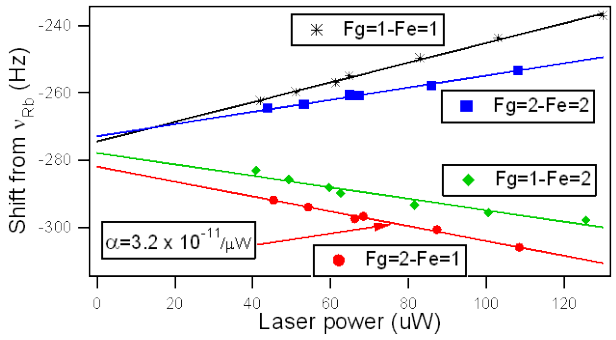


Figure 4. Shift of the CPT resonance as a function of the laser power (similar to the data of Fig. 3 but for the CPT scheme).

B. *Frequency light-shift coefficient (β)*

We also measured the shift of the clock transition as a function of the laser frequency for DR and for CPT. The results are reported on Fig. 5 and 6, respectively. Note that similar measurements were reported in [8], but not over such a large laser frequency scan and only for the case of a buffer gas cell. In Fig. 5, one may observe some differences as compared to the theoretical values presented in [9], in particular concerning the number of laser frequencies which give a zero-value intensity light-shift (we find two such laser frequencies instead of four occurrences given in [9]). Work is in progress in order to understand these differences and to analyze the analogue result for the CPT case.

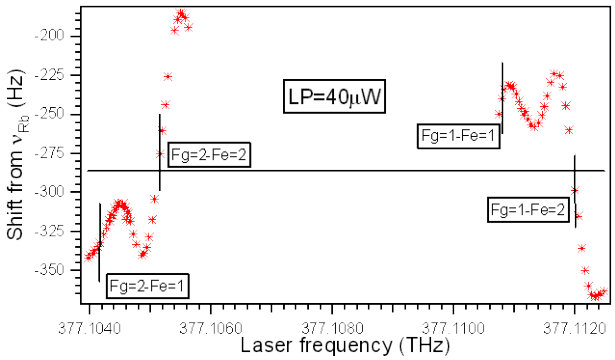


Figure 5. Shift of the DR line as a function of the laser frequency for the constant laser power of 40 μW.

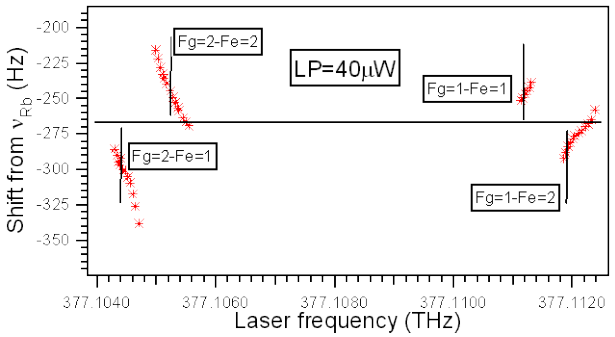


Figure 6. Similar measurements of β for the CPT clock scheme. Shifts are given relative to the unperturbed Rb hyperfine splitting frequency.

From the results shown in Fig. 5 and 6, one may extract the frequency light-shift coefficients (β) for any nominal laser frequency. For instance, when locking the laser to the F_{ground} = 2 to $F_{excited}$ = 1 sub-Doppler transition at an optical power of 40 μW, this coefficient is +310 mHz/MHz for the DR and -93 mHz/MHz for the CPT.

IV. Summary and Conclusions

We present our measurements and study of the intensity and frequency light shift in vapor-cell standards using a wall-coated cell and two clock schemes: DR and CPT. These results will be relevant in order to optimize the performances of a future compact atomic clock and determine its fundamental limits for medium and long-term stability.

Acknowledgment

We thank our colleagues F. Gruet, P. Scherler, and M. Durrenberger for their contributions to this work.

References

[1] C. Affolderbach, F. Droz, G. Mileti, "Experimental demonstration of a compact and high-performance laser-pumped Rubidium gas-cell atomic frequency standard", IEEE Transactions on Instrumentation and Measurements, Vol. 55, No. 2, April 2006, pp. 429-435.

[2] E. Breschi, G. Mileti, "Dark resonances in wall-coated cells for Rb clocks", Proceedings 24th EFTF, ESA/ESTEC, Noordwijk (NL), 2010.

[3] E. Breschi et al., "Study of light effects in the atomic-motion-induced Ramsey narrowing of dark resonances in wall coated cells", Phy. Rev. A, vol. 82, 063810, 2010.

[4] C. Affolderbach and G. Mileti, “A compact laser head with high-frequency stability for Rb atomic clocks and optical instrumentation”, Rev. of Scien. Instr., vol. 76, 073108, 2005.

[5] T. Bandi, C. Affolderbach, and G. Mileti, “Study of Rb 0-0 hyperfine double resonance transition in a wall-coated cell”, Proceedings 24th EFTF, ESA/ESTEC, Noordwijk (NL), 2010.

[6] C. Cohen-Tannoudji and J. Dupont-Roc, “Experimental Study of Zeeman Light Shifts in Weak Magnetic Fields”, Phys. Rev. A, vol. 5, pp. 968-984, 1972.

[7] W. Happer, “Optical Pumping”, Rev. Mod. Phys, vol. 44, pp.169-249, 1972.

[8] J. Q. Deng , G. Mileti, J. M. Lopez-Romero, D. A. Jennings, F. L. Walls, R. E. Drullinger, “Study of the frequency stability of laser-pumped Rb gas-cell frequency standards”, 11th EFTF, Neuchâtel, 1997.

[9] B. S. Mathur, H. Tang and W. Happer, “Light shifts in the alkali atoms”, Phy. Rev., vol. 171, pp. 11-19, 1968.

Reduction of the Magnetic Sensitivity of an Atomic Clock Against the non-Uniformity and Variations of the C-field

Louis Marmet
Institute for National Measurement Standards
National Research Council, Ottawa, Ontario, Canada
Email: louis.marmet@nrc-cnrc.gc.ca

***Abstract*— The frequency of an atomic clock may be stabilized against C-field variations by applying weak magnetic pulses tuned to the Larmor frequency. The small signal produced by these pulses is a Ramsey spectrum of the Zeeman transitions which, when added to the microwave Ramsey signal, may be adjusted to compensate exactly for the quadratic frequency shift of clock transition at a given value of the C-field. The cancellation is provided by a feed-forward mechanism which modifies the error signal without the need to adjust the C-field. This can be implemented in most atomic clocks by generating the required audio-frequency pulses on wires producing a magnetic field perpendicular to the C-field.**

I. Introduction

In an alkali-atom frequency standard, special care must be taken to produce a uniform C-field, stabilize it over extended periods of time and evaluate the large bias produced by the second-order Zeeman shift. The C-field can change for several different reasons such as variations of the external magnetic field, aging of the shield material, environmental temperature fluctuations, magnetization of the shields or hysteresis, mechanical shocks, and ageing of the electronic circuits. Because the frequency inaccuracy of an atomic clock is strongly dependent on the C-field, some methods have been proposed to reduce the magnetic field sensitivity. These methods can be classified by the type of mitigation used to reduce the magnetic field sensitivity. They fit into four general groups:

i) *Use atomic transitions with a weak dependence on the C-field.* The clock transition is selected to have a quadratic frequency dependence on the magnetic field and a weak magnetic field is used (e.g. the transition labeled 0-0, Fig. 1). A variation on this is to choose a large C-field value where the frequency dependence of the transition becomes null [1] (e.g. the transition between $|3,-1\rangle$ and $|4,-1\rangle$ labeled "$-1,-1$" in Fig. 1). However, this method requires an extremely uniform C-field and has not yet been implemented successfully.

ii) *Generate a very stable and uniform C-field.* Stable electronic circuits, permanent magnets and shielding offer possible methods to stabilize the C-field using passive stabilization or active feedback. A current generating the C-field can be stabilized to a constant value, magnetic shielding can be used to mitigate magnetic field variations, or permanent magnets can be used to generate the C-field. All these methods suffer from a large sensitivity on the temperature.

iii) *Use feedback from a measured value of the C-field.* Several methods have been proposed to measure the C-field and correct its value to keep it constant. One method proposes to use a dual frequency synthesizer [2] to run a clock at a high C-field while simultaneously measuring the C-field. Other methods involving some kind of adjustment to the C-field have been suggested [3]–[6]. However, it is known that such adjustments must be frequently applied because after a C-field change the frequency often relaxes back to almost where it started [7]. Most of the stability problems have been traced to the magnetic shields and their hysteresis effects. Measuring the C-field usually interferes with the operation of the clock, requires a complex system and affects the performance of the microwave system.

iv) *Use a feed-forward mechanism based on the value of the C-field.* For example, the "Multi-Coherent method" [8] uses a combination of many microwave and radio frequencies with optical pumping to achieve multiple measurements simultaneously. However, it only reduces the dependence of the frequency on the C-field by 33% (by using an average of all m_F states) and suffers from problems associated with an increased complexity of the microwave synthesizer which could affect the long-term frequency stability of the clock.

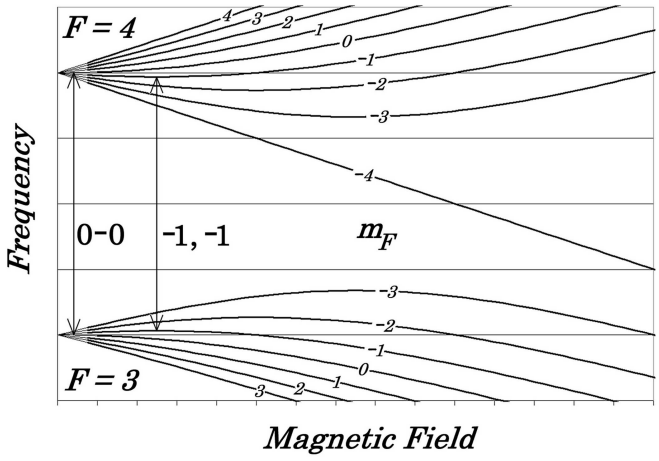


Fig. 1. Diagram of the hyperfine energy levels in a cesium atom. The relative frequency of each level is plotted as a function of the magnetic field.

978-1-61284-110-6/11/$26.00 ©2011 Crown

The stabilization method presented in this paper is a feed-forward mechanism which does not require changing the C-field nor changing the current producing it by more than a very small amount. This keeps the system operating in a linear regime. The feed-forward mechanism is provided by a small asymmetry produced on the microwave Ramsey fringe by coherent transfer of a small amount of population between the Zeeman levels. Two weak rf magnetic pulses are generated at a frequency near the Zeeman splitting of the hyperfine sub-levels and applied to the atoms between the microwave pulses. The magnetic pulses produce rf Ramsey fringes which are dependent on the linear frequency splitting between the hyperfine Zeeman sub-levels. When the C-field changes by a small amount, this rf Ramsey signal causes a frequency pulling on the microwave Ramsey fringes which can be adjusted to compensate exactly the quadratic Zeeman shift. With the proper amplitude adjustment of the rf magnetic pulses, the clock frequency can be made independent of small variations of the C-field. A detailed description is provided below with numerical examples given for the two cases of a beam clock and a fountain clock.

II. Description of the method

Any change in the value of the C-field will affect the frequency of a clock because it is referenced to an atomic transition which has a quadratic frequency dependence on the magnetic field $f_{00} = C_1 \left\langle B^2(z(t)) \right\rangle$, where $\langle\rangle$ denotes the time average, $B(z(t))$ is the value of the C-field at position $z(t)$ along the trajectory between the two microwave Ramsey excitations, and C_1 is a constant determined by atomic properties (in cesium $C_1 = 42.745\text{mHz/nT}^2$). A variation of the C-field away from its nominal value B_0 produces a frequency deviation on the microwave Ramsey fringe $\delta_R = C_1 \left\langle B^2(z(t)) - B_0^2 \right\rangle$. For small variations of the C-field away from its nominal value $B_0 >> B(z(t)) - B_0$ the frequency deviation of the microwave Ramsey fringe is $\delta_R \cong 2B_0C_1 \left\langle B(z(t)) - B_0 \right\rangle$. Therefore, at a given value of the C-field and as long as the deviation is small, the frequency offset of the clock will depend linearly on the C-field deviation from its nominal value.

To compensate this linear dependence of the frequency offset of the clock, one can take advantage of the linear frequency dependence of the Zeeman transitions. Several methods already exist to measure the C-field using the linear frequency shift [9], [10]. However, these measurements interfere with the normal operation of the clock and are only performed during the evaluation phase of the clock. Also, the measurements have to be repeated periodically to monitor the value of the C-field. In the method presented here, only one initial measurement of the C-field is required before running the clock. The feed forward mechanism keeps the frequency of the clock independent of small variations of the C-field.

The feed-forward signal is produced by applying rf magnetic pulses at a frequency f_C tuned near the Zeeman frequency $f_Z = C_2 \left\langle B(z(t)) \right\rangle$, where C_2 is a constant determined by atomic properties (in cesium $C_2 = 3.5$ kHz/μT). The field of the rf magnetic pulses must be aligned perpendicular to the C-field in order produce transitions between the $m_F = \pm 1$ states. The rf magnetic pulses produce Ramsey fringes (see Fig. 2), but in this case the signal is as a function of $f_C - f_Z$. A variation of the C-field away from its nominal value B_0 produces a frequency deviation on the rf Ramsey fringe $\delta_Z = C_2 \left\langle B(z(t)) - B_0 \right\rangle$. Thus, the frequency deviation of the microwave Ramsey fringe can be expressed as a function of the frequency deviation of the rf Ramsey fringe $\delta_R = 2B_0C_1\delta_Z/C_2 = \epsilon\delta_Z$. With weak rf magnetic pulses, the rf Ramsey fringes can be adjusted to produce a "frequency pulling" equal to $-\epsilon\delta_Z$, thus compensating for the quadratic Zeeman shift.

Thus, this technique uses a rf Ramsey interrogation within a microwave Ramsey interrogation. Since the Zeeman frequency shift is much larger than the Ramsey shift ($1/\epsilon \simeq 6800$ for a typical field of $6\mu T$ in a commercial cesium beam clock), the rf-pulses need only to have a small amplitude such that only a small fraction of the population is transferred. Also, since the C-field is sampled by each atom along its individual trajectory, spatial variations of the C-field are canceled as well as temporal variations. Thus, this method stabilizes the frequency of an atomic clock against temporal variations of the C-field and its non-uniformity.

III. Implementation examples

For the examples given below, a commonly used measurement cycle is assumed in the atomic clock. The source of atoms produces atoms distributed between the possible atomic states $|F, m_F\rangle$. A state selector only transmits the atoms in the states $|3, m_F\rangle$ (the cesium atom is used for this description). Ramsey interrogation [11] is used to interrogate the atoms with two microwave $\pi/2$-pulses at a frequency f_{MW} near the hyperfine transition frequency f_{00} between states $|3, 0\rangle$ and $|4, 0\rangle$ (the 0-0 transition or the clock transition). For the detection, a second state selector only transmits the atoms which are in any of the states $|4, m_F\rangle$ before they reach the detector. A typical microwave Ramsey signal measured as a

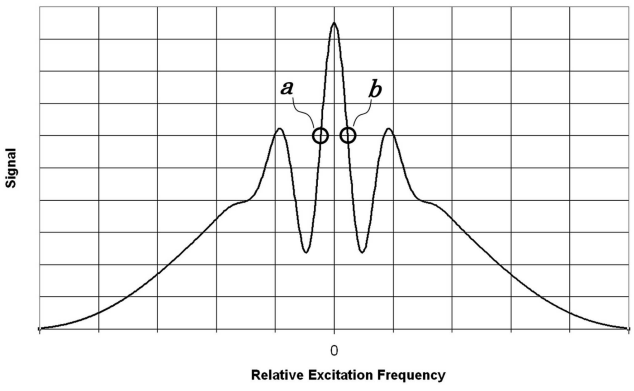


Fig. 2. Ramsey fringes and frequency points used by the locking circuit. The function is referred to as $R(f)$ in the text, with f being either the microwave frequency or the rf-detuning.

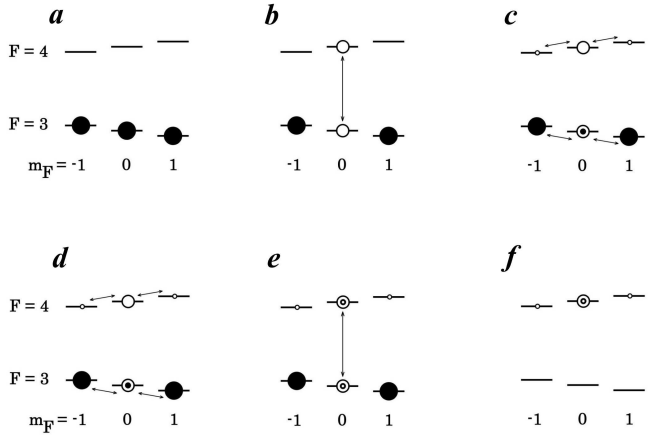


Fig. 3. Populations in a beam clock a) after the first state selector, b) after the first microwave Ramsey pulse, c) after the first rf Ramsey pulse, d) after the second rf Ramsey pulse, e) after the second microwave Ramsey pulse, and f) after the second state selector.

function of $f_{MW} - f_{00}$ is shown in Fig. 2. The local oscillator (clock frequency) is locked to the atomic reference using a frequency modulation scheme at the two frequencies *a* and *b* at $f_{MW} \pm \Delta_f$. The servo control adjusts the frequency of the local oscillator to keep the microwave Ramsey signals measured at points *a* and *b* at the same level. When this condition is met, the local oscillator is locked to the 0-0 transition, that is, f_{MW} is locked to f_{00}.

A. Compensation in a clock without magnetic state selection

This example describes the method as it could be implemented in a typical beam clock or in a fountain clock without state selection. The measurement cycle is assumed to start with all atoms in $F = 3$ states. Two rf magnetic field pulses, timed between the microwave Ramsey pulses, are applied on the atoms.

The states of interest for the microwave Ramsey interrogation are the states $|3, 0\rangle$ and $|4, 0\rangle$ (the notation $|F, m_F\rangle$ is used). Since the rf magnetic pulses have a small amplitude, these two states connect to the states $|3, -1\rangle$, $|3, 1\rangle$, $|4, -1\rangle$ and $|4, 1\rangle$. The contribution from the other states is negligible and not considered in this discussion. The population in each state is $a_{3\pm1}\,|3, \pm1\rangle$, $a_{30}\,|3, 0\rangle$, $a_{4\pm1}\,|4, \pm1\rangle$ and $a_{40}\,|4, 0\rangle$ (the populations are not normalized). Also, $\rho_{3\pm1} = |a_{3\pm1}|^2$, $\rho_{30} = |a_{30}|^2$, $\rho_{4\pm1} = |a_{4\pm1}|^2$, and $\rho_{40} = |a_{40}|^2$. Fig. 3 shows a diagram of the atomic populations, with cesium being used in this example.

a) After the first state selector, only the atoms with populated $|3, m_F\rangle$ states are transmitted. The populations are and $\rho_{3\pm1} = \rho_{30} = 1$ and $\rho_{4\pm1} = \rho_{40} = 0$.
b) The first microwave Ramsey pulse produces a coherent state-superposition between the states $|3, 0\rangle$ and $|4, 0\rangle$, with $\rho_{3\pm1} = 1$, $\rho_{4\pm1} = 0$, and $\rho_{30} = \rho_{40} = 0.5$.
c) The first rf Ramsey pulse produces several coherent state-superposition between the following states: $|3, -1\rangle$ and $|3, 0\rangle$, $|3, 1\rangle$ and $|3, 0\rangle$, $|4, -1\rangle$ and $|4, 0\rangle$, and $|4, 1\rangle$ and $|4, 0\rangle$. The resulting populations are $\rho_{3\pm1} = 1 - 3\epsilon/8$, $\rho_{30} = 0.5 + \epsilon/4$, $\rho_{4\pm1} = \epsilon/8$ and $\rho_{40} = 0.5 - \epsilon/4$. The value of ϵ is related to the amplitude of the rf magnetic field and is kept small (of the order of δ_R/δ_Z).
d) The second rf Ramsey pulse, is phase shifted by $\pi/2$ (that is, the rf frequency tuned to points *a* or *b* on Fig. 2) and produces more population transfers between the states: $|3, -1\rangle$ and $|3, 0\rangle$, $|3, 1\rangle$ and $|3, 0\rangle$, $|4, -1\rangle$ and $|4, 0\rangle$, and $|4, 1\rangle$ and $|4, 0\rangle$. The resulting populations are $\rho_{30} = 0.5 + \epsilon R(f_C - f_Z)$, $\rho_{4\pm1} = \epsilon R(f_C - f_Z)/2$ and $\rho_{40} = 0.5 - \epsilon R(f_C - f_Z)$, where $R(f_C - f_Z)$ is the Ramsey function plotted in Fig. 2.
e) The second microwave Ramsey pulse produces more population transfer between the states $|3, 0\rangle$ and $|4, 0\rangle$, resulting in the populations $\rho_{4\pm1} = \epsilon R(f_C - f_Z)/2$ and $\rho_{40} = R(f_{MW} - f_{00})$.
f) After the second state selector, only the atoms with populated $|4, m_F\rangle$ states are transmitted, resulting in the populations $\rho_{3\pm1} = \rho_{30} = 0$, $\rho_{4\pm1} = \epsilon R(f_C - f_Z)/2$, and $\rho_{40} = R(f_{MW} - f_{00})$.

Finally, the detector measures the total population resulting in a signal:

$$S = \sum_{F,m_F} \rho_{F,m_F} = R(f_{MW} - f_{00}) + \epsilon R(f_C - f_Z). \quad (1)$$

Without rf-pulses (e.g. $\epsilon = 0$), the detected signal would be $S_0 = R(f_{MW} - f_{00})$.

The feedback signal S uses the function $R(f)$ near the two operating points *a* and *b* (Fig. 2). A first order Taylor expansion of $R(f)$ gives the linear relation $R(f) = 1/2 + \beta(\Delta_f \pm f)$, where Δ_f is the offset frequency, the "+" sign is for point *a* and "-" is for point *b*. In order to compensate the frequency shift of the 0-0 transition, the contribution from the rf Ramsey signal is made negative by using a modulation which is out of phase with respect to the microwave modulation (i.e. when the microwave frequency is at point *a*, the rf-frequency is at point *b*, and vice versa.) The signal is then

$$S = \pm\beta[(f_{MW} - f_{00}) - \epsilon(f_C - f_Z)] + (1 + \epsilon)(1/2 + \beta\Delta f).$$

Using $f_{00} = C_1 B_0^2 + \delta_R$ and $f_Z - f_C = C_2 \langle B(z(t))\rangle - C_2 B_0 = \delta_Z = \delta_R C_2/(2B_0C_1)$, the signal is:

$$S = \pm\beta[f_{MW} - \delta_R(1 - \epsilon C_2/(2B_0C_1))] + K,$$

where K is a constant. The servo control modifies f_{MW} in order to keep the error signal S constant. The value of δ_R depends on the C-field, but with the appropriate choice $\epsilon = 2B_0C_1/C_2$, δ_R is multiplied by zero removing the dependence of the signal S on the value of the C-field.

Thus, by selecting the amplitude of the rf-pulses so that $\epsilon = 2B_0C_1/C_2$, the frequency deviation of the rf Ramsey signal will cancel the frequency deviation of the microwave Ramsey signal. The small value of $\epsilon < 0.0002$ needed for this method is to be contrasted to the large population transfer (effectively 100% or $\epsilon = 1$) used in [8] where the $m_F = 0$, ±1, ±2, etc. states are completely mixed by large rf-signals.

With the appropriate selection of the rf-pulses amplitude, the rf Ramsey signal added to the microwave Ramsey fringes

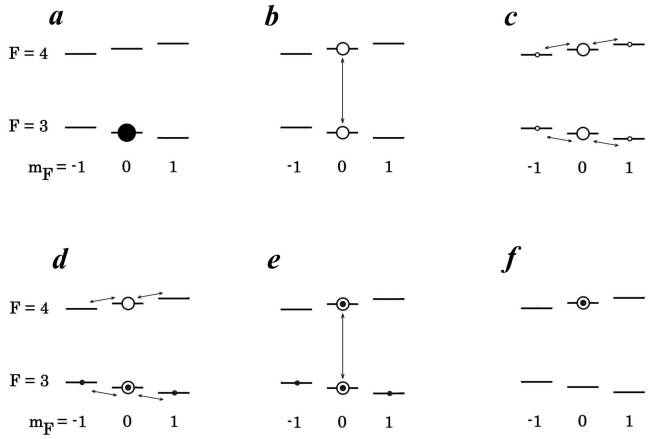


Fig. 4. Diagram of the populations in a fountain clock. The populations in the relevant states are shown by the circles: a) after the first state selector, b) after the first microwave Ramsey pulse, c) after the first rf Ramsey pulse, d) after the second rf Ramsey pulse, e) after the second microwave Ramsey pulse, and f) after the second state selector.

(the error signal steering the frequency of the clock) becomes independent of C-field deviations away from the nominal value determined by the frequency of the rf-oscillations f_C. Since the C-field is sampled by each atom along its individual trajectory (as represented by $\langle B(z(t))\rangle$), spatial variations of the C-field are canceled as well as temporal variations. Thus, this method stabilizes the frequency of an atomic clock against C-field non-uniformity and temporal variations by compensating for the quadratic frequency shift of transitions of the clock. The rf magnetic pulses produce a compensation directly at the level of the frequency servo loop. No feedback based on a measurement of the C-field has to be applied to the C-field dc-current.

Excitation electrodes producing a transversal field over the entire region of the C-field are assumed in this example. However, the excitation electrodes can also produce a longitudinal field over two short sections near the microwave excitation regions.

B. Compensation in a clock with magnetic state election

Some clocks use the initial state $|3,0\rangle$ at the beginning of the cycle, e.g. in some fountain clocks. In this case, the method takes advantage of the small difference between the Zeeman frequency splitting between the $|3,m_F\rangle$ states $f_{Z3} = C_3 \langle B(z(t))\rangle$ and the splitting between the $|4,m_F\rangle$ states $f_{Z4} = C_4 \langle B(z(t))\rangle$, where $C_3 = 3509.8$Hz/μT and $C_4 = 3498.6$Hz/μT in cesium.

a) After the state selector, the populations are $\rho_{30} = 1$, and $\rho_{3\pm1} = \rho_{4\pm1} = \rho_{40} = 0$. Fig. 4.
b) The first microwave Ramsey pulse produces a coherent state-superposition between the states $|3,0\rangle$ and $|4,0\rangle$, giving $\rho_{3\pm1} = \rho_{4\pm1} = 0$ and $\rho_{30} = \rho_{40} = 0.5$.
c) The first rf Ramsey pulse produces the populations $\rho_{3\pm1} = \rho_{30} = \rho_{4\pm1} = \epsilon/8$ and $\rho_{30} = \rho_{40} = 0.5 - \epsilon/4$. The value of ϵ is related to the amplitude of the rf magnetic field and is kept small (of the order of δ_R/δ_Z).

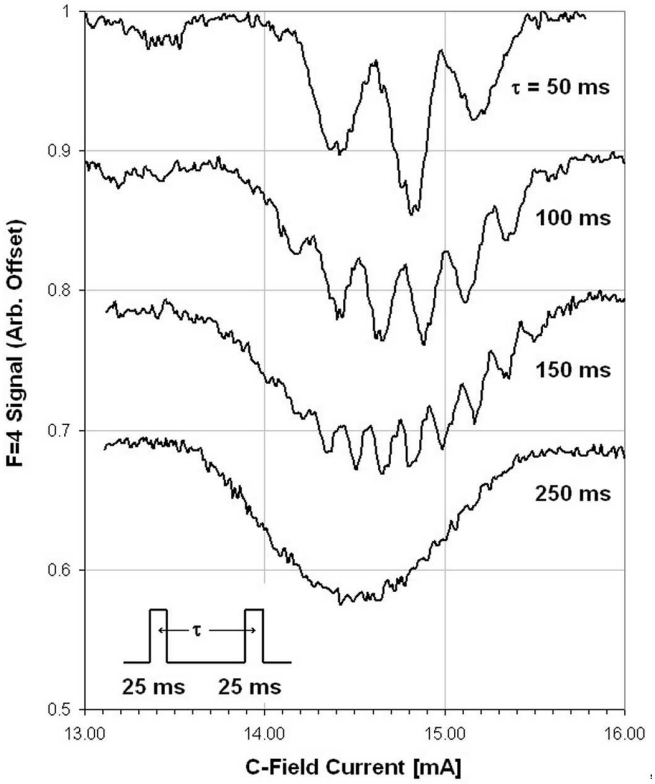


Fig. 5. Ramsey spectra of Zeeman states.

d) The second rf Ramsey pulse, is phase shifted by $\pi/2$ (that is, the rf is tuned to points *a* or *b* on Fig. 2). The resulting populations are $\rho_{30} = 0.5+\epsilon R(f_C - f_{Z3})$, $\rho_{4\pm1} = \epsilon R(f_C - f_{Z4})/2$ and $\rho_{40} = 0.5 - \epsilon R(fC - fZ4)$, where $R(f_C - f_Z F)$ is the Ramsey function plotted in Fig. 2 and f_{ZF} is f_Z calculated with the constants $C_2(F)$.
e) The second microwave Ramsey pulse produces the populations $\rho_{4\pm1} = \epsilon R(f_C - f_{Z4})/2$ and $\rho_{40} = R(f_{MW} - f_{00}) - \epsilon[R(f_C - f_{Z3}) + R(f_C - f_{Z4})]/2$.
f) After the pusher beam, $\rho_{3\pm1} = \rho_{30} = 0$ and only the atoms with populated $|4,m_F\rangle$ states are transmitted.

Finally, the detector measures the total population resulting in a signal $S = \sum_F \sum_{m_F} \rho_{F,m_F} = R(f_{MW} - f_{00}) + \epsilon[R(f_C - f_{Z4}) - R(f_C - f_{Z3})]/2$.

The feedback signal S is used in the same way as the signal of Eq. 1 with the correction term $R(f_C - f_{Z4}) - R(f_C - f_{Z3})$ providing the sensitivity on the C-field to be used as a first-order correction to the second-order Zeeman shift. One has to be careful to select an appropriate launch height/C-field combination for the technique to work.

IV. Measurement of the population transfer in a fountain clock

A measurement of the Ramsey fringes produced by rf magnetic pulses was made to demonstrate the technique. The launch height of the atoms was selected at 15cm above the Ramsey cavity. A pulse duration of 25ms was used with the rf pulse-pairs separated by different time intervals τ. Instead

of tuning the frequency of the rf-pulses, the current in the C-field electrodes was tuned to simulate a change in the value of the C-field. The results are shown in Fig. 5 for different times between the Ramsey pulses. As expected, the fringes become narrower as the time interval is increased. A shift of the central fringe with different time interval τ is the result of some C-field non-uniformity.

The contrast to the fringes decreases to zero when the time interval reached 250ms. This is due to the fact that with a C-field of 1.7mGauss, the Zeeman frequency splitting of the $F = 3$ states is 596.6Hz and the Zeeman frequency splitting of the $F = 4$ states is 594.7Hz. The ~ 2.0Hz difference implies that the states are in phase every 500ms, but are out of phase after 250ms.

V. Conclusion

The method presented in this paper improves the long term stability of a clock without adjustments of the C-field current. The rf magnetic oscillations produce a compensation directly at the level of the frequency servo loop. No feedback based on a measurement of the C-field is required to adjust the C-field dc current. The method also corrects for the spatial non-uniformity of the C-field since each atom individually probes the C-field along their respective trajectories. The method may be implemented in any atomic clock e.g. a fountain clock, a beam clock and in some cases, a hydrogen maser. The implementation is relatively straightforward and inexpensive. The long term stability of the microwave synthesizer is not compromised as this method does not require any modification of the microwave system. The method uses magnetic pulses in the audio frequency range of a few hundred Hertz which can be produced in most existing clocks using already existing coils. The stability of a clock is improved by a factor determined by the amplitude stability of the rf magnetic pulses.

References

[1] A. De Marchi, "The high C-field concept for an accurate cesium beam resonator," Proc. 7th European Frequency and Time Forum, Neuchâtel, Switzerland, March 1993.

[2] E. Rubiola, A. Del Casale, A. De Marchi, "A Dual Frequency Synthesis Scheme for a High C Field Cesium Resonator," Proceedings of the 47th IEEE International Frequency Control Symposium, 105, 1993.

[3] A. Stern, B. Levy, M. Bootnik, E. Detoma, G. Pedrotto, "Rubidium Frequency Standard with a high Resolution Digital Synthesizer." Proceedings of the 46th IEEE International Frequency Control Symposium, 108, 1992.

[4] A. Lepek, A. Stern, "Elimination of Magnetic Influence on Atomic Clocks," United States Patent 4,953,148 issued August 28, 1990.

[5] J. Rabian, J.-C. Potet, M. Buchon, "Atomic Clock," United States Patent 4,943,955 issue July 24, 1990.

[6] A. Stern, B. Levy, "Atomic Frequency Standard and System Having Improved Long Term Aging." United States Patent 6,614,321 issued September 2, 2003.

[7] J. A. Barnes, E. A. Rodrigo, "Fine Tuning of Atomic Frequency Standards," United States Patent 4,740,761 issued April 26, 1988.

[8] W. Happer, Y. Y. Jau, F. Gong, "Method and system for operating an atomic clock with simultaneous control of frequency and magnetic field." United States Patent 7,439,814 issued October 21, 2008.

[9] A. Clairon, S. Ghezali, G. Santarelli, P. Laurent, S. Lea, M. Bahoura, E. Simon, S. Weyers, and K. Szymaniec, "Preliminary accuracy evaluation of a cesium fountain frequency standard," in Proc. 5th Symp. Frequency Standards and Metrology. Singapore: World Scientific, 49, 1996.

[10] S. R. Jefferts, J. Shirley, T.E. Parker, T.P. Heavner, D.M. Meekhof, C. Nelson, F. Levi, G. Costanzo, A.D. Marchi, R. Drullinger, L. Hollberg, W.D. Lee, and F.L. Walls, "Accuracy evaluation of NIST-F1," Metrologia, Vol. 39, No. 4, 321, 2002.

[11] N. F. Ramsey, "A Molecular Beam Resonance Method with Separated Oscillating Fields," Physical Review, Vol. 78, Issue 6: 695-699; June 15, 1950.

Alternative Variables for Computing Sideband Pulling in Atomic Frequency Standards

Jon H. Shirley
National Institute of Standards and Technology
325 Broadway, Boulder, CO 80305

Abstract—Sideband pulling has long been considered a possible source of frequency bias in atomic frequency standards. We treat a two-level system excited by coherent radiation with amplitude and phase modulation. We create alternative variables to facilitate the second-order solution of the associated time-dependent Schrödinger equation. We extend earlier work to include time-dependent excitation and initial phase dependence.

I. Introduction

Frequency shifts due to the presence of sidebands on the exciting radiation have been of concern from the earliest days of atomic frequency standards [1]. Results valid for sidebands further removed from the carrier than the Rabi width were given long ago [2,3]. In 1978 Audoin et al. published a second-order result for the shift due to a single sideband valid for an arbitrary sideband separation [4]. Although this theory predicted a first-order shift, it was assumed to average to zero for the continuous signals from the atomic beam frequency standards of the time. Their result has been in use ever since to determine sideband intensity limits necessary to avoid a shift larger than other uncertainties.

With the advent of cesium fountains operating in pulsed modes, the possibility of modulation in synchronism with the cycling of the fountain signals was considered [5]. A deliberate application of synchronized phase modulation produced significant shifts in the primary standard NIST-F1 [6]. An accompanying theory agreed with the experimental results.

The purpose of the present paper is to extend the second-order results of Audoin et. al. to include contributions from synchronized pulsed operation and from two sidebands, equally spaced from the carrier, with arbitrary relative amplitudes and phases. Such sidebands may be re-expressed as the simultaneous application of amplitude and phase modulation at the same modulation frequency. We adopt this latter representation since amplitude and phase modulations interact with the dynamics of a two-level system in different manners. The primary focus of our analysis will be the dependence of any shifts on the initial phases of the modulations.

II. Background Theory

We begin with a two-level system excited by an electromagnetic field proportional to

$$B=\left[b+b_a\cos\left(\Omega t+\varphi_a\right)\right]\cos\left[\omega t-\left(b_p/b\right)\sin\left(\Omega t+\varphi_p\right)\right],$$

where $2b$ is the Rabi frequency for the excitation, and b_a and b_p are the corresponding Rabi frequencies for amplitude and phase modulation. These Rabi frequencies may be constant, or have time dependence. For half-sine-wave excitation we have $b(t)=(\pi/2)b_0\sin(\pi t/\tau)$. We assume b_a and b_p have the same time dependence as b. For a single sideband $b_a = b_p$ and $\varphi_a - \varphi_p = 0$ or π. When the rotating wave approximation is made, and the probability amplitudes are redefined by appropriate time-dependent phases to eliminate the rapid ωt time dependence [3,7], the resulting time-dependent Schrödinger equation for the system becomes

$$i\hbar\frac{d}{dt}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\hbar\begin{pmatrix}-\Delta & B_{\text{rot}}{}^*\\ B_{\text{rot}} & \Delta\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}. \qquad (1)$$

Here $\Delta=\frac{1}{2}(\omega-\omega_0)$ is half the detuning from the atomic resonance ω_0 and * denotes the complex conjugate. The coupling by the field is now represented by the complex quantity

$$B_{\text{rot}}=\left(b+b_a\cos\theta_a\right)\exp\left[-i\left(b_p/b\right)\sin\theta_p\right].$$

We have introduced the abbreviations $\theta_a=\Omega t+\varphi_a$ and $\theta_p=\Omega t+\varphi_p$. We impose on (1) the initial conditions at the beginning of excitation $\alpha(t=0)=1$ and $\beta(t=0)=0$. Then α is the probability amplitude that the system remains in its initial state and β is the probability amplitude that the system changes state. The Rabi transition probability is then $P(t)=|\beta(t)|^2$.

The corresponding Ramsey excitation probability is

$$P_{\mathrm{Ram}} = \left|\beta_{\mathrm{Ram}}\right|^2 = \left|\beta^{(2)}\right|^2 \left|\alpha^{(1)}\right|^2 + \left|\alpha^{(2)}\right|^2 \left|\beta^{(1)}\right|^2 + 2\,\mathrm{Re}\left(\gamma e^{2i\Delta T}\right),$$

where T is the period between the two excitations, and the coefficient γ is defined as

$$\gamma = \alpha^{(2)}\beta^{(2)}\alpha^{(1)}\beta^{(1)}\,*. \qquad (2)$$

The superscripts refer to the first and second excitation regions, respectively. The two excitations are assumed identical, except that the initial phase of both modulations has advanced by $\Phi \equiv \Omega(T+\tau)$ for the second excitation.

The presence of the perturbing modulations causes a frequency shift given by

$$2\Delta_{\mathrm{res}} = \frac{\mathrm{Im}\,\gamma}{T\,\mathrm{Re}\,\gamma},$$

If we expand γ, defined by (2), in powers of the perturbation (subscripts) to second order, we find

$$2\Delta_{\mathrm{res}} = \frac{\mathrm{Im}\,\gamma_1}{T\,\mathrm{Re}\,\gamma_0} + \frac{\mathrm{Im}\,\gamma_2}{T\,\mathrm{Re}\,\gamma_0} - \frac{\mathrm{Im}\,\gamma_1\,\mathrm{Re}\,\gamma_1}{T\left(\mathrm{Re}\,\gamma_0\right)^2}. \qquad (3)$$

The imaginary part of γ_0 was omitted because it is an odd function of Δ and merely makes a small change in the value of T in the denominator (cf. Sec. 3.2 in [7]). Note that the first-order correction to the denominator makes a contribution in the last term. To find the perturbation expansion of γ we need the perturbation expansions of α and β. These are difficult to derive from (1).

III. Alternative Variables

To simplify the derivation of the perturbation expansion we define alternative variables as follows:

$$W(t) = e^{ia(t)}\left[\mathrm{Re}\,\alpha(t) + i\,\mathrm{Im}\,\beta(t)\right]$$
$$\text{and } Z(t) = e^{-ia(t)}\left[\mathrm{Re}\,\beta(t) - i\,\mathrm{Im}\,\alpha(t)\right],$$

where a(t) is defined as the integral of b(t). The variables paired in the definitions are those coupled by the excitation. The exponentials in the definitions incorporate the unperturbed, on-tune evolution of the two level system. These new variables obey the same normalization criterion as α and β:

$$\left|W(t)\right|^2 + \left|Z(t)\right|^2 = 1.$$

In terms of these variables the Rabi transition probability becomes

$$P = \tfrac{1}{2} - \tfrac{1}{2}\left[\mathrm{Re}\left(e^{-2ia}W^2\right) - \mathrm{Re}\left(e^{2ia}Z^2\right)\right].$$

These alternative variables obey the following two differential equations:

$$dW/dt = -i\Delta e^{2ia(t)}Z + ib_a\cos\theta_a(t)W - b_p\sin\theta_p(t)e^{2ia(t)}Z$$

and

$$dZ/dt = -i\Delta e^{-2ia(t)}W - ib_a\cos\theta_a(t)Z + b_p\sin\theta_p(t)e^{-2ia(t)}W$$

where $a(t)$ may be defined by a third simultaneous differential equation,

$$da/dt = b(t).$$

If these differential equations are multiplied by i, they form a Schrödinger-like pair with a two-by-two hermitian Hamiltonian. In the absence of detuning and the perturbation, the right-hand side vanishes. The variables W and Z then remain constant, equal to their initial conditions $W(0)=1$, $Z(0)=0$. Solutions in powers of the perturbation reduce to straightforward integrals.

IV. First-Order Shifts

For small shifts of the Ramsey resonance we may ignore the detuning in the excitation regions. The first-order solution for W is then due only to amplitude modulation,

$$W_1(\tau) = -i\int_0^\tau b_a(t)\cos\theta_a(t)\,dt,$$

and for Z is due only to phase modulation:

$$Z_1(\tau) = \int_0^\tau b_p(t)\sin\theta_p(t)e^{-2ia(t)}dt.$$

Although the expressions for the real and imaginary parts of γ in terms of the alternative variables are complex, their perturbation expansion is simpler, because the zero-order values of W and Z are their initial conditions. Thus we find

$$\mathrm{Im}\,\gamma_1 = -\tfrac{1}{2}\sin 2a\left[\mathrm{Re}\left(Z_1^{(1)}\right) - \mathrm{Re}\left(e^{2ia}Z_1^{(2)}\right)\right],$$

where the argument of a is understood to be the excitation period τ. The real part of γ_0 is $\sin^2 a\cos^2 a$. When we insert the integrals, apply a trigonometric identity, and substitute into (3), we obtain the first-order shift

$$\delta\omega_1 = \frac{-4}{T\sin 2a}\left[I_{CC}\cos a\sin\tfrac{1}{2}\Phi + I_{SS}\sin a\cos\tfrac{1}{2}\Phi\right]\cos\left(\tfrac{1}{2}\Phi + \varphi_p\right). \qquad (4)$$

The integral I_{CC} is given by

$$I_{CC} = \int_0^{\tau} b_p(t)\cos\Omega(t-\tau/2)\cos\left[2a(t)-a(\tau)\right]dt.$$

The integral I_{SS} is the same except that both cosines are replaced with sines. This result is valid for excitations symmetric about their half-way point. For asymmetric excitation additional integrals are required. The integrals can be done analytically for constant excitation to yield the previously published result [6]. They can be easily done numerically for half-sine-wave excitation.

The first-order shift given by (4) has a sinusoidal dependence on the initial phase and a rapid oscillation with the sideband offset due to the phase advance Φ between excitation regions. Figure 1 shows the envelope of this rapid oscillation for both constant and half-sine-wave excitation when the final cosine is replaced by unity.

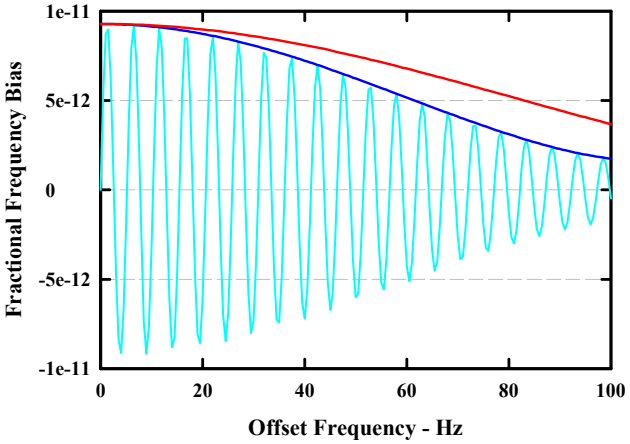


Figure 1. First-order sideband pulling. The light blue curve is for rectangular excitation. The dark blue curve is its upper envelope. The red curve is the envelope for half-sine-wave excitation.

V. Second-Order Shifts

The second-order shift can be found in a similar, if more laborious, fashion. The imaginary part of γ_2 has contributions from the second-order solution for Z and the product of first-order contributions from Z and W. These contributions involve the product of $\cos\theta_a$ and $\sin\theta_p$ at different times. A trigonometric identity allows this product to be written

$$2\cos\theta_a(t)\sin\theta_p(t') = \sin\left(\Omega t + \Omega t' + \varphi_a + \varphi_p\right) - \sin\left(\Omega t - \Omega t' + \varphi_a - \varphi_p\right).$$

Therefore the second-order shift will depend only on the sum and difference of the initial phases and not on their individual values. When the computations are carried out, the shift from the second term in (3) has the form

$$\delta\omega_2 = \frac{b_a b_p}{Tb^2}\left[\begin{array}{l} Q_0\cos\left(\phi_a-\phi_p\right)+\left(Q_C\cos\Phi+Q_S\sin\Phi\right)\cos\Theta \\ +\left(R_C\cos\Phi+R_S\sin\Phi\right)\cos\left(\phi_a-\phi_p\right)+R_0\cos\Theta \end{array}\right],$$

where $\Theta = \Phi + \Omega\tau + \varphi_a + \varphi_p$. This shift is due to the combined action of phase and amplitude modulation. The coefficients are sums of integrals depending on $a(\tau)$ and $\Omega\tau$, but of order unity. The Q coefficients arise from the interactions within each individual excitation region. The R coefficients arise from the interference between the first-order interactions in the two excitation regions. The terms with coefficients Q_C, Q_S, and R_0 do not contribute to the shift when the initial phases are averaged. But the remaining terms survive when φ_a - φ_p is constant. The Q_0, R_C, and R_S coefficients are the ones computed for constant excitation by Audoin et. al. [4]. We have verified their results except that they did not include τ in Φ. The power dependence of their results was discussed in [5]. We note that these terms vanish if $\varphi_a - \varphi_p = \pm\pi/2$, the same condition that makes the two sidebands have equal intensity.

For atomic beam standards the last term in (3) vanishes, since the initial phase dependence in $\mathrm{Re}\gamma_1$ averages to zero. However, if this term is added in, as may be needed for pulsed standards, the Q coefficients are modified, but the R coefficients cancel out. Thus if the shift is then averaged over initial phases, only the Q_0 term survives and there is no Ramsey-like oscillation in the shift. The modified Q_0 term is plotted in figure 2 for both constant and half-sine-wave excitation.

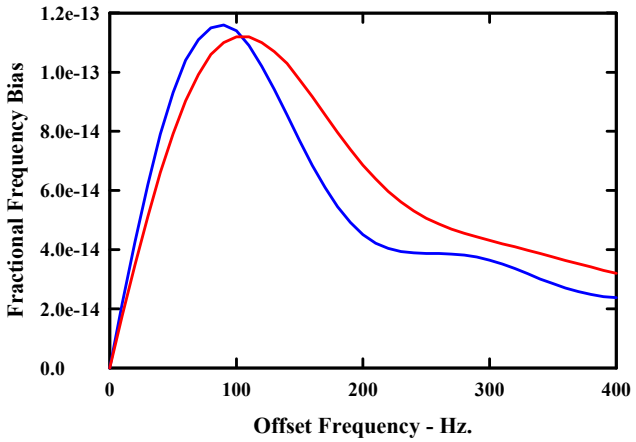


Figure 2. Second-order sideband pulling—initial phase-independent part. The blue curve is for constant excitation. The red curve is for half-sine-wave excitation.

SUMMARY

We have theoretically studied sideband pulling in atomic frequency standards in terms of amplitude and phase modulation. We have introduced alternative variables to simplify a perturbation treatment of the modulation. We find the first-order pulling, generated by phase modulation, for both constant and half-sine-wave excitation. We find the second-order pulling is generated by the combined action of phase and amplitude modulation. We report the initial-phase and Ramsey-time dependence of the second-order shift.

REFERENCES

[1] J. H. Holloway and R. F. Lacey, "Factors which limit the accuracy of cesium atomic beam frequency standards," Proc. Int. Conf. on Chronometry, Lausanne, June 1964, pp. 317-321.

[2] N. F. Ramsey, "Resonance transitions induced by perturbations at two or more different frequencies," *Phys. Rev.,* vol. 100, pp. 1191-1194, 1955.

[3] J. H. Shirley, "Some causes of resonant frequency shifts in atomic beam machines. I. Shifts due to other frequencies of excitation," *J. Appl. Phys*., vol. 34, pp. 783-788, 1963.

[4] C. Audoin, M. Jardino, L. S. Cutler, R. F. Lacy, "Frequency offset due to spectral impurities in cesium-beam frequency standards," *IEEE Trans. Instrum. Meas.,* vol. 27, pp. 325-329, 1978.

[5] F. Levi, J. H. Shirley, T. P. Heavner, D.-H. Yu, S.R. Jefferts, "Power dependence of the frequency bias caused by spurious components in the microwave spectrum in atomic standards," *IEEE Trans. UFFC,* vol. 53, pp. 1584-1589, 2006.

[6] J. H. Shirley, T. P. Heavner, S. R. Jefferts, "First-Order Sideband Pulling in Atomic Frequency Standards," *IEEE Trans. Intrum. Meas.*, vol.58, pp. 1241-1246, 2009.

[7] J. H. Shirley, W. D. Lee, R. E. Drullinger, "Accuracy evaluation of the primary frequency standard NIST-7," *Metrologia,* vol. 38, pp. 427-458, 2001.

Study of a High Performance Rubidium Atomic Frequency Standard

Da Zhong , Ganghua Mei
Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences
Wuhan 430071, the People's Republic of China
Email: zhongda@wipm.ac.cn

Abstract -- **In this paper, the study advances of a high performance rubidium atomic frequency standard (RAFS) will be reported. The RAFS is a traditional passive rubidium frequency standard. Some important performance test results have been obtained. The RAFS has realized short-term stability of <1.5E-12/1s , and long-term stability of <2E-14/1d**.

I. INTRODUCTION

The investigations of a high performance rubidium atomic frequency standard (RAFS) will be presented. This RAFS is a traditional passive gas-cell rubidium frequency standard. While its physics package employs a novel slotted-tube microwave cavity [1], and uses the separated filter technique (SFT). This kind of physics package features high signal-to-noise ratio, and low light frequency shift [2].

We have made a lot of experiments and tests with the RAFS unit, such as frequency stability, drift, phase noise, harmonics, spurious, magnetic sensitivity, and so on. Some important RAFS specifications have been obtained. The recent test results suggest that RAFS has realized short-term stability of <1.5E-12/1s , long-term stability of <2E-14/1d, and frequency drift of <3E-13/1d.

II. PRINCIPLE AND CHARACTERISTIC

Figure 1 shows a principle block diagram for our rubidium atomic frequency standard. The design of it is a typical gas-cell RAFS schedule. The physics package acts as a very stable frequency discriminator for a frequency around 6.834GHz. By using a frequency synthesizer and two frequency multipliers that are referenced to the 10MHz VCXO, the 10MHz output frequency can be stabilized to the rubidium hyperfine transition frequency.

The novel physics package mainly consists of cavity-cell assembly [3], spectral lamp, C-field and temperature controller, microwave frequency multiplier, and preamplifier. The cavity-cell assembly employs a patented slotted-tube microwave cavity, and utilizes the separated filter technique (SFT) .The slotted-tube microwave cavity has some advantages of higher cavity Q factor, large microwave filling factor, and can obtain a higher signal-to-noise ratio for the physics package.

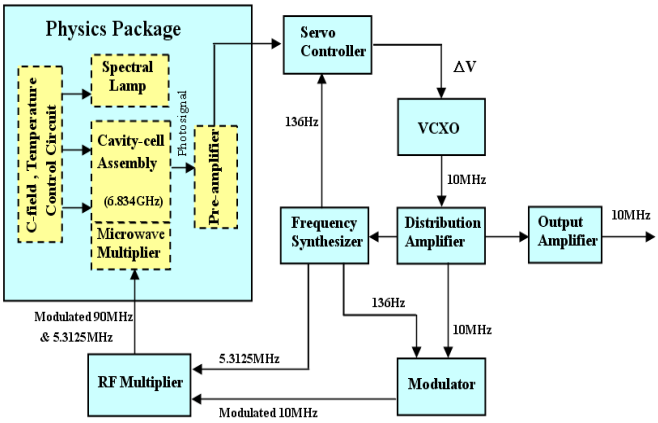


Figure 1. RAFS Principle Block Diagram

Light shift (LS) and temperature coefficient (TC) are the most significant factors that affect the long-term stability of a RAFS. The SFT design of the physics package enables us to deal with the gas parameters of absorption cell and filter cell separately, and buffer gas pressures in both cells can be optimized for the minimum LS and TC.

The spectral lamp is a RF-discharge lamp for optical pumping and detecting, it operates at about 110MHz. The

photosignal is amplified and filtered before being delivered to servo control circuit.

The servo controller is responsible for, synchronously detecting the amplitude and phase of the photosignals at 136Hz and 272Hz, and filtering the error signal to lock the 10MHz VCXO to the rubidium hyperfine transition.

Both the 5.3125MHz signal and the 136Hz phase modulation waveform are generated by the synthesizer. The modulated 10MHz signal is converted to 90MHz signal by RF multiplier. The two signals of 90MHz and 5.3125MHz are input to the microwave multiplier together.

The microwave multiplier is a SRD high-order harmonic frequency multiplier to complete 76th frequency multiplication from 90MHz to 6840MHz, meanwhile it also plays a role of frequency mixer to make frequency subtraction between 6840MHz and 5.3125MHz. Finally it will produce 6834.6875MHz microwave field in the cavity for exciting ^{87}Rb hyperfine transition. The microwave frequency multiplier has no independent resonating cavity, but shares the same slotted-tube cavity with ^{87}Rb atom transition microwave resonator.

III. MAIN PERFORMANCES TEST

After the RAFS unit has been completed, and its operating parameters have been optimized. We conducted some performance tests with RAFS unit.

The short-term frequency stability performance of the RAFS has been reported on the EFTF-IFCS-2009 conference [3]. In this paper, the long-term stability test result will be presented. Figure 2 shows the set-up for RAFS long-term frequency stability measurement.

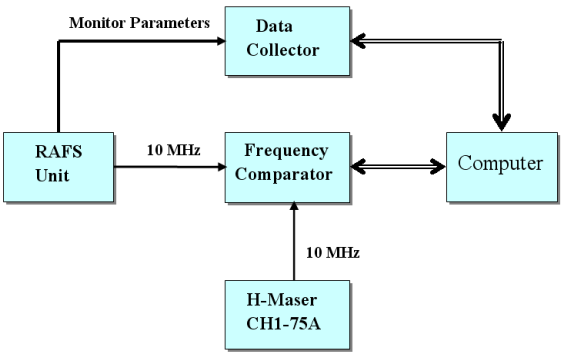


Figure 2. Set-up for RAFS Frequency Stability Test

The reference source used for RAFS test is the KVARZ active hydrogen maser CH1-75A which has an excellent frequency stability of <1E-15/1d. The frequency curve of RAFS is showed in Figure 3, about 78 days test data were collected with a sample period of 10 seconds, from 24 May 2010 to 12 August 2010. The frequency stabilities were calculated with the last 40 days test data in Hadamard deviation. The stability curve of Figure 4 indicates that the RAFS unit has achieved long-term frequency stability of 2.6E-14/10000s and 1.3E-14/1d, while its frequency flicker floor is around of 1.2E-14 level.

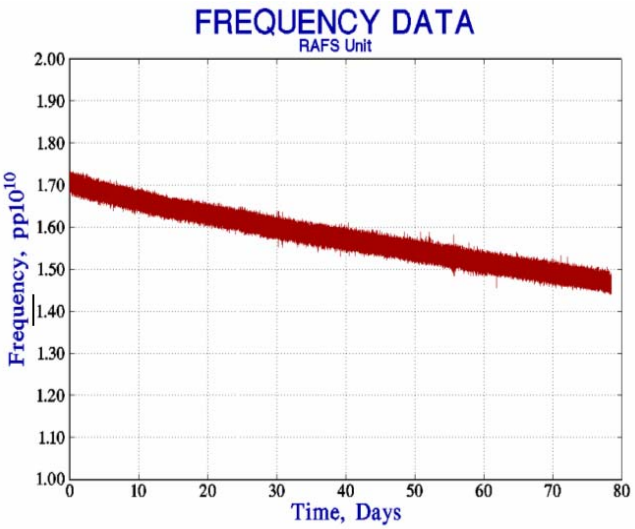


Figure 3. RAFS Frequency Test Curve

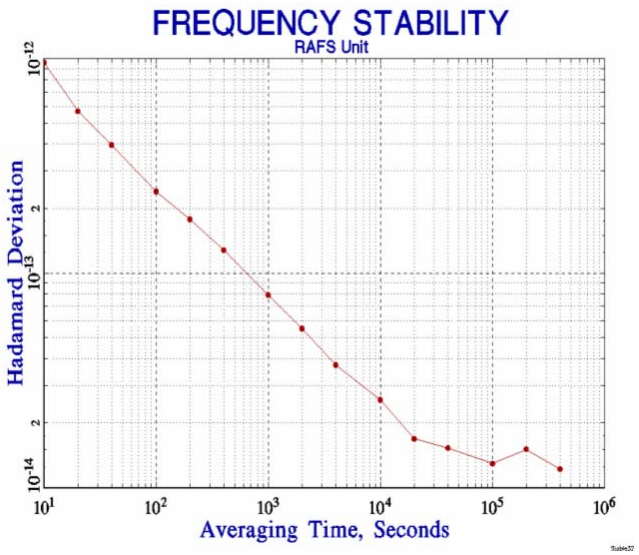


Figure 4. Frequency Stability of RAFS Unit

Based on calculation of the last 20 days data, the frequency drift is 2.5E-13/1d, which still tends to decrease as RAFS unit continue to work. It can be estimated that the drift will finally arrive to 1E-13/1d level or better.

Figure 5 presents phase noise curve of RAFS unit, measured by PN9000 phase noise test system. The measured results suggest that RAFS has a low phase noise, from 1Hz (<-95dBc) to 10kHz (<-147dBc).

The RAFS unit has obtained a low magnetic sensitivity of 1.2E-13/Gauss, benefits from the microwave cavity shielding scheme of three layers mu-metal.

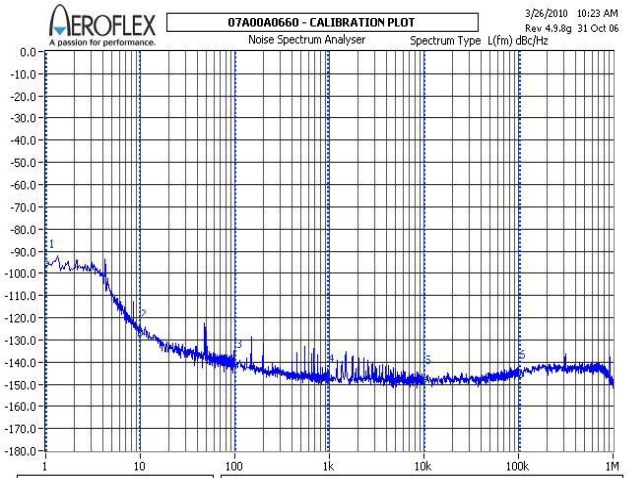


Figure 5. Phase Noise of RAFS Unit

IV. CONCLUSION

We have developed a gas-cell rubidium atomic frequency standard with a novel physics package. The performance test results indicate that the RAFS unit has achieved good frequency stabilities, the short-term stability is of $<1.5E\text{-}12/\tau^{1/2}$ ($1s \leq \tau \leq 1000s$), the long-term stability is of <2.0E-14/1d, and its frequency flicker floor is around of 1.2E-14 level. The RAFS unit also has low phase noise and low magnetic sensitivity.

REFERENCES

[1] G. Mei, D. Zhong, S. An, J. Liu and X. Huang, "Miniaturized Microwave Cavity for Atomic Frequency Standard", U. S. Patent, No. 6 225 870 B1, May 1, 2001.

[2] B. Xia, D. Zhong, S. An and G. Mei, "Characteristics of a Novel Kind of Miniaturized Cavity-Cell Assembly for Rubidium Frequency Standards"，IEEE Transactions on Instrumentation and Measurement, Vol.55, No.3, June 2006, pp. 1000-1005.

[3] Da Zhong, Baihua Xia, Shaofeng An, et al. ,"Investigation on Physics Package with Slotted-Tube Microwave Cavity for Rubidium Atomic Frequency Standard"， Proceedings of the EFTF-IFCS-2009 Joint Conference， pp. 1019-1022, April 2009.

An Ultra-Miniature Rb Atomic Clock Compatible with OCXO

Yuanhong Cao, Lin Yang, Qifang Li, Haiqing Zhao, Qiaoli Lin, Runchang Du
Spaceon Co. Ltd., Chengdu, China
thomascyh@hotmail.com

Abstract—**This paper shows a traditional Rubidium atomic clock – XHTF1031. The incorporate lamp and ultra-miniature waveguide cavity introduced in the physics package is the main means to downsize the volume of the clock; the electronics consisting of 6.8GHz direct synthesizer and digital servo system is to enhance the automatic functions which ensure massively fabrication available with low costs. At all, the clock has the physical dimensions of a ovenized crystal oscillator (OCXO), measuring 51mmx51mm (2'x2') and standing at a mere 25mm (1'). The Allan deviation of instability is 3E-11/s; the temperature coefficient is 3E-10 in the range of -20~60℃; the frequency aging is below 5E-12/d. The clock has the same dimension and pins' definitions with OCXO used in telecom, but higher performance, and would be popularly used in telecom equipments and base-stations instead of OCXO in future.**

I. INTRODUCTION

Traditional Rubidium atomic clock with a combination of high performance and volume has large engineer value and hearty life until now. Comparing with Hydrogen clock and Cesium clock, compact rubidium atomic clock with merits in aspect of low power, small dimension and excellent precession, is one of the most popular atomic clocks, and bulkily used in Telecom, IP-TV, time-keep system and the Global Navigation Satellite Systems (GNSS), Especially IEEE1588 protocol offering several nanoseconds of web synchronization [1]. In the above area, compact atomic clock will used instead of OCXO partly.

Rubidium clock uses optical pumping by an Rb spectrum lamp to create a population inversion between the ground state hyperfine energy levels F=2 and F=1 in absorption cell [2]. If the atoms of absorption cell are interrogated with microwave radiation, the change in light transmission through the cell can be checked by optical diode, and a discriminating signal in the light is to lock OCXO in the atomic clock. Employing the two naturally-occurring isotopes ^{85} Rb and ^{87}Rb is to enhance the efficiency of the optical pumping by a fortuitous overlap between the two optical absorption lines; and similarly, filling buffer gas (such as N2, NH4) to narrow the line width of the absorption of the clock transition between (F=2, m_F=0) and (F=1, m_F=0). Therefore, the Rb atomic clock has higher Signal-to-Noise ratio and more outstanding short-term ability. There is no vacuum system in Hydrogen clock and Cesium clock, therefore Rb clock has more viable in commercial fields.

When it comes to developing compact rubidium atomic, the most important work is to downsize the dimension of physics package of the clock, including lamp and cavity. In general, the effective means is to pay more attention to the latter than the prior, as for the cavity is the big part in Rb clock. In comparison, the prior can not be indefinitely reduced because of high temperature and exciting RF circuits located in it. Basing On traditional gas cell Rb clock scheme and different cavities, several Rb clocks as a module have been manufactured by Symmetricom, FEI, Spectratime in the world. In China, Xing Hua Time and Frequency (XHTF) joining in Spaceon at present has about forty years history in professional development of Rubidium clock. Until now, Spaceon has an outstanding achievement in miniature Rubidium clock, and designed a smallest gas cell Rb clock XHTF1031 in the world. The XHTF1031 is designed by new physics package and electronics, and has the same typical performances with the above-mentioned clocks. The detailed techniques are presented as below.

II. PHYSICS PACKAGE

A. Lamp

CLAPP oscillator has a high stability, and can be used as the driving circuit of the spectrum lamp [3]. The frequency of the oscillator is decided by L made up of a coil around the bulb, and by C which is in series with the coil. Because of high frequency scaled 100MHz, the distributing parameters will occur while the circuit is placed in the lamp oven. There is the optimal value of the L and C which need to be tuned in an extent. In order to cancel the temperature coefficient of the circuits, a diode is added in parallel with the b-e loop of transistor in the circuit, which would improve fluctuation and aging of Rb clock.

The configuration of assembling the parts to shape the architecture of the spectrum lamp mainly decides the size itself. In general, bulb and exciting circuits are deigned separately as two parts. As is known, the temperature coefficient is added by transistor in the circuits. Taking the

No. 2010XM161, principal technical term, Sichuang, China.

transistor into the case of lamp can be resolved. The spectrum lamp is called incorporate lamp. In the architecture, the bulb of lamp is held into solenoid and together inserts into a bold which is designed by Teflon which ensure the light density stable. All of these are fixed into a case which is designed by aluminum. The PCB is laid into the case simultaneously. In this way, the electromagnetic radicalization can be shielded effectively. Therefore, the power dissipation can be used to heat up the lamp, and the power of the lamp which is capable of reducing at an extent. In a word, the bottleneck in the aspect of the coefficient, the size and the power dissipation has been resolved in the incorporate lamp. The final appearance of the lamp is shown in Fig. 1.

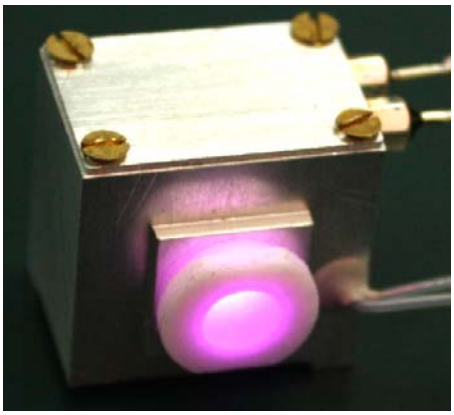


Figure 1. The lamp removed from the clock XHTF1031.

In the process of fabricating lamp, the bulbs are aged with high RF power and high temperature firstly. Secondly, the lamp will be supplied with higher voltage than it being in normal work. All the steps can ensure to fabricate lamp in bulk.

B. cavity

In Rb atomic clock, cavity is a container where absorption cell is located in. Generally, the cavity made of metal, such as copper or aluminum, is resonant with the microwave of atomic transitive frequency from synthesizer which is to interrogate atoms in the absorption cell in order to result in discrimination signal. At this rate, the volume of the resonant cavity is the biggest hamper while developing miniature Rb clock.

To reduce the size of cavity in Rubidium atomic clock, there are several types of cavity as below. The first is cylinder cavity of TE111 model, sometimes filled by Al_2O_3 ceramic [4]. Obviously, this cavity has too small Q quality to bring on poor short-term frequency instability while developing Rb clock by it. The second is magnetron cavity whose mode is similar to TE011, and has high Q quality. The cavity has been used in Hydrogen clock in advantage. But the complex architecture and difficulty in production by machine hamper its expanding to compact Rb clock.

Telecom technology brings digital synthesizer to microwave frequency band, and several synthesizers beyond 7G occur nowadays. The synthesizer has the ability to output the enough power level of 6.8GHz to interrogate atom. So, the cavity has no necessity to be resonant with atomic frequency.

Waveguide port is the avenue for microwave transmitting in a model. One port of the waveguide is enclosed as an emission source, and other port is enclosed excluding a hole laid in the center of bottom as the diaphragm the lamp light goes through. Simultaneously, absorption cell insets into the space in the cavity. In this way, one cavity is shaped in the Rb clock. While designing the cavity, the parameters including diameter and length must be calculated. Due to the boundary abnormity from absorption cell and the hole, the simulation in magnetism is necessary and effective auxiliary tool to calculate the parameters. The simulating is as in Fig. 2.

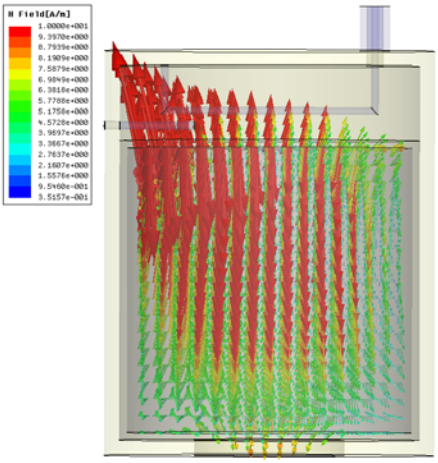


Figure 2. The Cavity of theRb atomic clock XHTF1031.

In Fig. 2, magnetic excitation is from a loop coupling the 6.8GHz microwave. In order to make excitation higher than electricity, there is a special hemicycle soldered with the coupling loop. A bolt made specially is to screw in and out to compensate the boundary abnormity in the cavity after inserting absorption cell. By the design, the tolerance of absorption cell size is bigger than the previous, and the cavity is more available to fabricate in bulk. There is no ceramic in side of the cavity; therefore the lower weight and higher performance in vibration and shock are implemented. In fact, the power level range of -20 to 0 dBm from frequency multiplier is ok to the cavity with screwing the bolt. In addition, the waveguide cavity is made of permalloy metal so as to reduce frequency shift resulted from magnetism.

III. ELECTRONICS

Electronics in Rb clock consists of frequency multiplier which is to multiply 10MHz from OCXO or TCXO to 6.8GHz, and servo which is to synchronously check and output an error to lock VCXO.

A. Frequency Multiplier

Traditional frequency multiplier is designed with step recovery diode (SRD), while SRD has the special property whose capacitance has different quality on the tradition while voltage shifts from positive to negative. 6.8GHz microwave to excite atom is from the high order harmonic wave by SRD. In this way, the resonance cavity must be used as the selective loop to enhance the 6.8GHz microwave, and Rb atoms in the absorption cell laid in the resonance cavity will output the

enough strength of discrimination signal by transmission of light intensity from lamp to lock VCXO. Owning to the configuration of frequency multiplier in Fig. 3, the above-mentioned cavity can be used in Rb atomic clock.

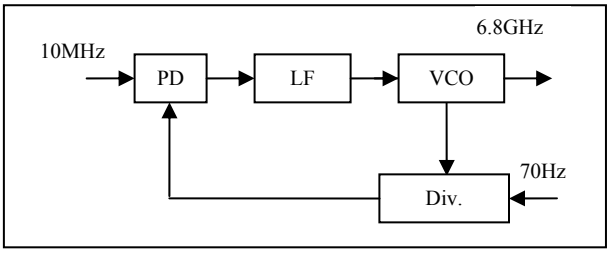


Figure 3. The frequency mitiplier scheme block.

In Fig. 3, the 10MHz from OCXO compares with the similar frequency feeding back from VCO in phase detector (PD), and the error voltage of the phase is filtered by Low Filter (LF). Finally, the 6.8GHz is locked in phase with 10MHz and modulated in frequency by 70Hz.

B. Digital Servo System

This part is embedded in a Micro-programmed Control Unit (MCU). The detailed configuration is as below in Fig. 4.

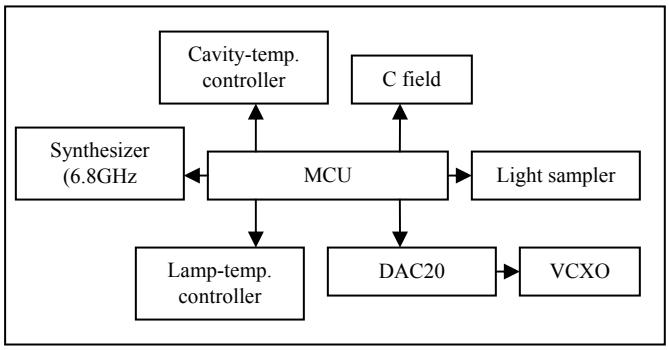


Figure 4. The digial servo system block.

- MCU is the center controller of the electronics in the Rb clock. The first function is to initialize external devices, including synthesizer, DAC20; the second is to sample the light intensity and temperature data and to process digital data to export an error to lock VCXO.
- C field is set by 12 bit DAC which can adjust the accuracy of 10MHz frequency standard signal with high resolution. Therefore, the Rb atomic clock can be disciplined by GNSS via tuning this DAC's transiting data, so the long term stability will up to the level of satellite atomic clock.
- DAC20 consists of two 12 bit DAC and is an equivalent 20 bit DAC, which is to lock the VCXO in normal work and to make the VCXO to be in holdover status in abnormal work.
- Light sampler is to gather intensity of the lamp light which contains work status of the lamp and atomic discrimination signal. The first one could assist to check the lamp's off-on statue to shift the voltage from low to high, and the second is to lock VCXO.
- Synthesizer as a frequency multiplier makes the 10MHz from VCXO to 6.834GHz directly and modulates the microwave with a low frequency 70Hz. The synthesizer is made up of a digital phase comparator, digital divider (Div.) and VCO.
- The cavity-temp. controller and lamp-temp. controller are to control the cavity temperature and lamp temperature respectively according to destination value of temperature automatically.

C. Monitor-control software

The monitor-control software is installed in the PC which communicates with Rb atomic clock with RS232 protocol. Otherwise, the clock's work status including lamp and cavity can be monitored real-timely. Furthermore, the software can also set the parameters including synthesizer's offset and divider coefficient. With the help of the software, the Rb clock can be massively produced at a low cost.

D. Archetechure design

The digital electronics and novel physics package are introduced, and the ultra-miniature Rb atomic clock can be shaped in a low size scale. Because OCXO with 51mmx51mm dimension is popularly used, as is necessary to design the similar Rb clock. Therefore the XHTF1031 has completed on this idea. The clock without case is shown in the. Fig.5. It is clear the synthesizer is enclosed in a metal shield, and the 6.8GHz transmits from the synthesizer to cavity by a microwave line with mini SMB connector.

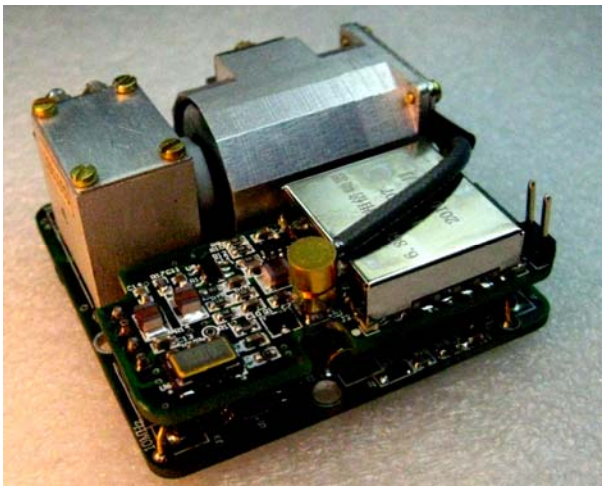


Figure 5. Spaceon Rb atomic clock XHTF1031 with the cover removed.

- The two PCBs are overlapped and fixed by 4 sockets. The frequency multiplier is enclosed in a case and directly soldered on the small PCB which is aligned with the cavity and the lamp.
- The case of Rb atomic clock is made of permalloy metal and is to reduce frequency shift resulted from magnetism again. Therefore the magnetic sensitivity will be reduced at large and finally to achieve the 2E-

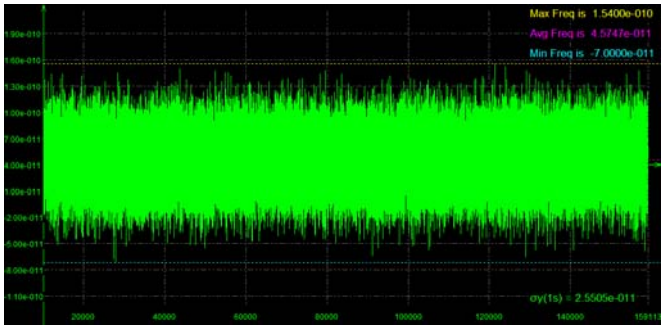
11/G. finally, the dimension is 51mm×51mm×25mm which is compatible with the ordinary OCXO. Especially, the baseplate of the clock has 4 screw hole and can be held safely, which makes the clock‘s frequency specifications tolerable on the condition of shock and vibration.

IV. TEST

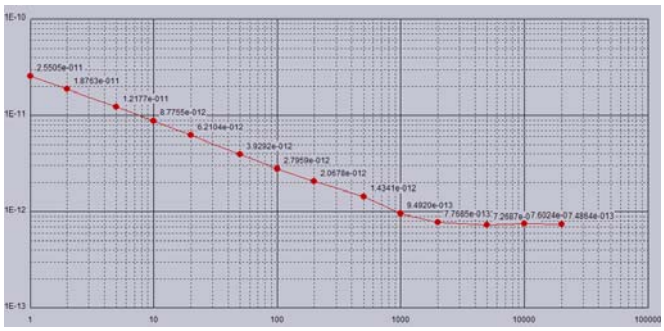
After making the key-technique design, five prototypes of the clock have been fabricated. Firstly, we have tested the five clocks comprehensively. The data and figures of testing are as below.

A. short-term frequency instability

Short-term frequency instability is relative with noise from synthesizer and physics package and the strength of atomic discrimination. In Fig.6, the short-term frequency instability of the clock XHTF1031 is given.



(a)



(b)

Figure 6. XHTF1031 short-term frequency instability. (a) The typical frequency accuracy of the clock. (b) The typical frequency Allan deviation of the clock.

In Fig.6, the short-term frequency instability is showed. The frequency data are sampled per 1 second in two day, therefore the quality of frequency point is enough to calculate the Allan deviation. From the diagram, the short-term frequency instability is about 2.6E-11/s, 9E-12/10s, 2.8E-12/100s, which is satisfied with root tao in Allan deviation, and is laid in the middle level of compact Rb clock.

Obviously, the data is better than ordinary OCXO’s, especially in the term of 10s and the time with big scale.

B. Temperature coefficient

Temperature coefficient is most difficulty in gas cell Rb clock while Doppler shift from hot atoms in absorption cell is appended in resonant frequency [5]. In addition, integrating filer technique in compact Rb clock is popularly used in order to downsize the clock’s volume and brings on bigger coefficient than the scheme of splitting filtering technique. In XHTF1031, the digital electronics is to compensate coefficient by the feeding-back error from the temperature sensor mounted the clock case. In this way, the perfect property of frequency vs. temperature is achieved finally in Fig.7 as below.

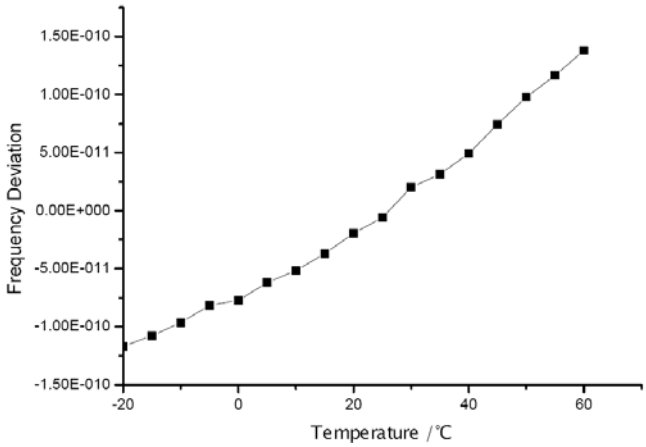


Figure 7. The clock XHTF1031coefficent in the range of -20 to 60 ℃.

The present temperature coefficient in the range of -20 to 60 ℃ is within 3E-10. Next step, the range of work temperature will be expanded to fit defense and avionics environment.

C. Aging

Owning to the bulb and absorption cell are made of a special glass which ingredients is researched by Spaceon, the aging of XHTF 1031 has the outstanding performance.

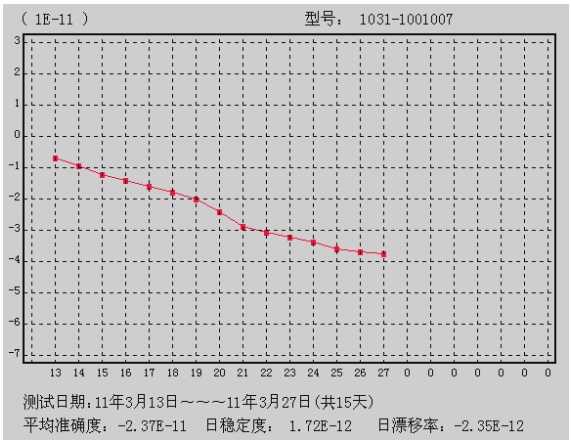


Figure 8. The clock XHTF1031’s Long-term aging test diagram.

In Fig.8, the typical data is below 5E-12/d, and better with two or more order of magnitude than ordinary OCXO's.

V. CONCLUSION

The ultra-miniature Rb atomic clock XHTF1031 basing on traditional gas cell Rb clock scheme is completed now. Furthermore, owning to introduced novel physics package and the digital electronics, the cost, power dissipation and dimension are reduced at large extent. Finally, the shape of the clock is 51mm x 51mm x 25mm (2' x 2' x 1') in the first version of the clock, which is better standing long-term frequency instability than OCXO, but compatible with it. In addition, the clock has better technique maturation than CPT clock occurring several years ago. Consequently, the clock will take on an important application position in future, and will be widely used in aspect of Global Navigation Satellite Systems (GNSS), IEEE1588 web, IP TV, Timing-keep system as frequency standard instead of OCXO with higher frequency performance two or more order of magnitude.

ACKNOWLEDGMENT

We will extremely thank Clive Green and Dr. Cosmo in Quartzlock England for their encouragement and valuable talk while developing the atomic Rb clock. Clive Green tries his best to spread Rb clock for Spaceon in the world also. In addition, some work for designing the Rb clock is done by Junqi Xiang, Xingwen Zhao, Yongjun Liu, Haitao Lin, Miao Liu, Xiaohua Pu, Xuejin Huang.

REFERENCES

[1] Mills D L. Internet Time Synchronization: the Network Time Protocol. IEEE Trans. Communications, 1991, 39(10): 1482.

[2] Audoin C., Guinot B.: The Measurement of Time, Frequency and Atomic Clock, Cambridge, University Press 2001, pp. 211-215.

[3] Jeff D. Crockett, Fullerton, Calif Lamp Oscillator For Atomic Frequency Standards. United States Patent 5,489,891.

[4] Y. Koyama, et al., " An Ultra-miniature Rubidium Frequency Standard with two-cell scheme,"Proc. 49th Annual Symposium on Frequency Control, pp.33-38, 1995.

[5] R. H. Dicke, the effect pf collisions upon the doppler width of spectral lines physical review volume 89, number 2 january 15, 1953.

Compact Frequency Standard with Cold Atoms

Stella Torres Müller, Renato Ferracini Alves and Vanderlei Salvador Bagnato
Instituto de Física de São Carlos
Universidade de São Paulo
São Carlos, Brazil
{stella,vander}@ifsc.usp.br

Daniel Varela Magalhães
Escola de Engenharia de São Carlos
Universidade de São Paulo
São Carlos, Brazil
daniel@sc.usp.br

***Abstract*— Our laboratory has been developing an atomic reference system in which an ensemble of cold atoms of cesium is produced already in the center of a microwave cavity. The timeline of the project is divided in three parts: proof of principle with an antenna, design of a microwave cavity with viewports to produce the magneto-optical trap (MOT) and a movable system with all the parts embedded in a single package. In this work we present the first two parts of the global project with some relevant aspects for the construction of this compact frequency standard.**

I. INTRODUCTION

Compact atomic standards are being developed in many laboratories using different configurations [1][2]. These systems are all intended to have small size and high performance, showing a very promising field in time and frequency that addresses portability. Following this tendency our group is working in a simplified configuration that makes use of cold atoms, trapped in a magneto optical cell.

II. FIRST ESSAYS

The first experiment we did followed ideas from [3], where the ensemble of ^{133}Cs atoms was initially prepared in a magneto optical trap (MOT) and released to expand. During the free fall the clock transition is excited using a microwave field applied with a horn. The main changes in our system were the use of a pyrex cell and a quarter-wave antenna, as illustrated in Fig. 1.

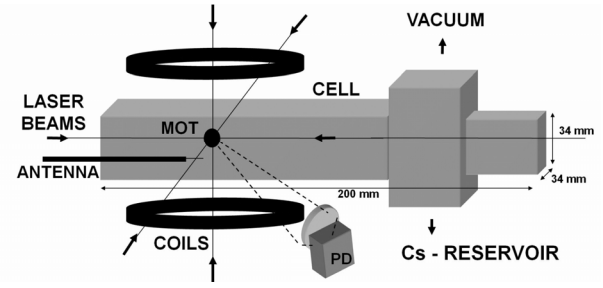


Figure 1. Vaccum cell and antenna used for the first experiment

This research is funded by FAPESP, CNPq and CAPES

Once mounted, the system showed a MOT with about 10^8 atoms that were interrogated with microwave fields in one and two pulses configurations. The detection was done using the same trap laser beams through fluorescence. The results observed for the hyperfine transition using both schemes of pulses are shown in Fig 2. They correspond to durations of 12ms in the single pulse and, in the second case, two 4ms pulses separated by an interval of 8ms.

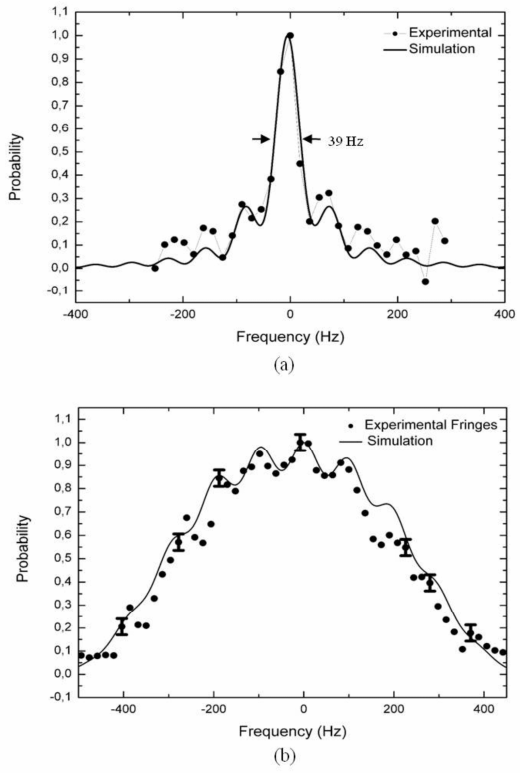


Figure 2. Resonance observed for the clock transition using (a) one and (b) two pulses

The observed results showed to be very promissing, with the linewidth for the single pulse configuration being 39Hz

and an Allan Standard Deviation of $9x10^{-11}\tau^{-1/2}$, when compared with a Cs commercial standard (Agilent 5071A)

The loss of contrast in the double pulse configuration was evaluated and the main contribution found was due to the spatial distribution of the field experimented by the atomic cloud in free expansion [4].

This configuration, despite the distribution of the field, has the very interesting aspect of compactness. Nevertheless, the use in a single pulse mode has shown a linewidth with reasonable contrast. The limitation in stability is atributed to SNR mainly.

III. Microwave Cavity with Cold Atoms

The next step was the evolution of the system for a better behaved microwave field. Since the aim of our experiment is not the use in microgravity, we decided to use the well known cilindrical resonator. Another point to be observed is the six beams configurations of the MOT, giving place to a cavity with at least six viewports. At the end, we finished with a cavity with ten viewports, all them with 10mm of diammeter, since this acts as cutoff waveguides for the used frequency. This number of viewports was necessary to guarantee place for the fluorescence detection, vacuum pump and Cs reservoir and microwave feeding with no adittion of many vacuum adapters. The microwave cavity was sculpted from a single block of 316L stainless steel, forming also the vacuum chamber with no weldings. A photograph of the chamber, with the built covers is shown in Fig. 3.

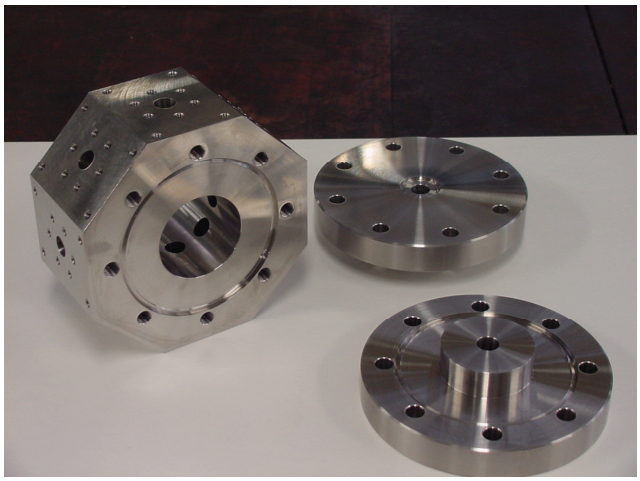


Figure 3. Covers and vacuum chamber body with the resonant cavity sculpted inside.

Because of the deformation of the resonant mode (TE011) due to all viewports, the dimensions of the cavity were first calculated with the use of FEM analysis. The cavity was then caracterized and the obtained quality factor is 1200. The temperature necessary to reach ressonance in the right frequency is 29°C. Considering the conductivity of the SS 316L, the quality factor gave us the confidence of a very good polishing and cleaning obtained in our mechanical workshop.

The MOT was them mounted with 9mm beams and the obtained cloud was measured with 10^9 atoms. The whole temporal sequence used for the experiment is generated using a data acquisition board (PCI-6025E National Instruments), that also performs all the measurements. The sequence include a phase of subdoppler cooling, allowing a cold sample with temperatures below 25μK.

After the subdoppler cooling phase, the cloud is released and submited to a sequence of two microwave pulses. The MOT beams are then turned on again and the fluorescence signal is collected with a PIN photodiode.

IV. Conclusions

The actual stage of our system is still far away the final shape we intent, since it is not optimized for minimum size of all the parts, like laser systems, microwave cavity, vacuum pump and optical bench. The target here was to obtain a vacuum chamber format that shows the feasibility of such system.

This chamber configuration permits also good optical access to the cloud of cold atoms, due to all viewports we were able to obtain, maintaining a reasonable quality factor for the cavity.

We considerate that the fluorescence detection is not the best method to use, since this results in a complete lose of cold atoms and increases the cycle time. One of the improvements we are implementing is the detection by absorption of an independent probe beam.

The actual system has now the objective to be a test bench for different configurations we plan to test, in order to minimize parts and subassemblies. Doing so we are passing to the third phase of our project, that is to design a system to be operated in an embedded configuration.

Acknowledgment

We thank A. Clairon, G. Santarelli and S. Bize for fruitful discussions and advices.

References

[1] F. X. Esnault, S. Perrin, S. Tremine, S. Guerandel, D. Holleville, N. Dimarcq, V. Hermann and J. Delporte, "Stability of the compact cols atom clock", Proceedings of the IEEE International Frequency Control Symposium, p1342-1345, 2007.

[2] V. Gerginov et al, "Component-level demonstration of a microfabricated atomic frequency reference", Proceedings of the Joint IEEE International Frequency Control Symposium and Exposition, p758-766, 2005.

[3] C. Monroe, H. Robinson and C. Wieman, "Observation of the cesium clock transition using laser-cooled atoms in a vapor cell," Optics Letters, vol. 16, no 1, pp. 50-52, January 1991.

[4] S. T. Müller, D. V. Magalhães, A. Bebeachibuli, T. A. Ortega, M. Ahmed, V. S. Bagnato, "Free expanding cloud of cold atoms as an atomic standard: Ramsey fringes contrast". Journal of the Optical Society of America. B, Optical physics, v. 25, p. 909, 2008.

Driving an extra small atomic resonator with low-power integrated RF frequency and laser locked loops

Jacques Haesler, Thomas Overstolz, Rony Jose James, Jörg Pierer, Matteo Contaldo, David Ruffieux and Steve Lecomte
Division Systems, Time & Frequency
Centre Suisse d'Électronique et de Microtechnique (CSEM) SA
CH-2002 Neuchâtel, Switzerland
jacques.haesler@csem.ch

***Abstract*— The first prototype of an eXtra Small Atomic Resonator (XSAR) was presented by CSEM in 2009. This paper describes the progress realized in 2010 towards the realization of a low-power miniature atomic clock. We present a second prototype as well as the first RF frequency and laser locked loops fully integrated on a chip.**

I. INTRODUCTION

Very accurate local clocks play a fundamental role in modern communication and navigation applications. High precision references enable fast communication data rates, while in navigation they allow longer holdover operation times in absence of synchronization signal, for example from the Global Positioning System (GPS). Even if this high accuracy can be achieved with atomic clocks, only the recent developments in photonics and MEMS processes allowed reaching the low power consumptions and small sizes needed for hand-held devices [1], paving the way to the realization of Miniature Atomic Clocks (MAC). However, reaching the target of overall device volumes < 1 cm^3 and power consumptions < 30 mW [2] requires further miniaturization and improved design of all the system aspects, including the RF control electronics, which has to be low-power without affecting the clock performances. State-of-the-art implementations as in [3] and [4] reached impressive low-power consumption performances, but using discrete electronic components which limit the system miniaturization. This work presents fully integrated RF frequency (RF IC v1.0) and laser (RF IC v2.0) locked loops developed at CSEM and designed for MAC applications, tested in combination with a breadboard level buffered ^{87}Rb cell system [5] based on coherent population trapping (CPT) interrogation.

The CPT interrogation [2] consists in modulating the bias current of a wavelength tuned VCSEL at an RF frequency (ν_{RF}) that is half the ground state hyperfine splitting of ^{87}Rb (ν_{HF} / 2 = 3.417 GHz). Fig. 1 illustrates the principle. Let's consider that a ramp on the laser diode (LD) bias current sweeps the pumping photon energy (E = hν) producing two peaks of absorption corresponding to transitions between each of the two ground states toward the excited states. When amplitude modulation of the LD is added, the absorption spectrum is composed of each individual laser wavelength (carrier plus sidebands). For $\nu_{RF} = \nu_{HF} / 2$, a transmission peak hundreds of Hertz wide (CPT signal) is observed and this resonance provides the high-Q frequency reference (> 10^6) of the MAC. To maintain the system locked around that peak and provide a highly stable clock, the frequencies of both the RF modulating signal and VCSEL must be controlled within feedback loops. In particular, locking of the RF frequency onto the atoms (RF IC v1.0) is obtained introducing frequency modulation of the RF carrier, which allows deriving a signed error signal used to form the RF frequency locked loop.

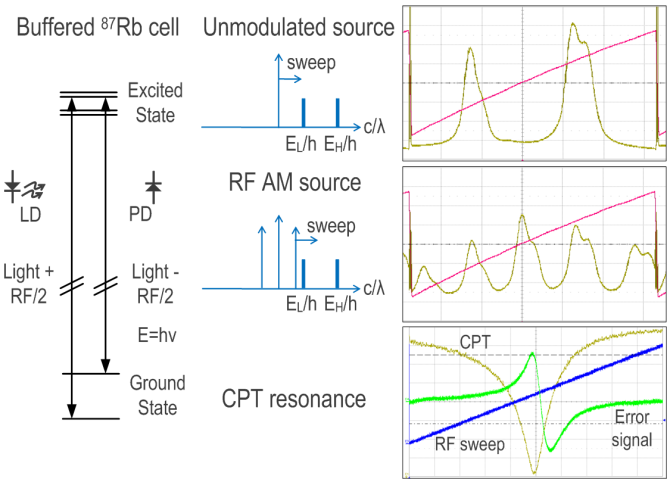


Figure 1. Physical principle of ^{87}Rb CPT interrogation.

II. PROTOTYPE EXTRA SMALL ATOMIC RESONATOR

The first prototype (v1.0) of an eXtra Small Atomic Resonator (XSAR) developed at CSEM had been demonstrated in [5]. It was built around a miniature 5 mm diameter glass cell filled with ^{87}Rb and a mixture of Ar and N_2 buffer gases. The prototype (Fig. 2 left) consisted in an assembly of miniaturized building blocks (the laser and optics

978-1-61284-110-6/11/$26.00 ©2011 IEEE

unit was only 4 mm high – heater and thermistor included) and laboratory-sized sub-systems (the glass atomic vapor cell unit was up to 20 mm long – heaters, thermistors and solenoid included). The performance of the complete clock driven by laboratory electronics and RF synthesizer was in agreement with the expectation: the short term frequency stability reached 5×10^{-11} @ 1 s.

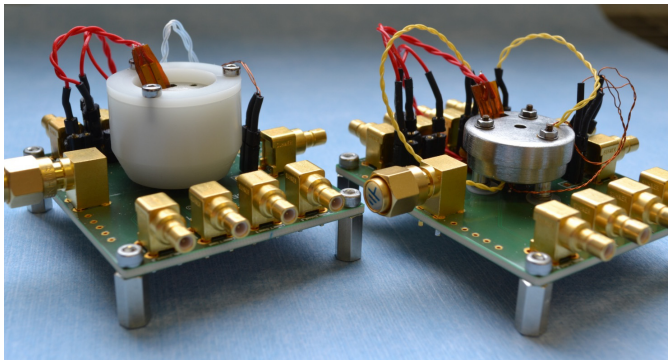


Figure 2. XSAR prototypes v1.0 (left) and 1.1 (right).

A step towards the small size and low-power objective has been tackled by mastering the fabrication of miniature atomic vapor cells using MEMS fabrication techniques (Fig. 3). Filling the cells with natural Rubidium has been realized with different techniques (alkali metal dispenser and alkali metal salt) and included controlled partial pressures of buffer gas mixture. Adding new functionalities to the cell and reducing the cell's external size are ongoing in order to reduce the overall physics package size.

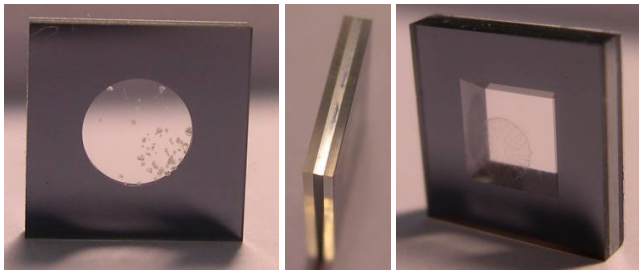


Figure 3. 10 x 10 mm MEMS cells with different filling approaches.

The second prototype (v1.1) of an eXtra Small Atomic Resonator (Fig. 2 right) could be assembled using the miniature MEMS cells. Driving this prototype with laboratory electronics allowed to reach frequency stabilities of the order of 8×10^{-11} @ 1 s. Driving the prototype with the integrated electronics (see below) is still not possible because MEMS cells could not yet be filled with the adequate isotope of Rubidium. The production of MEMS cells, filled with Rb-87, is ongoing. The performance results of the second prototype will trigger the design of the third generation prototype. It should integrate the low-power microwave synthesizer with integrated laser locked loop (see below), an advanced packaging for low power consumption, and a MEMS cell with reduced size and additional functionalities.

III. RF FREQUENCY LOCKED LOOP

Fig. 4 shows the block diagram of the integrated RF frequency locked loop (RF IC v1.0). A transimpedance amplifier (TIA) converts the photocurrent of the PD into a voltage. The signal is then chopped coherently with the FM modulation of the RF carrier before being integrated and low-pass filtered. The resulting voltage drives the frequency tuning port of a 40 MHz voltage controlled crystal oscillator (VCXO). The remainder of the system is equivalent to a direct modulation transmitter featuring a high resolution fractional-N PLL, whose loop bandwidth is chosen to minimize the RF close-in noise that will interact with the atoms. For increased flexibility, the low pass filter was implemented with external components. The standard 40 MHz crystal was chosen for small size (available down to 1.6 x 1.2 mm^2), possibility to derive the traditional 10 MHz reference after integer division and low in-band phase noise plateau thanks to a reduced PLL ratio. An external bias-T is used to couple the RF power and the LD biasing current before driving the VCSEL. The typical RF drive level is from –2 to +2 dBm. Frequency modulation of the RF carrier is performed at a programmable rate and depth with a 2-FSK unfiltered square wave signal. Both are typically in the kHz range with a lock-in amplifier loop bandwidth of 1 Hz. A resolution of the RF carrier frequency of 10^{-12} (ppt) is obtained with a 40-bit 2nd order ΔΣ modulator. The N divider, whose fractional value should be close to 85.425, is implemented with a cascade of dynamic dividers [6] to achieve low power dissipation.

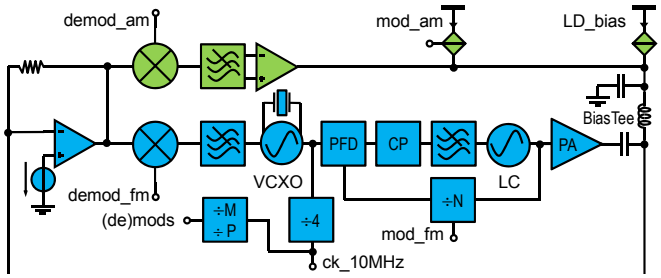


Figure 4. Architecture of the atomic clock RF frequency locked loop IC v1.0 (blue) and laser frequency locked loop IC v2.0 (green).

The photocurrent generated by the PD contains a strong DC component onto which a weak CPT signal, superimposed at the FM rate, has to be amplified. A low noise TIA (Fig. 5) with a band-pass filtering characteristic is implemented with only two transistors, an external capacitor and a current source. The two transistors form a current conveyor that reverse-bias the PD at a constant voltage independent of the photocurrent. At a higher frequency however, the capacitor freezes the gate of the top most transistor that presents at its source a large transimpedance. The latter is maximized with a large overdrive voltage and using a native transistor, resulting in high gain and much reduced noise compared to the PD shot noise. After amplification of the AC square wave signal by the transconductance and current mirror ratios, coherent chopping is performed so that the difference of its two components is integrated on a large external capacitor until the error signal vanishes. The resulting saw tooth voltage is further low-pass filtered with an external RC network before it drives the VCXO differential accumulation varactor. The differential VCXO based on [7], was designed for maximum tuning range (~ 100 ppm). A 7 bits switched capacitor bank with a large on/off ratio is used for coarse frequency tuning down to ppm

level so that ppt stability could be reached on the analog varactor.

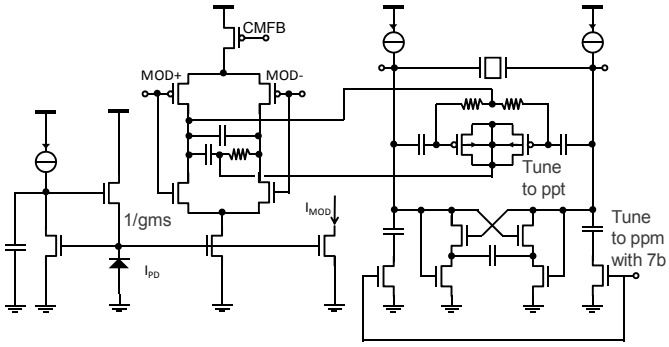


Figure 5. TIA, demodulator, loop filter and VCXO transitor level schematic.

The RF locked loop has been fabricated in a standard digital 0.18 µm CMOS technology. The chip has overall dimensions of 2.1 x 2.3 mm^2 comprised of pads and several test structures, which reduce the active area to about one half. Fig. 6 shows RF IC v1.0 integrated on a dedicated PCB as well as a micrograph with illustrated building blocks.

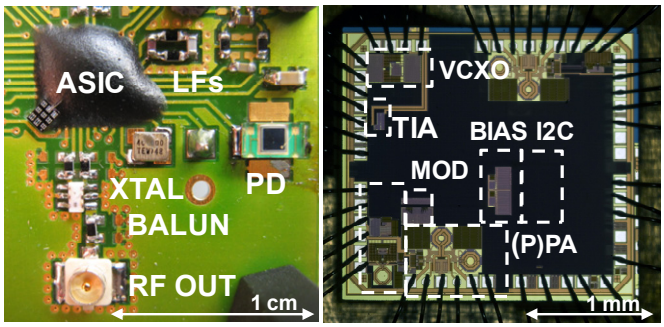


Figure 6. RF IC v1.0 on dedicated PCB (left) and chip micrograph (right).

Fig. 7 shows the measured phase noise of the RF signal (black), corresponding to −85.6 dBc/Hz at 1 kHz offsets from the 3.417 GHz carrier. At the PLL cutoff frequency of 250 kHz, the phase noise is −98 dBc/Hz. The phase noise of the RF signal is compared with what is achievable with typical laboratory equipment. The dark gray curve depicts the noise measured using a bulky oven controlled crystal oscillator (OCXO) in combination with an Agilent N5181A RF analog signal generator. The light gray curve shows the phase noise resulting from the use of the N5181A synthesis together with the integrated 40 MHz VCXO output after division by 4.

Fig. 8 depicts the MAC Allan deviation measured on the 10 MHz clock when locked through the IC RF loop to the physics package (□). The clock exhibits a 1-second intercept point of $\sigma_y(\tau = 1) = 4 \times 10^{-10}$, improving as $\sigma_y(\tau) \propto \tau^{-1/2}$ up to some tens of seconds. Work is ongoing to explain the abnormal drift observed for integration time $\tau > 20$ s also measured when using the IC VCXO in combination with the N5181A synthesizer (∇). The free VCXO frequency stability is also plotted (Δ). The last curve (○) represents the Allan deviation measured when locking the laboratory equipment to the ^{87}Rb cell.

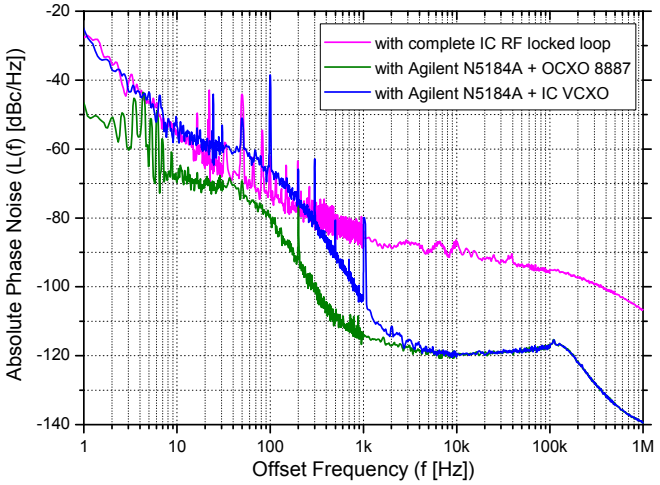


Figure 7. RF frequency locked loop phase noise: RF IC and standard laboratory equipments.

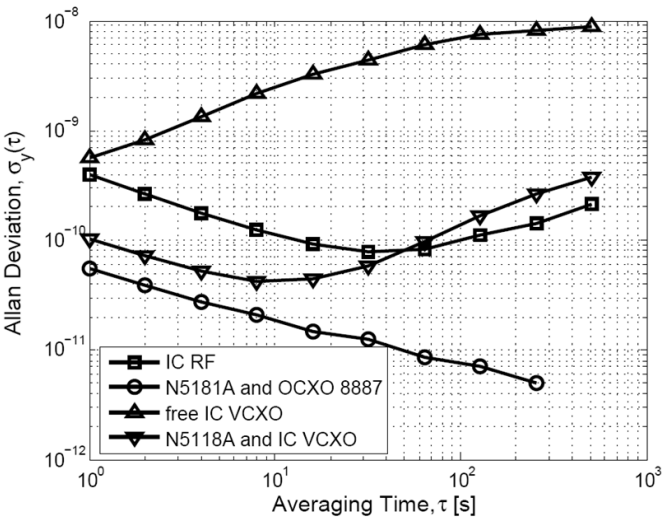


Figure 8. Measured Allan deviation on the implemented 10 MHz clock compared to results obtained with laboratory equipment.

The power consumption of the presented circuit is summarized in Fig. 9, together with a comparison with state-of-the-art RF electronics for MAC. The TIA and VCXO contribute to the overall consumption with only 0.31 mW, while the PLL comprised of the LC VCO consumes 7.89 mW. The overall consumption of 26.35 mW at an output power of 0 dBm suffers from the low efficiency of the power stage, which has been initially designed to reach output power up to 10 dBm. Redesign towards 0 dBm would necessitate the elimination of the power stage, resulting in an overall estimated power consumption of about only 12 mW.

IC consumption breakdown		
Block	Pow. Cons. [mW]	Supply [V]
TIA	0.17	1.8
VCXO	0.14	1.8
VCO	5.71	1.5
DIVIDERS	1.83	1.8
PFD+CP	0.3	1.8
MOD/DEMOD	0.05	1.8
PREPA	3.49	1.8
PA	14.66	1.5
TOT	26.35	-

Comparison with existing work			
	RF pow. cons. [mW]	$\sigma_y(\tau=1)$	IC integration
This work	26	4 E-10	yes
Lutwak09	51	5.1 E-11	no
DeNatale08	15	1 E-10	no

Figure 9. Chip consumption breakdown and comparison with existing work.

IV. Laser Frequency Locked Loop

Fig. 10 shows the schematic of the integrated laser frequency locked loop added in RF IC v2.0. The laser diode biasing is controlled via an external voltage obtained e.g. from a DAC. The voltage is copied internally and converted into a current across a resistor that is further mirrored towards the laser. A low frequency (10-100 kHz) square-wave modulated tiny current is superimposed to the main biasing current to alter slightly the optical frequency of the laser diode. The interaction with the atoms results in an amplitude modulated current after detection by the photodiode. The resulting signal at the TIA output is fed to a coherent demodulator and integrated on a loop filter whose implementations are similar to that of the frequency locked loop. The resulting tracking voltage is then converted to a current that is fed to the laser for fine optical frequency adjustment, hence closing the laser locked loop. Low frequency laser and RF frequency locked loops output signals are combined on an external bias tee before driving the laser diode.

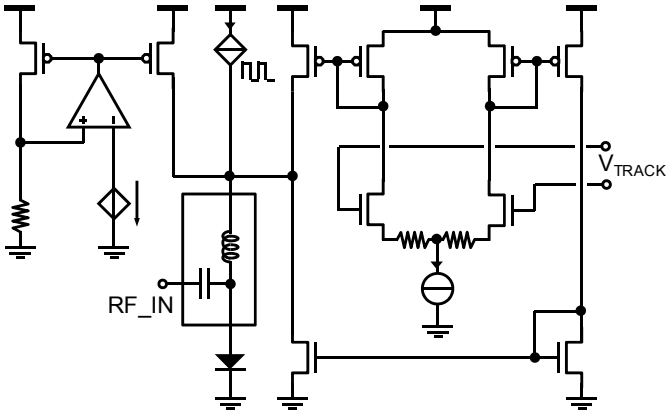


Figure 10. Schematic of the laser frequency locked loop.

The added functionality could be validated experimentally. Driving the laboratory CPT atomic clock described in [5] with RF IC v2.0 integrating the RF frequency locked loop as well as the laser biasing and frequency locked loop (Fig. 11) allowed to reach promising frequency stabilities of 1.5×10^{-10} @ 1s (Fig. 12). This system is to our knowledge the first RF frequency and laser locked loops for Miniature Atomic Clock (MAC) completely integrated on-chip.

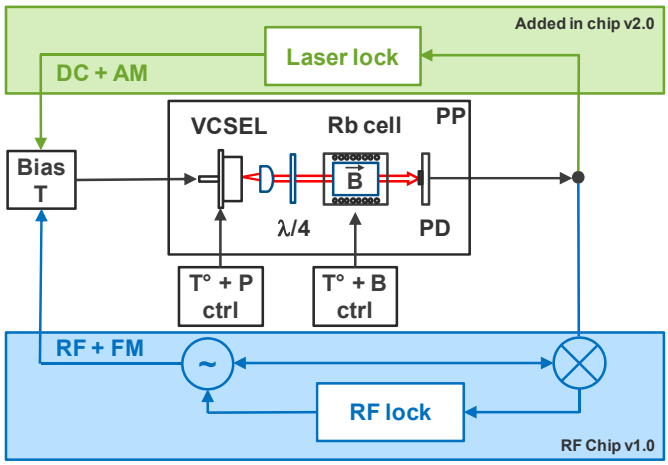


Figure 11. Graphical representation of RF IC v1.0 and v2.0 driving a laboratory CPT atomic clock.

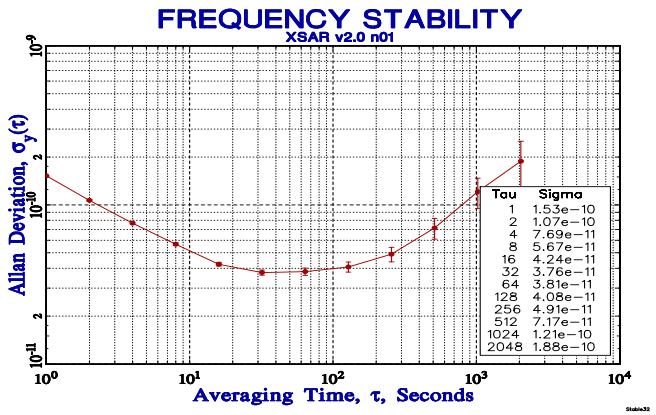


Figure 12. Frequency stability measured on the breadboard CPT atomic clock with the laser fully driven by RF IC v2.0.

Conclusion

The results of the second CSEM XSAR prototype have been presented. The prototype, integrating a MEMS-type atomic vapor cell, showed a frequency stability of 8×10^{-11} $\tau^{-1/2}$. The first version of a low-power integrated RF electronics has been characterized and integrated in a breadboard CPT atomic clock. With an overall chip consumption of 26.35 mW at an output power of 0 dBm, the clock showed a measured frequency stability of 4×10^{-10} $\tau^{-1/2}$. The second version of the RF chip, integrating the RF frequency locked loop as well as the laser biasing and the frequency locked loop, could also be successfully integrated in the clock. The measured preliminary frequency stability reaches a promising 1.5×10^{-10} $\tau^{-1/2}$. The next steps are to fully characterize the second version of the RF electronic unit, to miniaturize the MEMS-type atomic vapor cells adding new functionalities and to update the physics package towards

more integration and better thermal management. The project is ongoing and we expect to report new results soon.

ACKNOWLEDGMENT

This work was supported by CSEM's R&D activity financed by the SFOTE, the Canton of Neuchâtel, and the Cantons of Central Switzerland. Thanks to all CSEM collaborators participating in this project.

REFERENCES

[1] J. Kitching, S. Knappe, L. Liew, P. Schwindt, K. Shah, J. Moreland, L. Hollberg, "Microfabricated atomic clocks", IEEE International Conference on Micro Electro Mechanical Systems IEEE, pp. 1-7, February 2005.

[2] R. Lutwak, D. Emmons, W. Riley, R. M Garvey, "The chip-scale atomic clock – coherent population trapping vs. conventional interrogation", 34th Annual Precise Time and Time Interval (PPTI) Meeting, pp. 539-550, December 2002.

[3] R. Lutwak, "The chip-scale atomic clock – recent developments", IEEE International Frequency Control Symposium, pp. 573-577, April 2009.

[4] J.F. DeNatale, R.L. Borwick, C. Tsai, P.A Stupar, Y. Lin, R.A. Newgard, R.W. Berquist, M. Zhu, "Compact, low-power chip-scale atomic clock", IEEE Position, Location and Navigation Symposium, pp. 67-70, May, 2008.

[5] J. Haesler, J. Bennès, S. Lecomte, "Advances in the development of an eXtra Small Atomic Resonator (XSAR)", European Time and Frequency Forum (EFTF), Nordwijk (NL), April 2010.

[6] D. Ruffieux, J. Chabloz, M. Contaldo, C. Muller, F.-X. Pengg, P. Tortori, A. Vouilloz, P. Volet, C. Enz, "A narrowband multi-channel 2.4 GHz mems-based transceiver," IEEE Journal of Solid-State Circuits, vol 44, no. 1, pp. 228-239, January 2009.

[7] D. Ruffieux, "A high-stability, ultra-low-power quartz differential oscillator circuit for demanding radio applications," in Proceedings of the European Solid State Circuit Conference (ESSCIRC), pp. 85-88, September 2002.

Cs collisional frequency shift measurements in microcells filled with a Ne-Ar buffer gas mixture.

Rodolphe Boudot
Piotr Dziuban
Vincent Giordano
Christophe Gorecki
FEMTO-ST Institute
Besançon, France
Email: rodolphe.boudot@femto-st.fr

Danijela Miletic
Christoph Affolderbach
Gaetano Mileti
Laboratoire Temps-Frequence, Universite de Neuchatel
Neuchatel, Switzerland

Pawel Knapkiewicz
Jan Dziuban
Politechnika Wroclawska
Wroclaw, Poland.

***Abstract*— We report on the measurement of the dependence on temperature of the Cs clock frequency in microfabricated vapor cells filled with a mixture of Ne and Ar through the detection of Coherent Population Trapping (CPT) resonances. The temperature-dependence reduction of the Cs clock frequency is demonstrated in various cells. The inversion temperature at which the Cs clock frequency sensitivity is greatly reduced is found to be only dependent on the partial pressure of buffer gases and is measured to be lower than 80°C as expected with simple theoretical calculations.**

I. Introduction and Basics

The heart of vapor cell standards is an alkali vapor cell. A buffer gas or buffer gas mixture is generally added in the cell to slow the diffusion of the atoms and increase the time for the atoms to collide against the cell inner walls. Narrow linewidths can then be obtained through the Dicke effect [1]. At the same time, collisions between alkali atoms and the buffer gas perturbate hyperfine energy levels of alkali atoms, producing a clock frequency shift. In general, a buffer gas mixture is used. Each of them causes a clock frequency shift in opposite signs. A cancellation of the atomic clock frequency-temperature dependence can then be obtained by adjusting properly the buffer gas mixture composition. This technique is well-known for the development of high-performance compact Rb frequency standards [2]. We study here the Cs collisional frequency shift in vapor microcells filled with a mixture of Ne and Ar.

The frequency ν of the Cs hyperfine transition $|F=3,m_F=0\rangle \rightarrow |F=4,m_F=0\rangle$ can be described by:

$$\nu = \nu_0 + \alpha B^2 + \eta I(\omega) + \Delta\nu_{bg} \quad (1)$$

where ν_0 is the unperturbed Cs atom frequency (9.192 631 770 GHz), α is the second-order magnetic field coefficient (42.7 kHz/(mT)^2), B is the static magnetic field, η is the light shift coefficient and $I(\omega)$ is the intensity of the pumping light. $\Delta\nu_{bg}$ is often commonly called the buffer gas pressure shift and is the major concern of this study.
In a single buffer gas configuration, the pressure shift of the clock frequency $\Delta\nu_{bg} = \nu - \nu_0$ is generally described in a wide temperature range by [3]:

$$\Delta\nu_{bg} = P[(\beta + \delta(T-T_0) + \gamma(T-T_0)^2] \quad (2)$$

where P is the total buffer gas pressure at the reference temperature $T_0 = 273K$, T is the cell temperature, β is the pressure coefficient, δ is the linear temperature coefficient and γ is the quadratic temperature coefficient.
In the case of a buffer gas mixture, we define the ratio a between the pressure of buffer gas 2 and buffer gas 1 such as:

$$a = P_2/P_1 \quad (3)$$

Thus, the expression of the inversion temperature T_i for a given pressure ratio a is obtained by:

$$T_i = -\frac{\delta_1 + a\gamma_2}{2(\gamma_1 + a\gamma_2)} \quad (4)$$

Also, the pressure ratio a required to obtain an inversion temperature at $T = T_i$ is given by:

$$a = -\frac{\delta_1 + 2\gamma_1(T_i - T_0)}{\delta_2 + 2\gamma_2(T_i - T_0)} \quad (5)$$

Fig. 1 shows the estimated frequency shift of the Cs frequency versus the cell temperature for two different Ne-Ar pressure ratio and a total pressure of 6.6 kPa. Coefficients published in [4] are used.

For $a = 0$ (pure Ne cell), the temperature-dependence of the Cs frequency is expected to be cancelled at about 80°C as recently observed experimentally in chip-scale cells [5] and cm-scale glass blown cells [4]. Compared to pure Ne cells, the addition of Ar always shifts the inversion temperature to lower temperatures. The pressure ratio between Ar and Ne determines the inversion temperature.

II. Experimental set-up

Figure 2 shows a compact laboratory-prototype of the Cs CPT clock. It is fully described in [6].

The optical source is a 1.5 MHz-linewidth distributed feedback (DFB) laser diode. An optical isolator is used to prevent optical feedback. Two phase-coherent optical sidebands needed to perform the dark state spectroscopy are generated

978-1-61284-110-6/11/$26.00 ©2011 IEEE

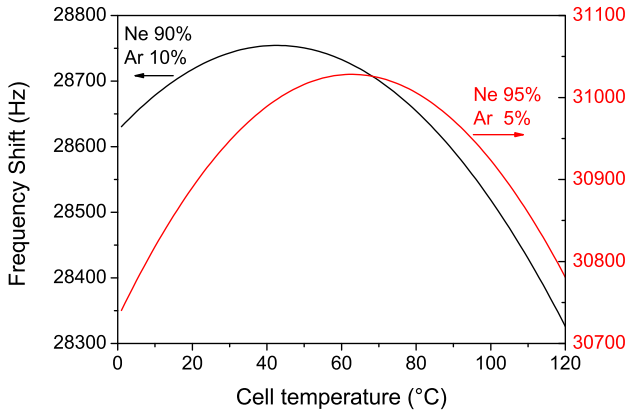


Fig. 1. *Theoretical estimations of the Cs frequency shift versus the cell temperature for cells filled with a Ne-Ar mixture. The total pressure is fixed to be 6.6 kPa. The ratio are Ne-Ar (90%-10%) and (95%-15%) respectively.*

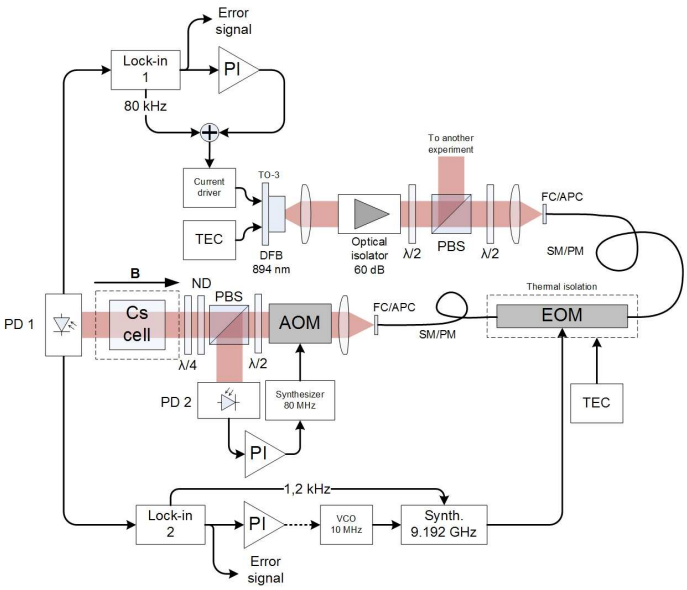


Fig. 2. Schematic of the Cs CPT clock used to test Ne-Ar microcells.

using a pigtailed electro-optic phase modulator. The EOM is driven by a low phase noise 9.192 GHz frequency synthesizer. The laser power is stabilized at the output of the EOM with an acousto-optic modulator (AOM)-based power servo loop. The laser beam is attenuated with a neutral-density filter and circularly polarized before to be sent in a 1.4-mm-long and 2-mm diameter Cs-Ne-Ar microcell shown in Fig. 3.

The cell temperature is stabilized at the mK level using a commercial temperature controller. The cell is surrounded by a solenoid to apply a static magnetic field of 10 μT and raise the Zeeman degeneracy. The ensemble is isolated from spurious magnetic fields using two concentric cylindrical mu-metal magnetic shields. The optical power transmitted through the cell is detected by a low noise Si photodiode PD1. A first servo loop is realized to lock the laser frequency using the

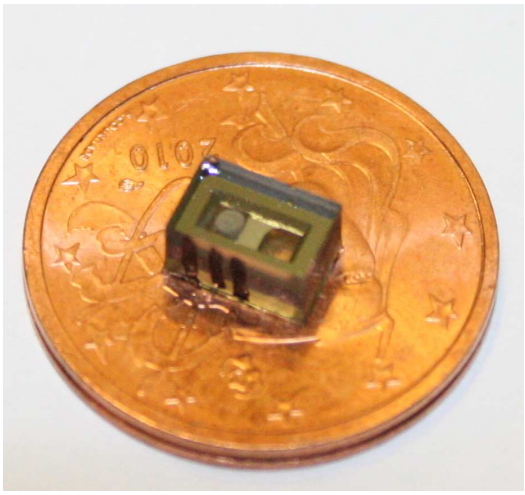


Fig. 3. *Photograph of a Cs-Ne-Ar microcell.*

lockin amplifier 1 while the second loop is used to lock the local oscillator frequency to the clock transition.

III. Experimental results

Several microcells have been realized. All the cells are expected to be filled with 85 % of Ne and 15 % of Ar. Fig. 4 shows a typical CPT resonance observed in a microcell filled with a buffer gas mixture of Ne and Ar.

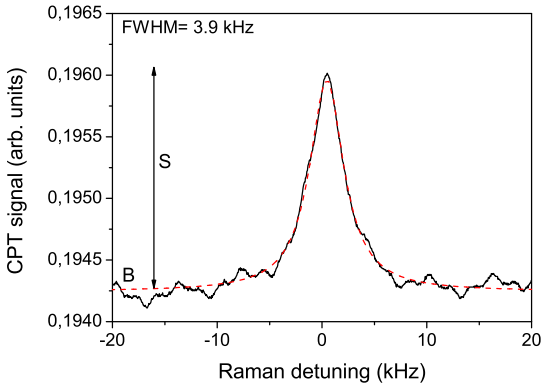


Fig. 4. *CPT resonance in a Cs-Ne/Ar microcell. The solid line is experimental data while the dashed line is a lorentzian fit. The laser power is 20 µW. The cell temperature is 75°C. The total buffer gas pressure is measured to be 8.5 kPa.*

The linewidth of the dark resonance is 3.9 kHz. Fig. 5 and 6 show the frequency shift of the Cs 0-0 hyperfine resonance as a function of the temperature for two different microcells. The buffer gas pressure shift measurement procedure is similar to the one described in [5].

The circles, stars, squares and triangles are the experimental data while solid lines are a second-order polynomial fit to the data. For both cells, the temperature-dependence of the Cs clock frequency is canceled for inversion temperatures lower than 80°C. Data fits are respectively $36170 + 15.39T - 0.14T^2$

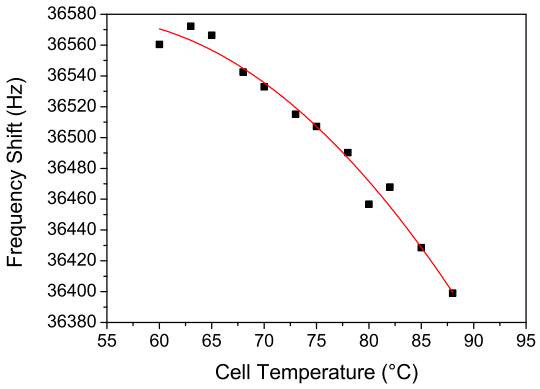


Fig. 5. *Frequency shift of the Cs clock versus the cell temperature in microcells filled with a Ne-Ar mixture. Cell 3.*

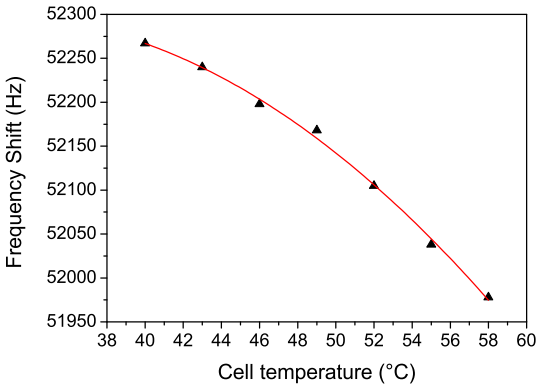


Fig. 6. *Frequency shift of the Cs clock versus the cell temperature in microcells filled with a Ne-Ar mixture. Cell 4.*

and $51823+29.10T-0.47T^2$ for the cells 3 and 4 respectively. The inversion temperature is measured to be 53.0 and 31.7°C for the cell 3 and 4 respectively. From these frequency measurements, we can estimate that the total pressure in the cell is 8.1 kPa and 12.5 kPa for cells 3 and 4 respectively. The Ne proportion is found to be 92.7 % and 87.7 % in cells 3 and 4 respectively. The uncertainty on the Ne proportion and the measured total pressure is typically $\pm$ 3.4 % and $\pm$ 0.9 kPa respectively for each cell.

The linewidth of the CPT resonance tends to be lower for increased buffer gas total pressure. The measured total pressure is found to be about 2-3 times smaller than the expected one. We note also that the measured pressure ratio is not in correct agreement with the expected one. Further tests are in progress to control accurately the actual filling total and partial pressures.

IV. Conclusions

The frequency-temperature dependence of the Cs clock frequency in microcells filled with a mixture of Ne and Ar was studied. A quadratic temperature dependence of the Cs frequency can be obtained with such microcells for inversion temperatures lower than 80°C. The inversion temperature only depends on the Ne-Ar pressure ratio and is shifted to higher temperatures for cells with increasing partial pressure of Ne. The total frequency shift depends on both the total and partial pressures. These results can be useful for the development of vapor cell frequency standards.

Acknowledgments

This work is supported by a collaborative project MAC-TFC of the European Commission (FP 7, ICT program, grant n° 224132).

References

[1] R. H. Dicke, "The Effect of Collisions upon the Doppler Width of Spectral Lines," Phys. Rev. **89**, 472-473 (1953).

[2] J. Vanier, R. Kunski, N. Cyr, J. Y. Savard and M. Tetu, "On hyperfine frequency shifts caused by buffer gases: Application to the optically pumped passive rubidium frequency standard," Journ. Appl. Phys. **53**, (8), 5387 (1982).

[3] J. Vanier and C. Audoin, "The Quantum Physics of Atomic Frequency Standards," Adam Hilger, Bristol (1989).

[4] O. Kozlova, R. Boudot, S. Guérandel and E. De Clercq, "Measurements of Cs buffer gas collisional frequency shift using pulsed coherent population trapping interrogation," *Proc. of the 2010 Conf. Precision Electromag. Meas.*, Daejon, Korea (2010).

[5] D. Miletic, P. Dziuban, R. Boudot, M. Hasegawa, R. K. Chutani, G. Mileti, V. Giordano and C. Gorecki "Quadratic dependence on temperature of the Cs 0-0 hyperfine resonance in a single Ne buffer gas microfabricated vapor cell," Elec. Lett. **46**, (15), 1069 (2010).

[6] R. Boudot, D. Miletic, P. Dziuban, C. Affolderbach, P. Knapkiewicz, J. Dziuban, G. Mileti, V. Giordano and C. Gorecki "First-order cancellation of the Cs clock frequency temperature dependence in Ne-Ar buffer gas mixture" Opt. Exp. **19**, (4), pp 3106-3114 (2011).

[7] R. Boudot, P. Dziuban, M. Hasegawa, R. K. Chutani, S. Galliou, V. Giordano and C. Gorecki, "Coherent Population Trapping resonances in Cs-Ne vapor microcells for miniature clocks applications", *Journ. Appl. Phys.*, **109,** 014912 (2011).

Multiple stack anodically bonded 4 mm thick Rb vapor cell

Yves Pétremand, Rahel Straessle and Nico de Rooij
Sensors, Actuators and Microsystems Laboratory (SAMLAB)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
2000 Neuchâtel, Switzerland
yves.petremand@epfl.ch

Matthieu Pellaton, Christoph Affolderbach and Gaetano Mileti
Laboratoire Temps-Fréquence (LTF)
Université de Neuchâtel
2000 Neuchâtel, Switzerland

***Abstract*—This paper presents a new fabrication method to manufacture alkali reference cells having dimensions larger than standard micromachined cells and smaller than glass-blown ones. This technology fills a gap in cell sizes and already available technologies.**

I. INTRODUCTION

Many efforts are made to miniaturize the overall size of atomic clocks. This miniaturization has many advantages (i.e. reducing the power consumption, the cost and the mass of the physics package of the clock, …) but one drawback is that it degrades the Allan deviation proportionally to the cavity size reduction as shown by Kitching et al. [1]. These chip scale devices are usually manufactured using MEMS (Micro Electro Mechanical Systems) technologies and anodic bonding [2]. For higher precision clocks, glass blown cells are generally used but this technique is limited to cell volumes bigger than few tens of cubic millimeters [3, 4].

Here we present a technology that fills the gap in achievable cell size between the chip-scale MEMS cells and the macroscopic glass-blown cavities. For applications where a cavity size with a critical dimension of few millimeters is required, the glass-blown technique does not allow good control of cell geometry or flat window surfaces, and standard MEMS technology based on DRIE (Deep Reactive Ion Etching) or KOH etching of silicon is not applicable due to the required etching depth.

II. TECHNOLOGY COMPARISON

This chapter will show the advantages and disadvantages of the possible cell fabrication techniques, depending on the application or the requirements. Some arguments are summarized in Table I.

A. Glass blowing technique

The glass blowing technique, discovered about 50 B.C. [5], is still used today. It consists of creating an object by gathering molten glass on the end of a hollow blowing pipe and inflating it like a bubble. It can be blown into a hollow mold to form it with good precision or freely shaped with simple tools.

Modern variants of this technique are used as state-of-the-art technology for the fabrication of macroscopic alkali cells used in the commercially available products. It is possible to obtain cells exempt of impurities and their long lifetime were already proven. The inconvenience of this technology appears when reducing drastically the size of the cell down to a couple of millimeters. The blowing induces rounded edges and does not allow obtaining a flat window surface quality. If the surface is not perfectly flat, light losses are observed. The minimal achievable reasonable cell size is of around 6-10 mm total length or diameter.

B. Microfabrication technology

MEMS technology development started a couple of decades ago, at first mainly for the microelectronics industry processes. Inventions like DRIE allowed creating structures in bulk substrates, mainly on silicon. The advances in bonding techniques allowed creating more and more complex and ever smaller structures in three dimensions.

TABLE I. COMPARISON BETWEEN THE TWO PRESENTED TECHNOLOGIES

Technology	Advantages	Issues
Standard glass-blowing	Well-know and clean technology	Relatively poor window quality for small cavity size
	Long cell lifetime proven	Minimum achievable size of around 10 millimeters
Microfabrication technology	Perfectly flat window	Expensive technology
	Wafer-level fabrication	Alkali dispensing method difficult to implement
	Freedom concerning the shape of the cell	Long term aging not yet fully studied

This work was supported by the Swiss National Science Foundation (Sinergia grant CRSI20-122693/1).

One of the major advantages of this technology is the possibility to process several samples in parallel at wafer level and the freedom in the shapes that can be given to structures as it is defined by a photolithographic mask. Furthermore, for fabrication of alkali cells by bonding several of these wafers together to form a cavity, the obtained sidewalls have a very low roughness and it is possible to obtain perfectly polished window surfaces. On the other hand, this technology requires heavy and expensive equipment. Several alkali filling methods were already reported [6, 7] but the precise amount of alkali is difficult to control. Working microfabricated cells were presented during the last years and accelerated thermal aging tests were already performed [8] but yet the long term behavior is not fully studied.

III. VAPOR CELL DESIGN

The designed chip is represented in Fig. 1. Its dimensions are the followings:

- Internal diameter: 4 mm
- Internal height: 4.05 mm
- Outer footprint: 6x6 mm
- Outer height: 5.05 mm.

It is composed of a 3 mm thick drilled core made out of glass and bonded to two rims of silicon on each side, creating a 4.05 mm long cylindrical cavity. This cavity is closed with glass lids on both sides.

IV. FABRICATION PROCESS

A schematic view of the process is shown in Fig. 2. After cleaning the required substrates (Fig. 2 A), our approach is to start with a thick glass wafer in which through holes are drilled. This glass wafer is anodically bonded to a first silicon wafer patterned by DRIE (Fig. 2 C). The result obtained is bonded to a glass wafer to create the bottom windows of the cavities. This stack of three wafers is again bonded to a silicon wafer patterned by DRIE on the other side of the thick glass substrate (Fig. 2 E). These bonding steps are performed in a Süss bonding machine at 400°C and up to 1000 V.

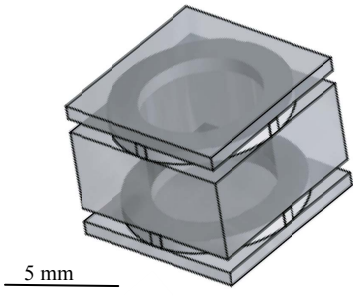


Figure 1. CAD image of the vapor cell.

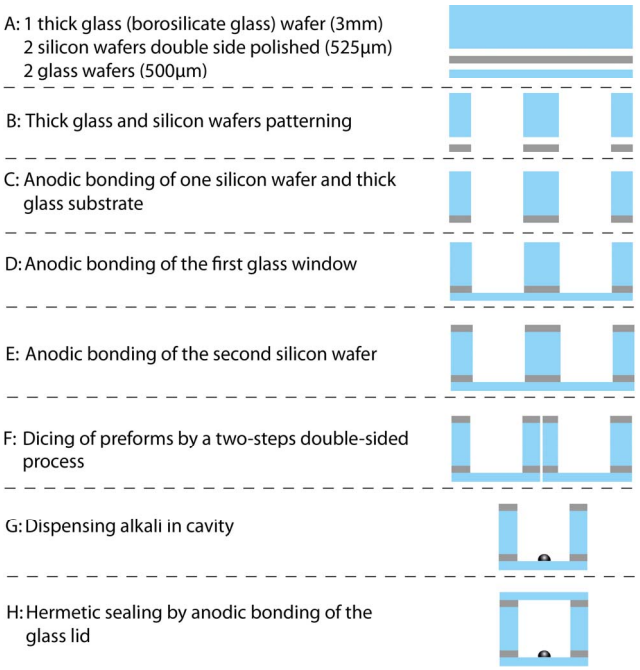


Figure 2. Fabrication process flow.

The obtained stack is then diced using a double-sided dicing process. The dicing is performed on both sides due to the large thickness of the structures (4550 µm). The alignment of the blade is performed on dicing marks patterned on both silicon wafers.

The diced pre-forms are then filled with a droplet of alkali and a glass lid is anodically bonded on top. The dispensing technology was previously presented by Pétremand et al [9]. The method is based on the condensation of alkali vapor produced by a commercially available dispenser from SAES Getters.

V. FABRICATION AND SPECTROSCOPY RESULTS

The cell fabricated according to the described process is depicted in Fig. 3. The left picture shows a top view of the cell. A small metallic spot of alkali is visible in the center of the cavity. The right picture shows an isometric view of the device. It can be seen that the outer dimensions are almost

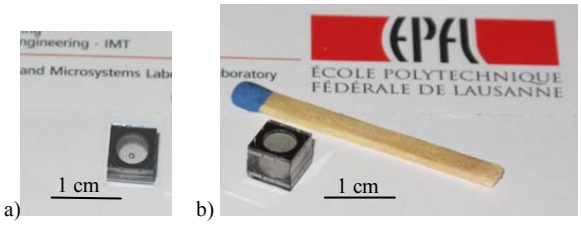


Figure 3. Pictures of the obtained cell. a) Top view. A small spot of alkali metal can be seen in the cavity. b) Isometric view.

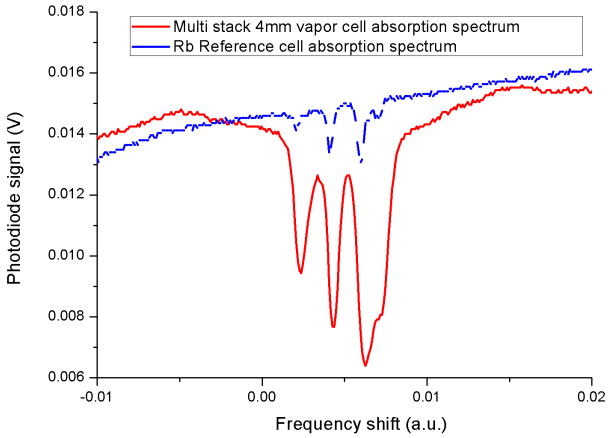


Figure 4. Absorption spectrum measured through the 4 mm thick Rubidium cell.

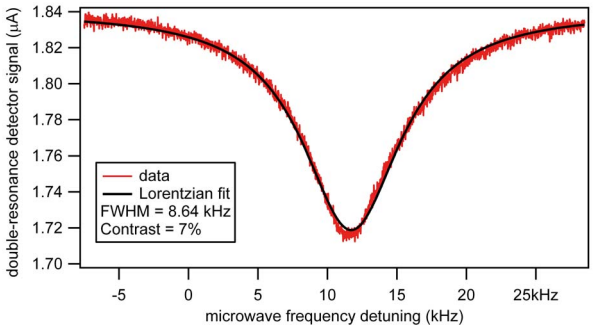


Figure 5. Double-resonance signal of the clock transition observed in the 4mm thick cell.

cubic. Preliminary test showed the expected Rb absorption spectrum (Fig. 4) for a cell heated to about 100°C (here, absorption lines are broadened due to the buffer-gas filled into the cell).

Double resonance measurements were performed with the alkali cell as shown in Fig. 5. The initial tests gave estimated clock short term stability of $7 \times 10^{-11}\ \tau^{-1/2}$. These results were obtained without optimization of the parameters and will be improved in a near future.

VI. Conclusion

A new type of middle-sized cells fabricated using microfabrication technology is shown. Spectroscopy measurements were performed and double resonance signal was measured. Initial measurements show a clock signal compatible with a clock stability of at least $7 \times 10^{-11}\ \tau^{-1/2}$.

The presented technology opens the perspective for massively parallel production of Rb vapor cells of few millimeters dimensions by using MEMS / microfabrication technologies. These cells are of interest for use in compact conventional Rb atomic clocks using small vapor cells.

Acknowledgments

The authors would like to thank the CSEM cleanroom staff, Idonus sarl and all the project partners for their contribution and help.

References

[1] J. Kitching, S. Knappe and L. Hollberg, “Miniature vapor-cell atomic-frequency references”, Applied Physics Letters, vol. 81, num. 3, p. 553-555, 2002.

[2] S. Knappe, “MEMS atomic clocks”, Comprehensive Microsystems, vol. 3, p. 571-612, 2008.

[3] J. Haesler, J. Bennès and S. Lecomte, “Advances in the development of an eXtra Small Atomic Reference (XSAR)”, Proceedings EFTF, 2010.

[4] C. Schori, G. Mileti, B. Leuenbeger, P. Rochat, "CPT Atomic Clock based on Rubidium 85", Proceedings EFTF, 2010, p.125.

[5] Frank Starr, “A resource for glass”, The Corning Museum of Glass Education Dept., 1998.

[6] L-A. Liew, et. al. “Microfabricated alkali atom vapor cells,” Applied Physics Letters, 84, p. 2694 (2004).

[7] F. Gong, Y.-Y. Jau, K. Jensen, and W. Happer, “Electrolytic fabrication of atomic clock cells”, Rev. Sci. Instrum. 77, 076101 (2006).

[8] M. Hasegawa et al., “Microfabrication and thermal behavior of miniature cesium-vapor cells for atomic clock operations”, MEMS 2011, Cancun, Mexico, January 23-27, 2011.

[9] Y. Pétremand et al., “Low temperature indium-based sealing of microfabricated alkali cells for chip scale atomic clocks”, Proceedings EFTF, 2010.

Investigations on improved Rb cell standards

Thejesh Bandi, Florian Gruet, Christoph Affolderbach and Gaetano Mileti
Laboratoire Temps – Fréquence (LTF)
Institut de Physique, Université de Neuchâtel
Neuchâtel, Switzerland
Email: gaetano.mileti@unine.ch

Claudio E. Calosso
Optics division
Istituto Nazionale di Ricerca Metrologica (INRIM)
Torino, Italy

***Abstract* — In this article we present our development of a high performance laser-pumped Rubidium gas-cell atomic frequency standard based on the continuous microwave-optical double resonance interrogation. A larger resonance cell is employed and no laser noise cancellation is required to reach a short-term frequency stability below $4\times10^{-13}/\sqrt{\tau}$ for $40 < \tau < 1000$ s.**

I. Introduction

Since last few decades the development of laser-pumped rubidium standards has become the subject of interest in view of improving the stability, mainly by replacing the lamp with a laser for optical pumping [1]. The initial experiments using lasers already proved to be improving the signal-to-noise and shot-noise limits attainable in contrast to that of lamps [2]. Since then, the constantly improving technology of compact laser sources and their availability has made the subject more reliable. For instance, the recently available Distributed-Feedback (DFB) laser sources are compact and robust. A compact, laser-pumped clock using an Extended Cavity Diode Laser (ECDL) had been already demonstrated with a short-term stability of $3\times10^{-12}\,\tau^{-1/2}$ [3]. And we note our recent work with improved performance of $<4\times10^{-13}/\sqrt{\tau}$ and reaching 1×10^{-14} at 10^4 s [4].

Our motivation is to achieve the stability of $<1\times10^{-12}/\sqrt{\tau}$ between 1 to 1000 s (short- and medium-term) and $<1\times10^{-14}$ (drift removed) up to one day. To achieve such a goal, the design requirements of our clock include; (i) the physics package (PP), that is able to give a DR linewidth of <1 kHz, and a discriminator slope of the clock signal > 1 nA/Hz. The detection noise due to the laser after passing through PP should be <5 pA/√Hz. The optimized buffer-gas mixture on the clock cell should give a temperature coefficient of $<1\times10^{-11}$ /°C, this parameter is essential in reaching medium and long-term stabilities of $< 10^{-14}$ level. Magnetic shields are required around the Rb cell to have $<10^{-13}$/G sensitivity. (ii) A stabilized laser source, with 780 nm wavelength and power output of few tens of microwatts in order to optically pump the atoms in D2 transition of ^{87}Rb to one of the hyperfine ground states. The linewidth of the laser should be of few MHz range to resolve all the sub-Doppler transitions, which is useful for laser locking and stabilization to an auxiliary, evacuated Rb cell. The Relative Intensity Noise (RIN) and FM noise should be $<10^{-11}$ Hz^{-1} and <10 kHz/√Hz, respectively. The power stability of the laser needs to be <0.1 %/day, at fixed environmental conditions. And (iii) microwave Local Oscillator (LO) synthesizer that can generate a stable 6.835 GHz and a power output of few dBm is required. A tuning range of 2 MHz with steps of <1 Hz resolution is essential. Most importantly, the phase noise in the clock frequency modulation, f_m range, when referred to a 10 MHz source should be ≤ -137 dBrad^2/Hz.

The flow of this paper highlights the important steps, tests and results in building the clock to achieve the goal envisaged by our motivation.

II. Experimental Setup

In Fig. 1 we show the block diagram of the laser-pumped Rb atomic frequency standard. We briefly describe below its three main components: the physics package, the laser head and the electronic control system.

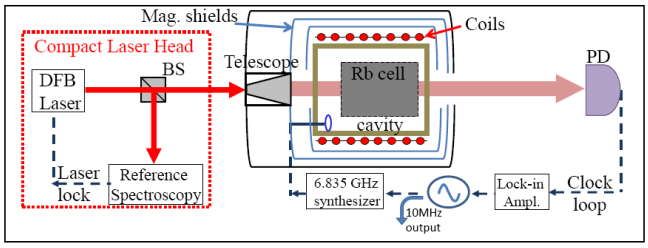


Fig.1. Schematic of the Double-Resonance clock setup.

A. Physics Package (PP)

As presented in [5], the atomic resonator is made of a 25mm-diameter glass-cell filled with rubidium and a mixture of Argon and Nitrogen as buffer gases. This clock cell is placed in a newly developed magnetron-type microwave cavity resonant around 6.8 GHz and surrounded by a C-field coil and two magnetic shields. The overall volume of the Physics Package (PP) is below one liter.

This work was supported by the Swiss National Science Foundation, the European Space Agency, and the Swiss State Secretariat for Education and Research – Space Affairs Division.

B. *Laser head*

Following our previous realization of a compact extended-cavity laser head, which includes a saturated-absorption reference cell [6], we have developed an improved version based on a DFB laser source [7] and its variant that comprises also an Acousto-Optical Modulator (AOM) that allows a fine tuning of the laser frequency towards a "magic" zero intensity light-shift point [5]. For the study presented here, we have used the version without AOM and operating around 780 nm. The overall volume of the laser head is below one liter. Fig. 2 shows the Allan deviation of the beat frequency of two identical compact non-AOM laser heads. The stability after 100 s is degraded by the external thermal fluctuations in the room, though during the clock operation we had better control over the thermal isolation of the laser head.

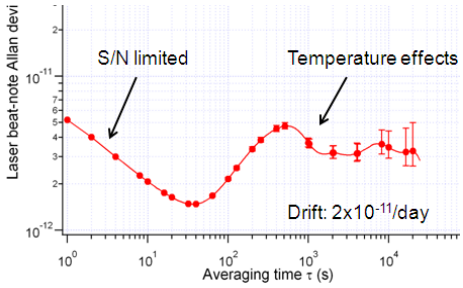


Fig. 2. Stability of two identical 780nm compact laser head units.

C. *Electronics control system*

For this work we used home-made laboratory electronics modules consisting digital temperature controllers for the PP, the Laser Head and a low-noise current source with the lock-in for the DFB laser source. The microwave source (or Local Oscillator, LO) and the computer-controlled lock-in were realized in INRIM and are described in [5, 8, 9, 14].

III. SPECTROSCOPIC STUDIES AND CLOCK STABILITY

A. *Double resonance signal and signal-to-noise limit*

Fig. 3 shows a typical double resonance signal, which corresponds to a shot-noise limited frequency stability of $1.4\times10^{-13}/\sqrt{\tau}$. Presently, the laser AM and FM noise and a residual contribution of the LO phase noise prevent us from reaching this performance and set a signal-to-noise limit of $2.9\times10^{-13}/\sqrt{\tau}$.

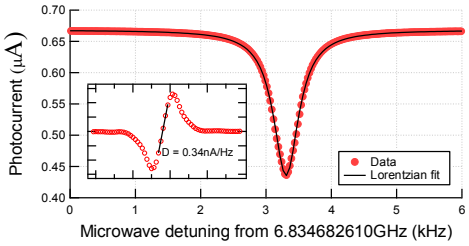


Fig. 3. Double-Resonance signal and the clock signal discriminator slope in the inset is also shown.

B. *Light-shift*

One of the main limits for the medium-term and long-term frequency stability of laser-pumped Rb clocks results from the AC Stark effect, through the laser intensity (α-coefficient) and frequency (β-coefficient) light-shift phenomena. Fig. 4 and 5 depict these two effects in our experiment when locking the laser source to the different possible sub-Doppler spectroscopy lines and at different optical pumping intensities. The influence of light-shift proves negligible in the short-term (below 1×10^{-14} for $1s < \tau < 100s$) while limiting the stability to around 2×10^{-14} for $\tau = 10'000$ s.

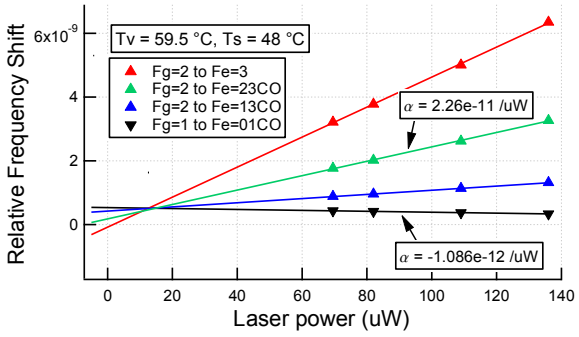


Fig. 4. Intensity light-shift (α) for different laser locking transitions of D2 line. Locking to the Fg = 1 to Fe = 01CO results in a minimum α-coefficient.

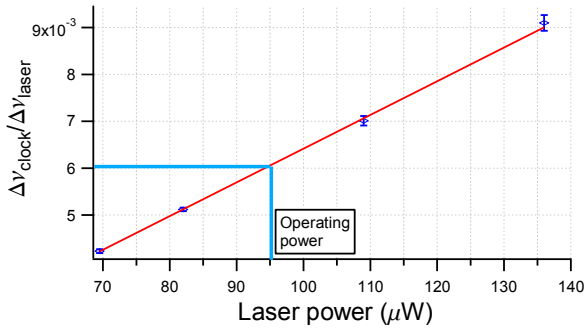


Fig. 5. Frequency light-shift coefficient (β) as a function of interrogating laser power.

C. *Temperature coefficient and other effects*

The relatively large total frequency shift in gas cell frequency standards is due to collisions with the buffer gas. Using an in-house cell filling system, we have filled the clock cell with a mixture of Argon and Nitrogen that prevents the Rubidium atoms from colliding with the cell's inner walls (buffer gas effect) and strongly suppresses their fluorescence (by quenching) that would otherwise reduce the optical pumping efficiency. The pressure of the mixture gives rise to a constant shift, which we have used in reducing the intensity light-shift for Fg = 1 to Fe = 01CO transition, whereas the ratio of this mixture allows fine tuning of the temperature coefficient inversion point, as shown on Fig. 6. When the cell is operated at a temperature close to 59°C, an overall temperature stability of the resonant cell within a few mK is sufficient to reduce the resulting fractional clock instabilities to below 10^{-14}.

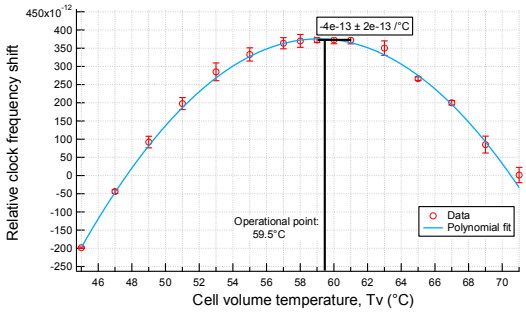


Fig. 6. Temperature coefficient of the clock cell. Ratio of the buffer gases determines the inversion point.

The other sources of clock instability, such as the unavoidable thermal gradients, the spin-exchange effect and the microwave power coefficient are presently under study.

D. Clock stability

Fig. 7 displays our present clock frequency stability results of $4\times10^{-13}/\sqrt{\tau}$ for 40 s < τ < 10000 s. The inset of the graph shows the raw frequency data measured. The signal-to-noise limit (red-line) $2.9\times10^{-13}/\sqrt{\tau}$, the shot-noise limit (green-line) $1.4\times10^{-13}/\sqrt{\tau}$ and the measured stability (black-line) $4\times10^{-13}/\sqrt{\tau}$ are shown. Our measurement system has a limit with $1/\tau$ dependence, which is shown in the blue-dotted-line. We suspect that the bump around 20 s is due to noise in the measurement system loop.

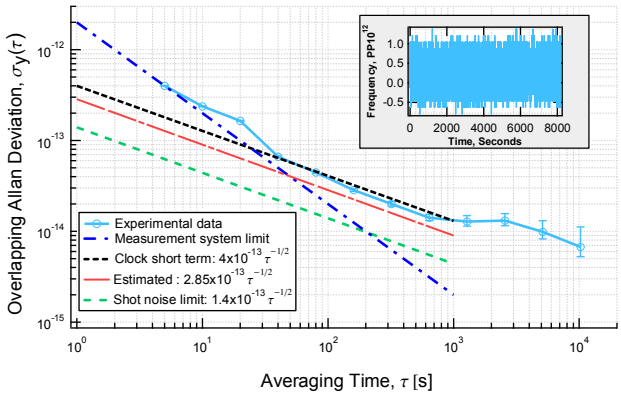


Fig. 7. Allan deviation of the clock and the inset shows the measured frequency data. Linear drift of 2.5×10^{-13}/day is removed [4].

IV. CONCLUSION

In this paper, we presented an improved Rb standard and its important parameters: the light-shift effect and temperature coefficient which are well-controlled to reach the stability around 10^{-14} over a period of 1day. The compact DFB laser-heads' beat-note stability was also presented. The clock Allan-deviation reaching the estimated goals of $<1\times10^{-12}/\sqrt{\tau}$ between 1 to 1000 s (short- and medium-term) and $<1\times10^{-14}$ (drift removed) up to one day was shown along with the estimated limits.

In comparison to the previous realization of the continuous wave (CW) operation using laser [3], we have shown an improved standard. Our clock has a smaller volume in comparison to vapor-cell clocks in pulsed operation [9], passive hydrogen maser (PHM) [10], HORACE [11], space optical clock [12] and Cs-beam clock [13], and gives the stability that is still sufficient for the next generation clocks for navigation systems, such as GALILEO and future telecommunication applications.

ACKNOWLEDGMENT

We thank M. Pellaton for fabrication of the cells and the other members of LTF, in particular P. Scherler and M. Durrenberger who have contributed in realizing this clock.

REFERENCES

[1] J. Vanier, C. Mandache, "The passive optically pumped Rb frequency standard: the laser approach", Appl. Phys. B: Lasers and Optics, vol. 87, pp. 565-593, 2007.

[2] G. Mileti, P. Thomann, "The impact of laser optical pumping on the physics and on the performances of the rubidium frequency standards", 5th Symposium on Frequency Standards and Metrology, editor James C. Bergquist, World Scientific, 1996, pp. 392-393

[3] C. Affolderbach, F. Droz and G. Mileti, "Experimental demonstration of a compact and high-performance laser-pumped rubidium gas cell atomic frequency standard", IEEE Trans. Instr. Meas., vol. 55, no. 2, pp. 429-435, 2006.

[4] T. Bandi, C. Affolderbach, C. Calosso, and G. Mileti, "High-performance laser-pumped rubidium frequency standard for satellite navigation", Electronics Letters, in press.

[5] T. Bandi, C. Affolderbach, G. Mileti, "Studies on an improved compact physics package for Rb cell standards", proc. on 42nd annual Precise Time and Time Interval (PTTI) systems and applications meeting, 2010, pp. 329-342.

[6] C. Affolderbach, G. Mileti, "A compact laser head with high-frequency stability for Rb atomic clocks and optical instrumentation", Review of Scientific Instruments, vol. 76, 073108, 2005.

[7] S. Micalizio, A. Godone, F. Levi, C. Calosso, T. Bandi, M. Pellaton, F. Gruet, C. Affolderbach, G. Mileti, "Pulsed optically pumped Rb clock with optical detection: first results", proc. of 24th European Time and Frequency Forum (EFTF), 2010, ESA-ESTEC, Noordwijk (NL).

[8] C. E. Calosso, F. Levi, E. K. Bertacco, A. Godone, and S. Micalizio, "Low-noise electronic design for the 87Rb coherent population trapping maser", IEEE Trans. Ultrason. Ferroelectr. and Freq. Control, vol. 52, no. 11, pp. 1923-1930, 2005.

[9] S. Micalizio, A. Godone, C. Calosso, F. Levi, F. Gruet, C. Affolderbach, "Pulsed optically pumped Rb clock with high frequency stability performances", proc. of Joint conference of the IEEE-IFCS/EFTF, 1-5 May 2011.

[10] M. Belloni et al., "Space passive hydrogen maser -performances, lifetime data and GIOVE-B related telemetries", Proc. of the 24th EFTF, 13-16 April 2010, ESA/ESTEC, Noordwijk (NL).

[11] F.X. Esnault et. al., "HORACE : A compact cold atom clock for Galileo," J. Adv. Space Res., vol. 47 (5), pp. 854-858, 2011.

[12] S. Schiller et al., "The space optical clock project", Proc. International Conference on Space Optics, 2010, Rhodes / Greece.

[13] S. Lecomte et al., "Performance demonstration of a compact, single optical frequency cesium beam clock for space applications", Proc. IEEE Frequency Control Symposium & 21st European Frequency and Time Forum, 2007, pp. 1127 – 1131.

[14] C. E. Calosso et al., "Electronics for the pulsed rubidium clock: design and characterization", IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 54 (9), pp. 1731-40, 2007.

A Fully Functional Ramsey-CPT Atomic Clock

Qun Zhou, Chao Guan, and Xiaosong Zhu
The Application Center of New Technology,
Electronic Engineering Institute,
Hefei, 230037, People's Republic of China,
E-mail: guanchao2008@163.com

Yeqing Li
School of Electronics Engineering & Computer Science,
Peking University,
Beijing, 100871, People's Republic of China,
E-mail: daphne_8896@163.com

Abstract—**The abstract presents a fully functional Ramsey-CPT atomic clock, which has the characteristics of compact structure, small volume, low cost, strong adaptability, high stability, convenient for adjustment, unique design, and easy-to-commercialization. The Ramsey-CPT atomic clock combines the method of separated oscillatory fields and the phenomenon of coherent population trapping, so as to obtain narrower line width and better signal-to-noise ratio of the resonance signal for improving the frequency stability. It not only keeps the advantages of miniaturization and low power of the CPT atomic clock, but also can obtain higher frequency stability than CPT atomic clock about one order of magnitude.**

I. INTRODUCTION

Coherent population trapping (CPT) phenomenon is a quantum interference generated by atoms interacting with coherent light [1]. A new chip-scale passive atomic clock implement is currently a frontier technology in the field of atomic clocks. However, the CPT atomic clock is, at present, theoretically the only one that can be realized as a chip-scale atomic clock, by making use of good coherent properties of laser to prepare atomic "CPT state".

The scope of application of subminiature, even chip-scale atomic clock can be greatly expanded, especially used in measuring instruments, the portable communication and the navigation equipments.

In this system to be described, two coherent optical radiation fields are applied to a low-cost natural Rubidium (Rb) atoms ensemble. The fields are in resonance with the D_1 transitions [2] between the two hyperfine levels of the ground state and one of the hyperfine levels of the excited state, forming a so-called Λ system.

This paper aims to do some further research on Ramsey-CPT phenomenon and to promote the study of device level and producibility of Ramsey-CPT atomic clock. The Ramsey-CPT atomic clock combines the method of separated oscillatory fields [3] and the phenomenon of coherent population trapping, so as to obtain narrower line width and better signal-to-noise ratio of the resonance signal for improving the frequency stability. Using vertical-cavity surface-emitting laser (VCSEL) [4] as optical source, we propose an appropriate physics package and design a comprehensive servo system, forming a fully functional Ramsey-CPT atomic clock for the purpose of developing high-performance and commercialization of Ramsey-CPT atomic clock.

II. THEORETICAL ANALYSIS

A. Operating Principle

Fig. 1 shows the schematic illustration of the practical realization system of the Ramsey-CPT atomic clock. The frequency conversion circuit and the digital step attenuator are controlled by the microcontroller to generate the required microwave signal. With the digital step attenuator generating any form of microwave pulse signal, we can achieve more superior discriminator curve. Through the Bias-Tee, the signal couples with the drive current of the VCSEL to achieve two coherent lights, which phase difference is constant and the frequency difference is equal to the frequency of the microwave signal.

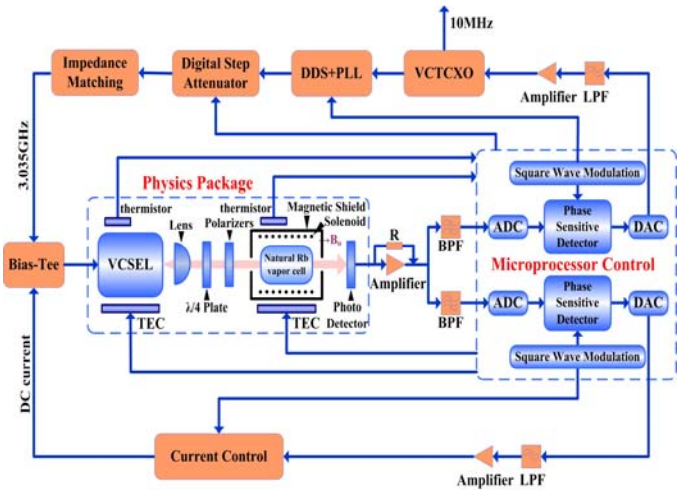


Figure 1. Schematic illustration of the practical Ramsey-CPT atomic clock. The bias-tee is used to couple the VCSEL injection current and the microwave modulation frequency. The λ/4 plate provides circular polarization required in the excitation process by transition selection rules. B_o is the applied small magnetic induction providing an axis of quantization for the atomic ensemble.

The form of the CPT system consists of two parts: the physics package and the servo circuit system. The physics package forms the optical path of laser beam from the VCSEL emitter to the photo receiver, and it consists of the laser used to generate the coherent light, the alkali metal atoms cell, the photodetector used to receive the radiation, and so on. The servo circuit system includes two locking loop: (1) laser frequency locking loop; and (2) microwave frequency locking loop. The first loop is to lock the laser wavelength to the atomic D1 transition line for ensuring the stability of CPT incentive. The second loop is to lock the modulated microwave frequency to the CPT hyperfine resonance frequency for providing a high-performance frequency signal.

B. Comprehensive Analysis

Based on the related data and previous studies [2], we can obtain the relationship between characteristics of Ramsey-CPT system and influencing factors as shown in Fig. 2, which can be used as an important theoretical guide for future design and optimization of the system.

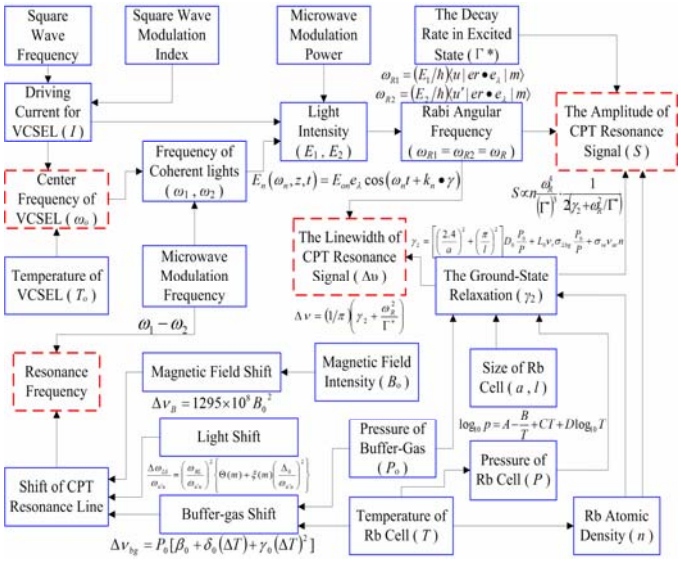


Figure 2. Block scheme of comprehensive analysis of the system.

As is clearly seen in this block scheme, the influencing factors are complex with each other. The shift of CPT resonance line consists of light shift, buffer-gas shift, and magnetic field shift. The line width and the amplitude of CPT resonance signal are the main object of comprehensive analysis of the Ramsey-CPT system.

The factors to be considered are mainly in: (1) the operating temperature of VCSEL and Rb vapor cell, (2) the modulating frequency of microwave, (3) the transmitted light intensity, (4) the weak magnetic field, and (5) the driving current of VCSEL.

Therefore, in practice, we need overall consideration to develop a high-performance Ramsey-CPT atomic clock.

III. IMPLEMENTATION METHOD

A. Physics Package

Depending on the size and volume limited and considering the organic combination with servo circuit system, we realize a physics package required by the system in which the laser beam can be collimated and adjustable, both temperature and magnetic field can be digital quantization, and the light intensity can be stepping adjustments.

Fig. 3 shows the photograph of the physics package which consists of three parts. The first part includes VCSEL, lens, and λ/4 plate. We can control the diameter of the light spot by adjusting the distance from lens to VCSEL. The second part consists of two polarizers which can quantitatively change the light intensity by rolling the bearing with 32-divided scale on it. The third part includes magnetic shield, Rb vapor cell, solenoid, magnetic sensor, and photodetector. The magnetic sensor is used to monitor the weak magnetic field. The magnetic shield, which is produced to a cavity using permalloy, can reduce the effects brought by stray magnetic field and earth's magnetic field.

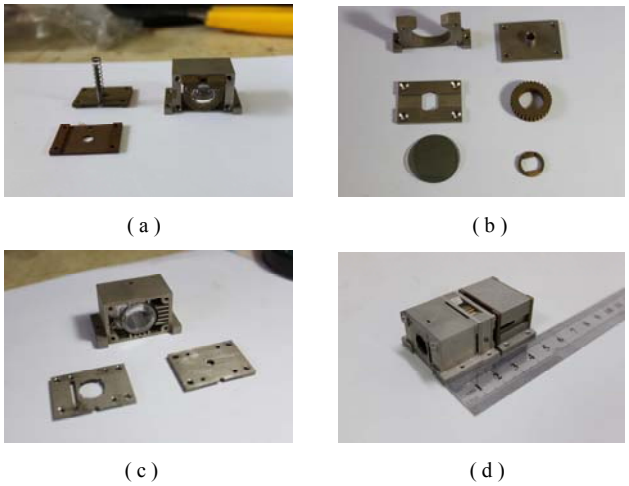


Figure 3. Photograph of the physics package. (a) The first part; (b) The second part; (c) The third part; and (d) the whole physics package.

The beam divergence of VCSEL is about 10 ° – 25 °, so we need to collimate the beam via a collimating lens. A plano-convex lens with 10mm focal length is used here. The beam, after passing through the lens, becomes approximate parallel beam. Then, the laser beam is circularly polarized and attenuated. Finally, the photodetector is required to detect the laser beam which has interacted with Rb atoms.

In short, the physics package designing is the key for obtaining a subminiature, adjustable and measurable system. It can be further development.

B. Comprehensive Servo System

Using comprehensive design method of microwave and high precision circuit and the technology of high density printed circuit board (PCB) designing, we design and implement comprehensive servo system of Ramsey-CPT atomic clock as shown in Fig. 4.

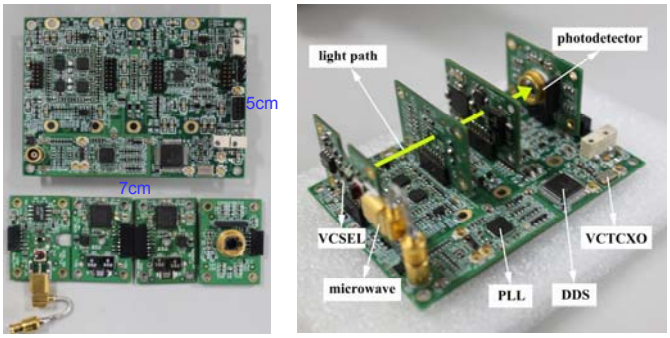


Figure 4. Photograph of the comprehensive servo system.

The servo circuit system consists of five printed circuit boards, which are VCSEL driving circuit, temperature control circuit for VCAEL, temperature control circuit for Rb cell, photo detector and magnetic field monitor circuit, and motherboard circuit. They are connected by the four connectors on the top of the motherboard.

The small and digital servo system has been designed for realizing automatic scan and locking control, monitoring the temperature and magnetic field, and so on.

The detection signal outputted from the physics package is amplified to an appropriate range, sampled by ADC, and then sent to the microprocessor to make the phase sensitive demodulation. At this point, we can obtain the differential curve of the absorption spectrum which is regarded as correction curve for locking the laser frequency. Finally, the feedback loop can be closed and locked. Based on the same approach, the microwave laser modulation frequency can be identified and then locked to the maximum of the CPT resonance transmission line.

C. *The Integral System*

Make the physics package connect to the servo system, and then we can obtain the overall architecture of the fully functional Ramsey-CPT atomic clock system. Its photograph can be seen in Fig. 5. This CPT atomic clock system has a volume about 60cm^3, and is convenient to adjust and measure.

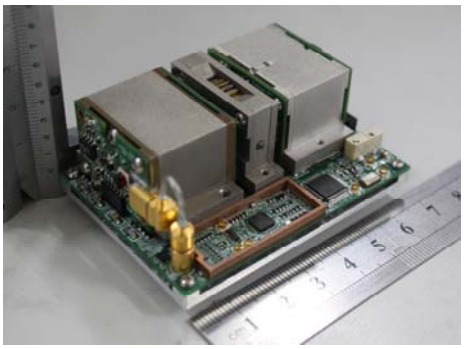


Figure 5. Photograph of the integral system.

In additional, based on the integral system, the chip-scale atomic clock can be developed through simplifying the structure of the physics package and the digital servo circuit, especially the physics package, because they have enough space for improvement.

IV. CONCLUSION

In conclusion, we describe the operating principle and the comprehensive analysis of the Ramsey-CPT system. Based on these theories, we have outlined the practical realization and well characteristics of the fully functional system. The whole system has the characteristics of compact structure, small volume, low cost, strong adaptability, high stability, convenient for adjustment, unique design, and easy-to-commercialization. It is believed that these characteristics of the Ramsey-CPT system can be improved in future developments.

REFERENCES

[1] E. Arimondo, in *Progress in Optics*, edited by E. Wolf (Elsevier, Amsterdam, 1996) , Vol. 35, p. 257.

[2] J. Vanier, "Atomic clocks based on coherent population trapping: A review," *Appl. Phys. B*, vol. 81, pp. 421–442, 2005.

[3] N. F. Ramsey, "A Molecular Beam Resonance Method with Separated Oscillating Fields", Phys. Rev., vol. 78, p. 695, 1950.

[4] D. K. Serkland, G. M. Peake, K. M. Geib etc. "VCSELs for Atomic Clocks," in *Proc. of SPIE*, Vol.6132 613208-1, 2006.

Application of Precision Engineering in laser transverse pumping ^{133}Cs brain magnetometers

Takaya Watanabe
World Technology Instrument Co., Ltd.
533 Kamoshidacho, Aoba-ku, Yokohama city, 227-0033 Japan

Abstract

The transverse relaxation time of the laser pumping of ^{133}Cs atomic vaporing. Even if the temperature rises, the extension of the energy distribution of the electron is suppressed the effect of a thermal extension by the quantization and can lower the oscillation threshold to the quantum-dot structure of the VCSEL without causing the decrease in the gain to the utmost limit. The design of the system that measures the brain magnetism at 100fT level. The cause brain magnetic field from 10^{-13}T to 10^{-14}T or more can be detected.

1. INTRODUCTION

The first step of Table1 shows various magnetic fields in the natural world, and an artificial magnetic field and the third step show the range of motion of the magnetic field weighing device in the second step. If the effect of Hanle of happening when the pumping magnetometer on electron does the diffusion conduction while shining and doing the precession movement is used, the sensitivity of 10^{-16}T can be obtained. Figure 1 and Figure 2 show the principle of transverse optical pumping. [1] Only slight element H_0 (Permanent magnet) of the direction of z remains about 0 when sweeping it so that the magnetic field element of the direction of x may pass 0 when a horizontal pumping to enter the circularly polarized light from the semiconductor of luminescence on the side laser into direct-current magnetic field H_0 vertically is done and the transverse magnetization is detected. In this paper, the optical pumping magnetometer researches the practical use of the diagnosis of the bodily function of the brain magnetism.

2. ^{133}Cs brain magnetometer

Figure 3 shows the magnetometer of ^{133}Cs developed brain. ^{133}Cs gas absorption cell that is the sensor of this brain magnetometer is put in a permanent magnet.

There is a transverse pumping to make the pumping light a circularly polarized light as a method of greatly improving the sensitivity of the optical pumping magnetometer and to enter into the direct-current magnetic field vertically. As for the magnetometer of the brain, the one that operates at the same temperature as the temperature of the living body is desirable. 10ms or more is necessary to monitor the neuron activity the transverse relaxation time of the laser pumping of ^{133}Cs atomic vaporing. The design of the system that measures the brain magnetism at 100fT level is as follows.

Figure 4 heats the operating temperature of the gas absorption cell of which the easing gas is neon to the Cs atom to 28℃-46℃, passes the gas absorption cell the VCSEL (Vertical Cavity Surface Emitting Laser) light of 852.1 nm in luminescence wave length modulated in the gas

978-1-61284-110-6/11/$26.00 ©2011 IEEE

absorption cell by the Raman transition frequency, and shows the relation between the sealing pressure and the easing time of the neon when Laser pumping is done. The easing time becomes the lowest at the time of a temperature it near 28.4℃ in melting point of the ^{133}Cs atom 28℃.

Figure 5 is a luminescence spectrum of VCSEL. Even if the temperature rises, the extension of the energy distribution of the electron is suppressed the effect of a thermal extension by the quantization and can lower the oscillation threshold to the quantum-dot structure of the VCSEL without causing the decrease in the gain to the utmost limit. [2]

The photocount distribution that was wider than the Poisson distribution becomes a subPoisson distribution with only the half in the state of coherence by the negative feedback of the semiconductor laser before putting the negative feedback on the semiconductor laser. [3] Figure 6 puts the negative feedback on the semiconductor laser and is a result of having precisely measured the noise level of photocurrent directly after the detecting phase as for the output light. A usual excitation method is asymptotic in the standard quantum limit of the coherent state. If the pump swing is suppressed with negative feedback loop, the standard quantum limit can be exceeded as shown in Figure 7.

Figure 8 is two step type A-D converter of the pipelined architecture that processes the lamor frequency signal. The A-D conversion of high speed and high resolution is achieved. It is sampled by track/holding (T/H) circuit, and the input signal is maintained as pipeline T/H, and controlled as an analog signal voltage value in the timing control part. After the voltage value is converted into a digital value of m bit with the pre. A-D converter, A-D is converted into the analogue signal again with the D-A converter. The difference with the analogue signal that delays the analogue voltage value maintained in this analogue signal and track/hold circuit by an analog delay line is requested. The rest error amplifier converts the voltage of this difference into a digital value of m bit with the main A-D converter. The output of these two A-D converters is matched and the A-D conversion value of the m+n bit is obtained. The signal from ADA and main A-D enters the error correction machine. The signal from the error correction machine becomes a digital output. An analog DeLay line is necessary so that it may make amends for the operation delay time of the D-A converter following pre A-D converter and it.

Figure 9 is an analog lowpass filter that passes DC-40Hz of the lamor frequency signal. The analogue signal is processed with the low pass filter by the FIR method of the sampling frequency 80Hz. In this low pass filter, the signal of the frequency 43Hz or more has been suppressed to 60dB or more. As for the signal that passed this low pass filter, the image is processed as image brain magnetic wave signal of the A-D conversion.

Figure 10 shows the measured brain wave. A horizontal axis is an alpha wave, a beta wave, and the ordinate is strength of the magnetic field of the brain. The brain wave shows the effect of the ionization current that happens by the dendritic process of about 50,000 neurons in the synaptic transmission.

3. Magnetic shaking

Figure 11 shows magnetic shaking mechanism. [4] Magnetic shaking Magnetic shaking is to put a big exchange magnetic field more than to some

degree on the magnetic substance because it always keeps the magnetic wall of the magnetic substance showing in Figure 11 in vibrating state of the movement. It explains the mechanism of this magnetic shaking by using the relation between a distance and magnetic energy of the magnetic wall. As for energy e of the magnetic wall, the resilience that the magnetic wall receives by the function of distance s of the magnetic wall is shown with de/ds. Does the magnetic wall move very greatly and the size of the most important permeability in the magnetic shielding is decided to pressure for each unit magnetic field that joins the magnetic wall. For instance, the magnetic wall turns regularly in the abcd loop in the state that only the shaking magnetic field hangs when the resilience of the magnetic wall is shown typical by the harmony of the sine function and the linear function like figure. Under such a condition, the loop moves easily by several cycles and changes into the state of efgh if magnetic field ΔH of a small amplitude whose frequency is lower than the shaking magnetic field that corresponds to small pressure Δp (= 2IsΔH and Is are the saturation magnetizations) at the difference level of f point and b point is put further. This increases and is shaking under magnetic shaking rapid the incremental permeability to a minute magnetic field.

4. Simulation with thermic fluid

Figure 12 shows the outline of two nonstationary dimensional analyses process where two or more work-sheets were made to correspond to the time series. When the work-sheet group until prescribed time completes once, all the calculations are automatically executed only by changing the value of the parameter that calculates.

Conclusions

In a magnetic shaking container where the vibration magnetic field where a part of the brain magnetic field measurement part by an optical pumping had been liberated was impressed, this developed bodily function diagnosis device measures the magnetic field distribution of a primary change according to the nervous activity of brain neurons. The cause brain magnetic field from 10^{-13}T to 10^{-14}T or more can be detected.

References

[1]. Claude Cohen-Tannoudji “Atoms In Electromagnetic Fields (World Scientific Series on Atomic, Molecular and Optical Physics)”Second edition 2004

[2]. H.Fujino, S.Koh,S.Iba,T.Fujimoto, and H.awaguchi,” Circularly polarized lasing in a (110)-oriented quantum well vertical-cavity surface-emitting laser under optical spin injection,”Appl.Phys.Lett.vol.94, no.13,131108, 2009

[3]. Y.Yamamoto,N.Imoto, and S.Machida, “Amplitude squeezing in a semiconductor laser
using quantum nondemolition measurement and negative feedback,” Phys.Rev. A,Vol.33. no.5, 3243 (1986)

[4]. I.Sasada, S.Kubo, R.C.O’Handley, and K.Harada, “Low-frequency characteristic of theenhanced incremental permeability by magnetic shaking,” J. Appl. Phys. 67(9), 5583(1990)

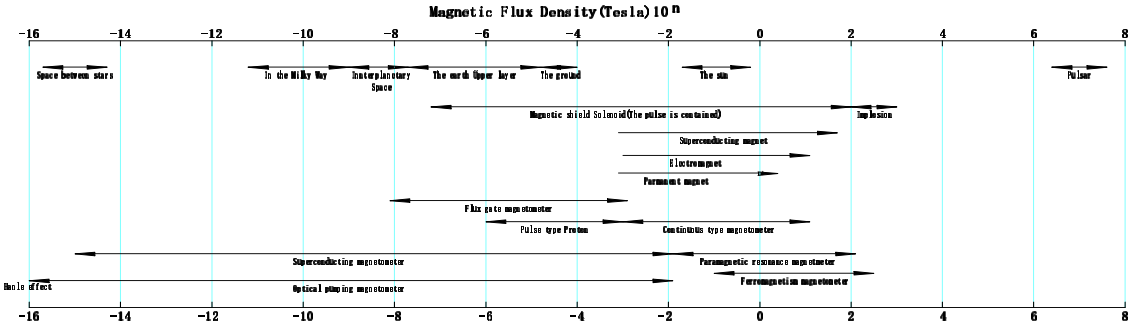


Table1. Range of magnetic field

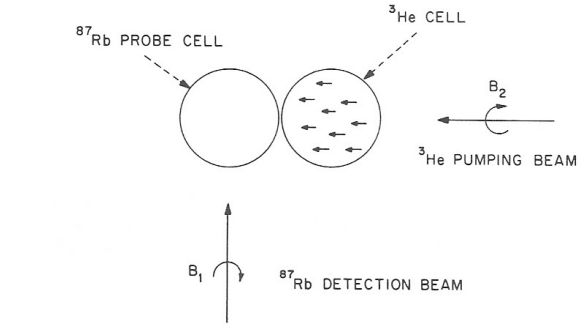


Fig.1 Transverse pumping principle chart

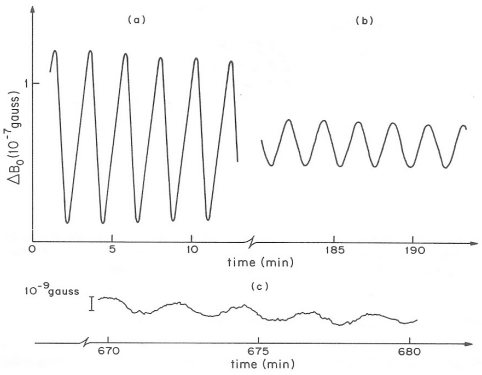


Fig.2 Result of a measurement

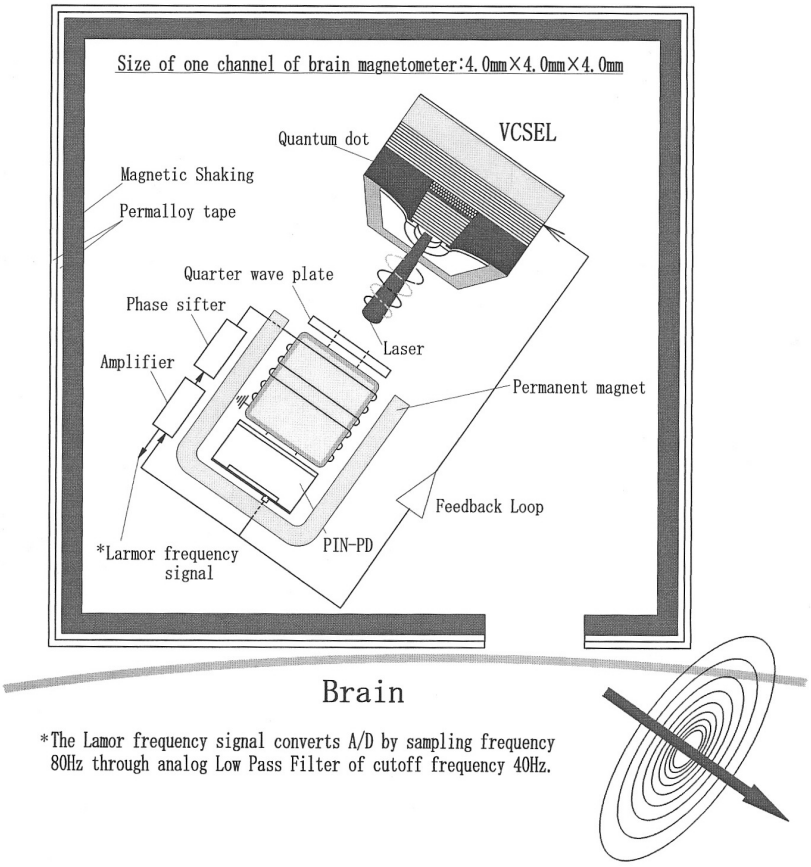


Fig.3 Developed ^{133}Cs brain magnetometer

T2: Transverse relaxation

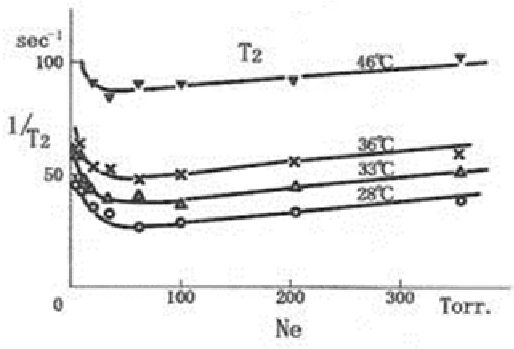


Fig. 4 Dependence of Ne buffer gas pressure of T_2

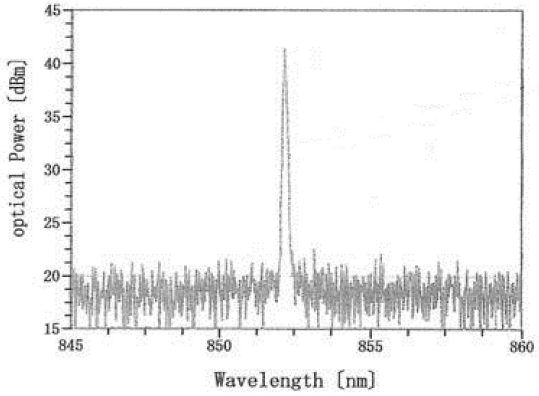


Fig.5 Luminescence spectrum of quantum dot VCSEL

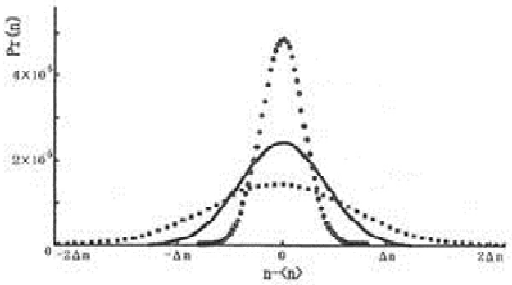


Fig.6 Photoelectron statistics for laser

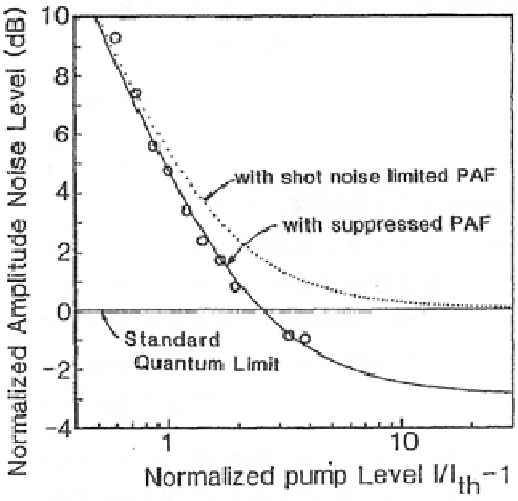


Fig.7 Squeezing of photon-number in feedback loop laser

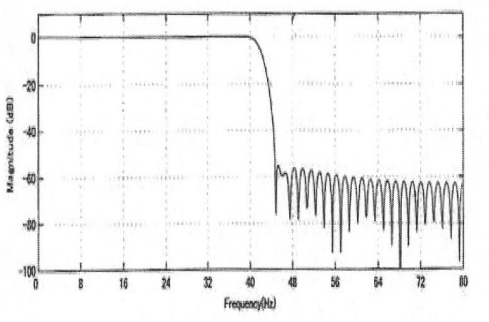


Fig.8 Analog low pass filter that passes DC-40Hz

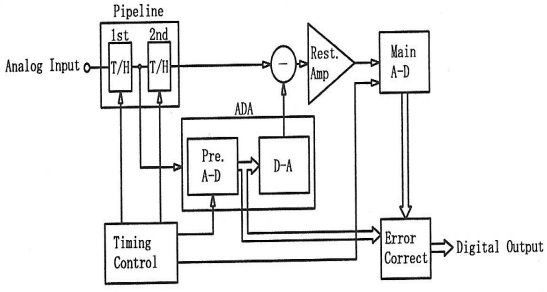


Fig.9 Two step type A-D-converter

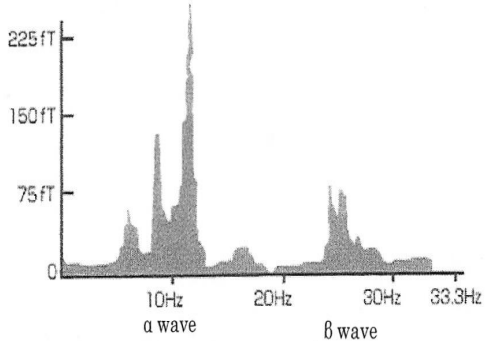


Fig.10 The measured brain waves

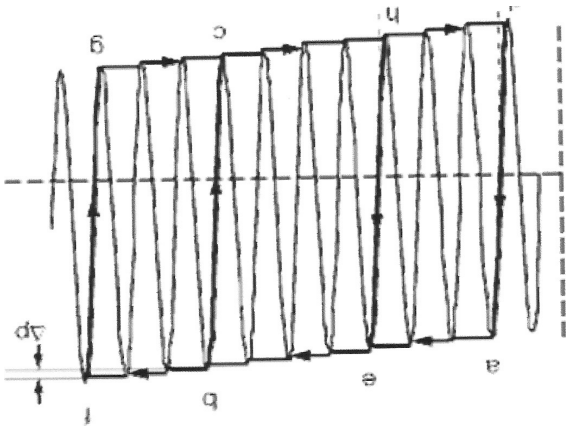


Fig.11 Magnetic shaking mechanism

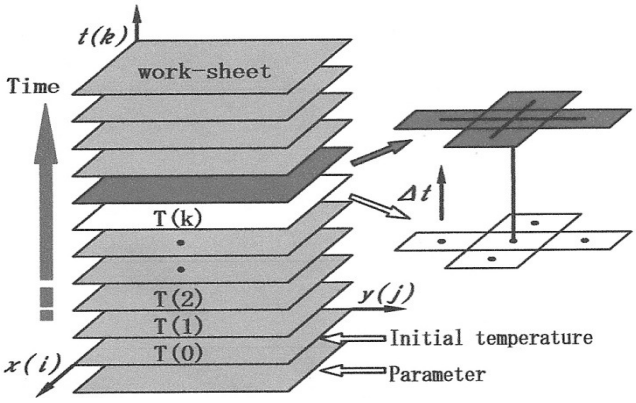


Fig.12 Outline of two nonstationary dimensional analyses process

Polymer coated thin film plate acoustic resonators (FPAR) for gas sensing applications

L. Arapan, I. Katardjiev, V. Yantchev
Dept. Solid State Electronics
Uppsala University
751 21 Uppsala, Sweden
veya@angstrom.uu.se

G. Alexieva, V. Strashilov
Dept. Solid State Physics and Microelectronics
University of Sofia
1164 Sofia, Bulgaria

I. D. Avramov, E. Radeva
Georgy Nadjakov Institute of Solid State Physics
1784 Sofia, Bulgaria

Abstract— **Mass sensitivity of thin aluminum nitride (AlN) film S0 plate wave resonators is theoretically and experimentally studied. Here, two-port 888MHz synchronous thin film plate acoustic resonators (FPAR) are micromachined and subsequently coated with plasma-polymerized hexamethyldisiloxane (pp-HMDSO) thin films of various thicknesses. Systematic data on frequency shift and insertion loss versus film thickness are presented in a comparative manner. Measurements in gas phase environment are further presented in a comparative manner.**

I. INTRODUCTION

During the past 20 years, micro-acoustic gas sensors utilizing various types of surface acoustic waves (SAW) have demonstrated improved performance [1]. As the wave field of SAW is confined to the vicinity of the substrate surface, their propagation characteristics become sensitive to any kind of surface perturbation and more specifically to mass loading. The detection is performed in either delay-line or resonator-stabilized oscillation, the latter being preferred in gas sensing due to the higher Q and lower device loss resulting in substantial noise reduction and resolution improvement. At present, these sensors are well known and have already entered the stage of commercial interest [2]. Quartz-based SAW gas sensors exploiting Rayleigh SAW (RSAW) and SH-SAW (Love, STW) have been extensively studied [3]. The results indicate sensor resolutions in the ppb range due to high sensitivity and very low noise levels [4,5].

Contemporary thin film electroacoustic technology brought a new dimension to the acoustic wave sensors development. Thin film bulk acoustic resonators (FBAR) operating in the lower GHz range, employed in silicon-integrated miniaturized gas sensors, have demonstrated high mass sensitivity [6, 7].

This work is supported by the Swedish Research Council (VR) and the National Scientific Fund of Bulgaria under contract VU-F-203/2006.

In this paper we study the mass sensitivity and the applicability of thin film plate acoustic resonators (FPAR) in gas sensing applications.

FPARs represent a class of thin film acoustic devices utilizing the lowest order symmetric Lamb wave (S0) propagating in thin aluminum nitride (AlN) membranes [8,9]. The S0 mode is known for its low dispersion, high velocity (10000 m/s) and moderate electromechanical coupling [10] which makes it a high velocity alternative to the conventional SAW. Further, the FPARs topology, while similar to that of the SAW-based devices (see Fig. 1), is micro machined on AlN membranes on a silicon (Si) substrate [9] thus making it compatible with integrated circuits.

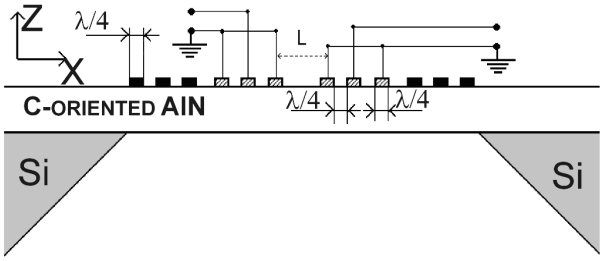


Figure 1. 2 Port FPAR Schematic

Two-port FPARs have demonstrated insertion losses of around 5dB and an unloaded Q of 3000 at a frequency of 0.9GHz [10]. Also, 0.9GHz FPAR-based oscillators have exhibited low phase noise [11]. These results, along with the expected high mass sensitivity of the S0 Lamb mode in acoustically thin plates, motivated us for the present study. Here the mass sensitivity of the FPAR is theoretically and experimentally studied in view of combining high sensitivity

with low losses. The results are further supported by measurements in gas phase environment.

II. THEORETICAL CONSIDERATIONS

The mass sensitivity of the S0-Lamb wave in an acoustically thin plate loaded by a sensitive layer is described by using the equations of motion complemented by a full set of boundary conditions for a waveguided solution with polarization in the XZ-plane. Details are omitted here for brevity, since they have been thoroughly discussed elsewhere [12]. Mass sensitivity is assessed through the fractional change of the acoustic velocity as a function of the sensitive layer thickness. It is revealing to calculate and compare theoretically the mass sensitivity of the S0 Lamb wave to those of the RSAW and Love wave propagating on layer coated AT-cut quartz. The RSAW calculation can readily be done using the algorithm described in [13], while the Love wave calculation employs the layer boundary conditions derived in [14].

In this specific case an AlN acoustic plate with thickness-to-wavelength ratio d/λ=0.167 as well as an AT-cut quartz substrate, both loaded with the glassy polymer pp-HMDSO are theoretically considered [15]. We found that that the S0 mode demonstrates a considerably higher mass sensitivity as compared to the RSAW and Love waves. However, increased sensitivity must not necessarily be the best choice for a practical system. Generally, improved mass sensitivity should correlate with improved damping especially at thicker layers [16]. This aspect is further theoretically studied for the S0 Lamb mode in comparison to RSAW and Love waves on AT-cut quartz. The S0 mode attenuation is found comparable to that of the RSAW and Love waves for relatively thin layers h/λ < 1%. At larger layer thicknesses the S0 attenuation is significantly higher than the attenuation of the RSAW mode. The latter determines the frame where the use of S0 Lamb wave can be beneficial. The use of relatively thin sensing films seems to be preferable, similarly to the case of surface transverse waves (STW) [16].

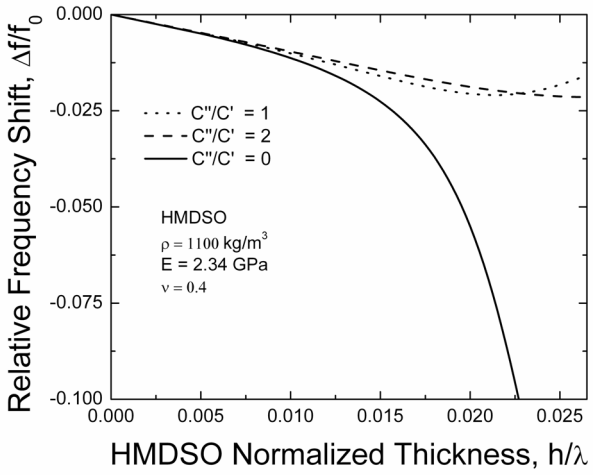


Figure 2. Theoretical mass sensitivity as function of layer viscosity.

Further, the layer viscosity itself influences the S0 mass sensitivity especially for pp-HMDSO layers with relative thicknesses h/λ > 1% (see Fig. 2). In this range the sensitivity deteriorates significantly compared to the case of a lossless layer. This effect is explained by the field distribution of the dominant longitudinal component within the composite plate as calculated by the algorithm proposed in this work. It shows that the viscosity tends to decrease the energy trapping to the low impedance layer, which in turn reduces significantly the sensitivity at thicker layers. Quite a similar effect has been discussed for polymer coated SH-SAWs [13] although its nature has not been discussed so far.

III. FPAR EXPERIMENTAL LOADING BY PP-HMDSO

In this section we compare the theoretical relative frequency shifts of the S0-mode with experimental values. The experiments were performed with a 2-port FPAR with acoustic wavelength λ=12μm (grating pitch 6μm). The FPAR device was micromachined on a freestanding AlN membrane with a thickness d=2μm resulting in a d/λ value of 0.1667. Then the FPAR has been coated with pp-HMDSO thin films of various thicknesses. The actual thickness of the sensing films has been measured during the deposition. After each deposition step the frequency shift as well as the insertion loss were recorded with a Vector Network Analyzer. In Fig. 3 the experimental frequency shift is compared to the theoretically determined one presuming a lossless pp-HMDSO film. For layer thicknesses h/λ < 1% the agreement between the lossless theory and experiment is excellent, while for thicker layers a progressive discrepancy is observed. This point comes to support the loss of sensitivity mechanism discussed in the previous section. The experimental data is seen to be in excellent agreement with the lossy theoretical predictions for $C''/C' \approx 1$ (ratio between imaginary and real part of elasticity).

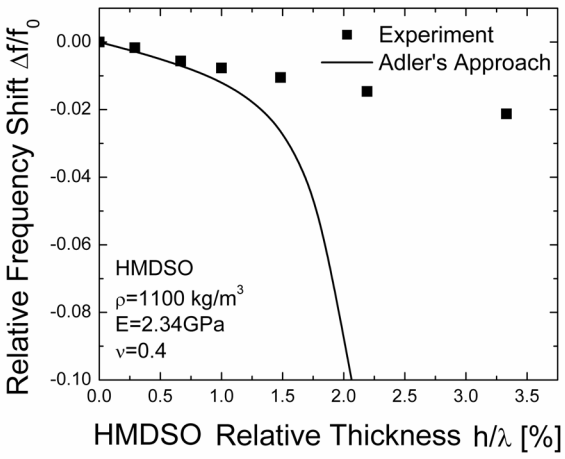


Figure 3. Theoretical lossless vs experimental mass sensitivity

In Fig. 4 the recorded changes in FPAR insertion loss as a function of the pp-HMDSO thickness are shown. Relative losses of up to 2dB are observed for layer thicknesses h/λ<1%. In Fig. 5 the close-to-resonance frequency response of a 2-port synchronous FPAR is shown prior the deposition of an

acoustically thin pp-HMDSO layer with h/λ=0.6%. The FPAR demonstrates around 4 MHz (4600ppm) frequency downshift at the expense of only 0.6 dB loss increase due to the HMDSO loading. The coated FPAR demonstrates an insertion loss of around 7dB and a loaded Q of 700. Note that an FPAR with such a performance was readily employed in a low noise frequency generator, which is a prerequisite for achieving high measurement resolution.

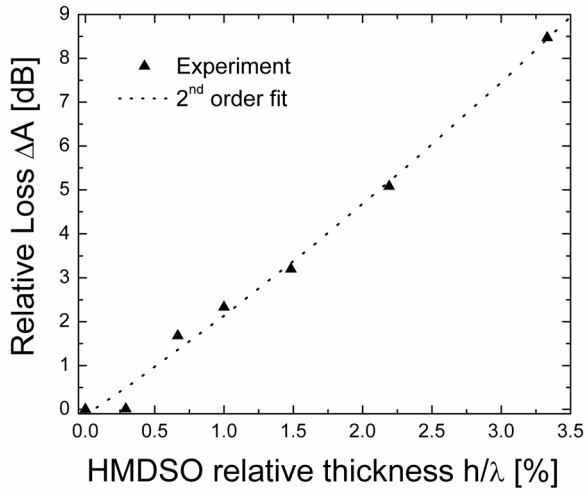


Figure 4. Change in insertion loss of a pp-HMDSO coated FPAR

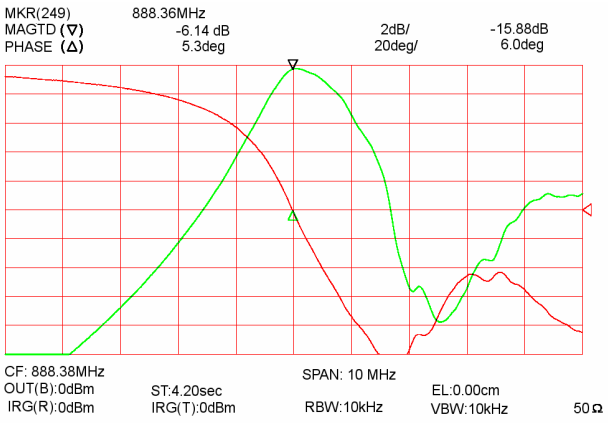


Figure 5. Narrowband frequency (upper curves) and phase (lower curves) characteristics of a 2-port FPAR

IV. GAS SENSITIVITY EXPERIMENTS

Here we further considered a comparison of the FPAR gas sensitivity to that reported for other HMDSO-coated resonant devices using different types of acoustic modes. In [4] HMDSO-coated STW resonators at 700 MHz and SAW resonators at 430 MHz have been tested for their reaction against a number of gases including xylene. That study was aimed at achieving optimum HMDSO-layer thickness promoting maximum sensitivity. At the optimum thickness the wave energy is concentrated at the surface area of the film and strongly interacts with the sorbed gas [4]. Optimum values of HMDSO thickness were 100 nm (h/λ=1.4%) for the STW case and 280 nm (h/λ=3.9%) for SAW. The resulting sensitivities for xylene have been 6.4 Hz/ppm and 3.1 Hz/ppm, respectively. The conclusion drawn has been that STW-resonators have higher gas sensitivities at the same wavelength – a result that is in direct agreement with the theoretical comparison here, although Love waves are chosen as a representative of the SH-SAW family. To check this comparison experimentally, an FPAR device with an HMDSO layer as thick as 380 nm (h/λ=3%) was tested. The frequency downshift due to the layer deposition was 27 MHz (30 000 ppm). The HMDSO layer has not been optimized with respect to gas sensitivity and represents a relative thickness (h/λ) somewhere in between optimum values for STW and SAW, respectively. The device was gas probed with xylene vapors. The non-flow experimental method of applying the gas that we used for xylene probing has been described in detail in [17]. The changes of the resonant frequency and insertion loss have been recorded by a scalar network analyzer. The experimental sensitivity of the 380 nm HMDSO coated FPAR device to xylene concentration is shown in Fig. 6.

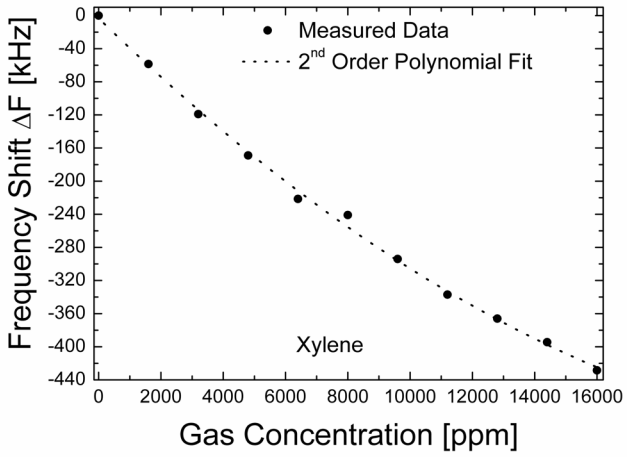


Figure 6. Sensitivity versus xylene for the FPAR with 380 nm HDMSO

The sensitivity curve is nearly linear with a 33 Hz/ppm slope, which is about 5 times higher than the sensitivity of the STW resonator from [4], operating at about the same frequency. At the same time the FPAR loss varied within 1 dB over the whole concentration range, which has a negligible effect on the sensor oscillator performance.

Further, the relative sensitivity to xylene defined as the ratio of the absolute sensitivity [in Hz/ppm] to the unloaded resonance frequency was calculated for FPAR, STW and SAW. These sensitivities were found to be in relation 5.0/1.3/1.0 for FPAR vs. STW vs. SAW, respectively, which is in a good agreement with the theoretical expectations.

It is further to note that HMDSO selectivity has also been assessed within this study. HMDSO coated FPARs have shown sensitivities towards ethanol, acetone and formalin in the range of about 1Hz/ppm which is more than one order of magnitude lower than the sensitivity to xylene. The latter results seem reasonable if one bears in mind that HMDSO has

been known as highly hydrophobic. Xylene molecules, also hydrophobic, should be adsorbed on the polymer surface at much larger quantities than the polar molecules of ethanol, acetone and formalin. Similar results have independently been observed in [18], where pp-HMDSO has demonstrated much higher sensitivity to the hydrophobic (non polar) toluene and diethyl ether as compared to the sensitivity to the polar acetone and ammonia. The above explanation is further supported through gas probing experiments with acetic acid vapors (see Fig. 7).

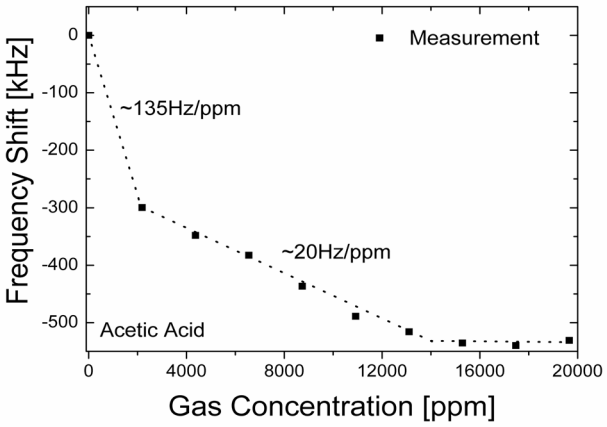


Figure 7. Sensitivity versus acetic acid vapors

The HMSDO-coated FPAR has shown very high sensitivity of 135 Hz/ppm exhibited at low concentrations, up to 2500 ppm, followed by a linear part with slope of 20 Hz/ppm that extends to about 20000 ppm gas concentration, where frequency reached the saturated downshift of – 550 kHz. These features are also noticeable in the behavior of the loss whose increase again does not exceed the 1 dB range. A high reaction of HMDSO towards the acetic acid has been already established by QCM assisted experiments [18] but the effect has not been clarified. To look for a possible explanation, we repeated the check of the FPAR against xylene. The reference equilibrium frequency has been shifted downward by approximately 200 kHz and the gas sensitivity has decreased for higher gas concentrations (see figure 8). The most plausible reason for this kind of behavior is that the covalent surface bonding produced by the acetic acid chemisorption possibly leads to the formation of polar functional groups of silanol and acetic ester types that irreversibly add extra material to the surface. These groups seem also to contribute to adsorption of additional quantities of acetic acid by physisorption saturation mechanism. The covalent bonding ultimately results in a decreased specific adsorption area for xylene molecules in the repeated experiment. This may not be significantly manifested during the first stage of the adsorption (small gas concentrations) but definitely limits the adsorption capacity as observed at higher concentrations.

Further research is needed to explore in details the effects of membrane scaling on the sensor performance.

Note that the weak dispersion of the S0 Lamb mode allows to improve the sensitivity of the FPAR without any change of the resonant frequency, i.e. decreasing the membrane thickness.

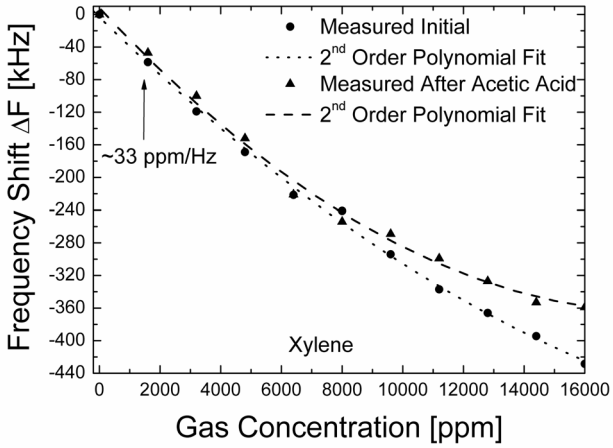


Figure 8. Sensitivity versus xylene before and after acetic acid

This unique opportunity has also been discussed in contour-mode resonant configurations employing the same mode [19]. More generally, the improvement in sensitivity is expected to be combined with reduced affinity to the viscous losses in the sensitive layer, which in turn imposes stronger limitations on the gas sensitive coating thickness. The latter applies also to frequency scaling where the used membranes are thinner and the layer losses larger. In other words, the optimum FPAR sensor design is application specific, i.e. requires a trade off between FPAR mass sensitivity and layer losses in order to achieve maximum resolution.

V. CONCLUSIONS

These initial results from the gas sensing experiments clearly support the theoretical predictions for high mass-sensitivity of the studied resonator type. The considerable sensitivity improvement over one of the most efficient acoustic wave mode gas sensors– the STW, for the same types of polymer film and probing gas, and at comparable frequencies and losses, reveals the great potential of the S0 mode for sensor applications. Another issue of practical importance is the fact that this high sensitivity is accompanied by a weak effect on the resonator loss even at high concentrations of aggressive gas phase compounds interacting with the polymer-coated surface. The authors believe that these features may be attractive for integrated circuit compatible FPAR sensor systems in the near future.

ACKNOWLEDGMENTS

We thank Dr Iliyan Kolev for helpful discussions.

REFERENCES

[1] H. Wohltjen, R. Dessy, “Surface acoustic wave probe for chemical analysis,” Parts I–III, Anal. Chem., vol. 51, pp. 1458–1475, Aug. 1979.

[2] M. Rapp, J. Reibel, A. Voigt, M. Balzer, O. Bulow, "New miniaturized SAW-sensor array for organic gas detection driven by multiplexed oscillators," Sens. Actuators B, vol. 65, pp. 169–172, June 2000.

[3] I. D. Avramov, S. Kurosawa, M. Rapp, P. Krawczak, E. Radeva, "Investigations on plasma polymer coated SAW and STW resonators for chemical gas sensing applications," IEEE Trans. Microwave Theory Tech., vol. 49, pp. 827-837, Apr. 2001.

[4] I. D. Avramov, M. Rapp, S. Kurosawa, P. Krawczak, E. Radeva, "Gas sensitivity comparison of polymer coated SAW and STW resonators operating at the same acoustic wave length," IEEE Sensors Journal, vol. 2, pp. 150 – 159, June 2002.

[5] C-L. Hsu, C-Y. Shen, R-T. Tsai, M-Y. Su, "Surface acoustic wave ammonia sensors based on ST-cut quartz under periodic Al structure," Sensors, vol. 9, pp. 980-994, 2009.

[6] R. Gabl, M. Schreiter, E. Green, H.-D. Feucht, H. Zeininger, "Novel integrated FBAR sensors: a universal technology platform for bio-and gas-detection," Proc. IEEE Sensors, 2003, vol. 2, pp. 1184-1188.

[7] M. Penza, P. Aversa, G. Cassano, D. Suriano, W. Wlodarski, "Thin-film bulk-acoustic-resonator gas sensor functionalized with a nanocomposite langmuir–blodgett layer of carbon nanotubes," IEEE Trans. Electron. Dev., vol. 55, pp. 1237 – 1242, May 2008.

[8] J. Bjurström, I. Katardjiev, V. Yantchev, "Lateral-field-excited thin film Lamb wave resonator," Appl. Phys. Lett., vol. 86, no. 15, art. no. 154103, Apr. 2005.

[9] V. Yantchev, I. Katardjiev, "Micromachined thin film plate acoustic resonators utilizing the lowest order symmetric lamb wave mode," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 54, pp. 87–95, Jan. 2007.

[10] V. Yantchev, L. Arapan, I. Katardjiev, "Micromachined thin film plate acoustic resonators (FPAR): Part II," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 56, pp. 2701–2710, Dec. 2009.

[11] L. Arapan, I. Avramov, V. Yantchev, "Thin film plate acoustic resonators for integrated microwave power oscillator applications," Electron. Lett., vol. 47, pp. 453-454, 2011.

[12] V. Yantchev, I. Katardjiev, "Propagation characteristics of the fundamental symmetric Lamb wave in thin aluminum nitride membranes with infinite gratings," J. Appl. Phys., vol. 98, art. no. 849101, 2005.

[13] J. Kondoh, S. Shiokawa, M. Rapp, S. Stier, "Simulation of viscoelastic effects of polymer coatings on surface acousticwave gas sensor under consideration of film thickness," Jpn. J. Appl. Phys., vol. 37, pp. 2842–2848, 1998.

[14] V. Yantchev, V. Strashilov, "Surface transverse waves in polymer-coated grating configurations," J. Appl. Phys., vol. 91, pp. 5700–5705, 2002.

[15] D. Hegemann, H. Brunner, C. Oehr, "Plasma treatment of polymers to generate stable, hydrophobic surfaces," Plasmas and Polymers, vol. 6, pp. 221-235, 2001.

[16] V. Yantchev, V. Strashilov, M. Rapp, U. Stahl, I. D. Avramov, "Theoretical and experimental mass-sensitivity analysis of polymer-coated SAW and STW resonators for gas sensing applications," IEEE Sensors J., vol. 2, pp. 307-313, Aug. 2002.

[17] V. Strashilov, G. Alexieva, V. Velichkov, R. Mateva, I. D. Avramov, "Polymer-coated quartz microbalance sensors for volatile organic compound gases," Sensors Lett., vol. 7, pp. 203-211, 2009.

[18] H. Aizawa, K. Noda, R. Naganawa, K. Yamada, M. Yoshimoto, "Gas sorption of acetone, diethyl ether, toluene, acetic acid, and ammonia on plasma-polymerized hexamethyldisiloxane films coated with quartz crystal microbalance," Journal of Photopolymer Science and Technology, vol. 22, no. 6, pp. 743-745, 2009.

[19] G. Piazza, M. Rinaldi, C. Zuniga, "Nanoscaled piezoelectric aluminum nitride contour-mode resonant sensors," in Proc. IEEE Sensors Conference, 2010, pp. 2202-2207.

Calibration of Non-contact Temperature-compensated SAW Resonant Torque Sensors

V. Kalinin
Transense Technologies plc
Bicester, Oxfordshire, UK
victor.kalinin@transense.co.uk

***Abstract*—A new calibration method has been proposed for passive wireless torque sensors based on SAW resonators. It requires individual calibration of the sensors at the production stage only at room temperature and individual measurement of the difference frequency drift at zero torque over the operating temperature range. This approach greatly simplifies calibration and reduces its cost in the case of high volume manufacturing. Analysis of calibration statistics for the torque transducers has shown that the total measurement errors achieved in this case are very close to those obtained in the case of the individual sensor calibration in several temperature points over the entire operating temperature range.**

I. Introduction

Passive wireless or contactless sensors based on SAW resonators have been developed and found industrial application for temperature [1, 2], pressure [3] and torque measurement [4, 5, 6]. Recently, contactless resonant SAW torque sensors have been developed for various automotive applications such as electrical power assisted steering [7] and powertrain [8, 9]. Due to technological tolerances of the SAW die manufacturing and packaging processes, the SAW sensors usually require individual calibration of each unit in order to achieve a good accuracy around 1% full scale (FS). It is especially applicable to torque sensors due to additional contribution to the spread of sensor characteristics made by a glue interface between the SAW die and the part where shire strain is measured.

Automotive torque sensors are very demanding from the point of view of calibration since they need to produce a temperature compensated torque reading within a wide range from -40°C to +125°C and sometimes even at higher temperatures. This means that calibration of each torque transducer (application of a known torque) has to be performed at a number of temperature points more or less uniformly distributed within the operating temperature range. In the case of SAW temperature and pressure sensors, a large number of them can be placed into an oven and calibrated simultaneously because of their small size. The size of the SAW torque transducers can on the contrary be quite large so that only very few of them can be placed into the oven and calibrated simultaneously (see Fig. 1 showing torque calibration rig just for one flexplate transducer used to measure the output engine torque [8]). Large thermal mass of the rig including mechanical components required for torque application can increase time needed to reach a steady state temperature at each calibration point up to 2-3 hours so that the whole calibration process can take from 15 to 30 hours. As a result the calibration significantly increases the sensor manufacturing cost making it unacceptable for high volume automotive applications.

One of the solutions of the calibration problem has been recently proposed in [10]. The idea of the method is to reduce the number of the temperature calibration points from five-nine to just two. This approach gives extra calibration errors below ±1% FS for the flexplate transducers if the two-point individual calibration data are combined with the generic characteristics of the SAW sensors. It reduces calibration time at least by a factor of five. Nevertheless it still requires application of torque to a transducer at one elevated temperature, e.g. +80°C. This significantly complicates design of a calibration rig and increases the calibration time in comparison with calibration only at room temperature.

Figure 1. Calibration rig for the SAW flexplate torque transducer

978-1-61284-110-6/11/$26.00 ©2011 IEEE

The aim of this paper is to propose an alternative method of calibration when torque is applied to each individual SAW sensor only at room temperature. This method also relies on previously measured generic temperature characteristics of a particular type of the senor. It has been developed for torque sensors used in kinetic energy recovery system (KERS) of Formula 1 vehicles.

The structure of the paper is as follows. Design of the SAW KERS torque sensor and its generic characteristics are described in Section II. The next section presents a gradual development of the calibration method from a trivial approach to a more refined one. The method in its fully developed form is discussed in Section IV where it is applied to a number of the KERS SAW torque sensors. Section V presents some conclusions.

II. SAW KERS Torque Sensor

The sensor has been developed for F1 vehicles to monitor the amount of torque transmitted trough a short shaft from an electric motor to an engine [9]. The main features of this application are: (a) a very high rotation speed up to 18000 rpm, (b) very strong and fast torsional vibrations, (c) the temperature range from +20°C to +160°C, and (d) a limited space available for the sensor installation. These features dictated the sensor design based on two diametrically opposite SAW sensing elements, the 4x6 mm^2 SAW die being directly bonded to the KERS shaft as shown in Fig 2, and the cylindrical co-axial RF rotary coupler.

One of the sensing elements, HFSAW, contains three one-port resonators on Y+34° cut quartz. Two of them, with resonant frequencies $f_1 \approx 437$ MHz and $f_2 \approx 435$ MHz, are at ±45° to the X axis of the substrate and the shaft axis so that $F_{m1} = f_1 - f_2$ linearly depends on the applied torque M. The third resonator, with the frequency $f_3 \approx 433$ MHz, is used for temperature compensation since $F_t = f_2 - f_3$ depends on temperature T differently compared to F_{m1} allowing an independent measurement of T [5].

The second sensing element, LFSAW, contains only two SAW resonators with frequencies $f_4 \approx 431$ MHz and $f_5 \approx 429$ MHz positioned at ±45° to the shaft axis so that $F_{m2} = f_4 - f_5$ also linearly depends on M. Orientation of LFSAW relative to the HFSAW is such that $F_m = (F_{m1} + F_{m2})/2$ is sensitive to torque and insensitive to the bending force applied in both planes.

Each resonant frequency is measured by the interrogation unit wirelessly by means of launching 2 µs RF interrogation pulse through the RF rotary coupler and picking up free oscillations of the resonator after the pulse is over [7, 8]. The resonator excitation, the response sampling and the spectral analysis take 155 µs. The torque values are updated every 310 µs. The standard deviation of the measured difference frequency F_m is approximately 300 Hz.

Typical calibration characteristics of the KERS sensor measured at five temperature points within the torque range ±50 Nm (FS = 100 Nm) are shown in Figs. 3 and 4. The mathematical model of the sensor is as follows:

$$M = \begin{cases} C_p(T)\left(F_m - F_{m0}(T)\right)/(10^3 \times 2^{18}), & F_m \geq F_{m0}(T) \\ C_n(T)\left(F_m - F_{m0}(T)\right)/(10^3 \times 2^{18}), & F_m < F_{m0}(T) \end{cases} \quad (1)$$

$$F_t = 100a_1 - a_2T - 10a_3M/64 + a_4T^2/1600 + a_5T^3/102400 \quad (2)$$

where $C_{p,n}$, the reciprocal sensitivities of F_m to positive and negative torque respectively, the offset $F_{m0} = F_m$ at $M = 0$ and the coefficients $a_{1\text{-}5}$ are found from the measured data by means of LMS method. The numerical coefficients in (1) and (2) are selected in such a way that all the calibration parameters can be represented by two-byte integers in the interrogator's DSP memory. $C_{p,n}$ and F_0 as functions of temperature are stored in the form of lookup tables. Torque and temperature are found from the measured values of $F_{m,t}$ by means of an iterative solution of (1), (2).

The parameters of the model (1), (2) vary from one unit to another one. The degree of variation was investigated in

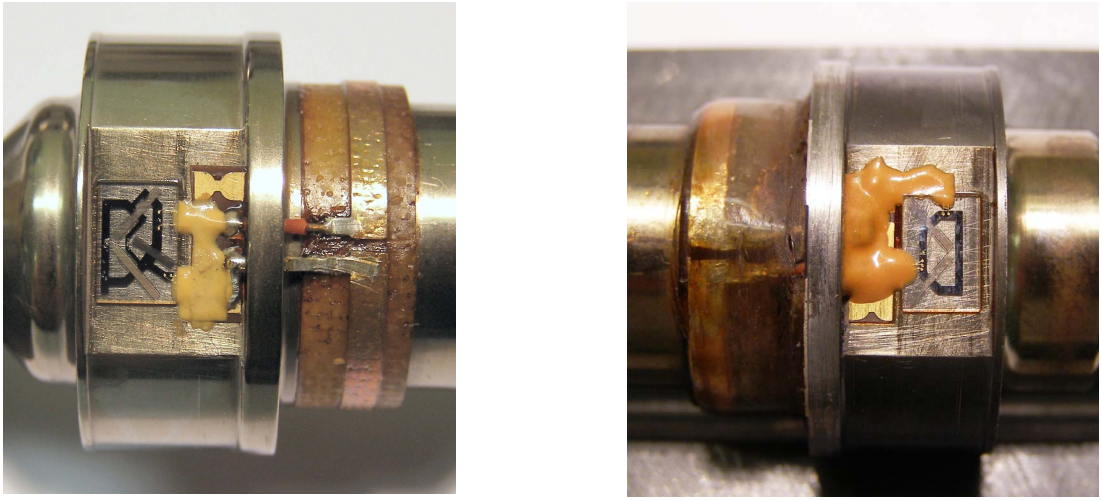
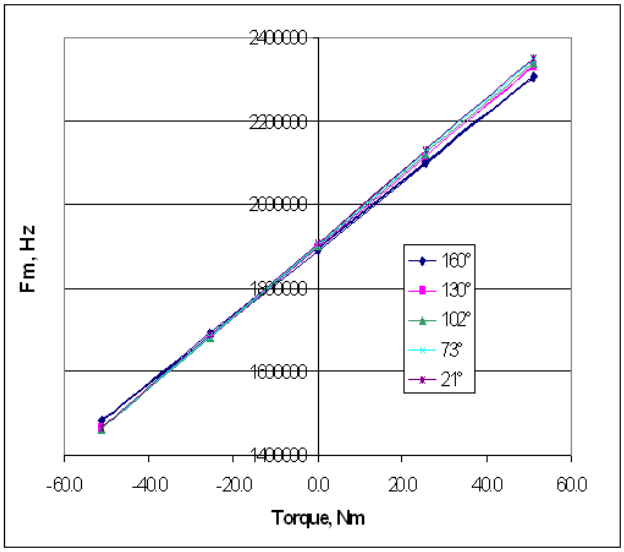


Figure 2. SAW sensing elements, HFSAW (left) and LFSAW (right), bonded to the two opposite sides of the KERS shaft



Figire 3. Variation of F_m with torque at different temperatures

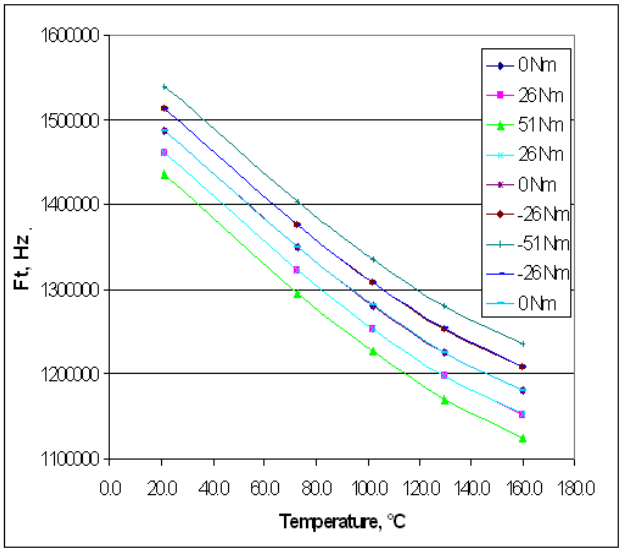


Figure 4. Variation of F_t with temperature at different torque values

[10] for the flexplate transducer and a similar variation has been observed for the KERS sensor. By averaging the values of $a_{1\text{-}5}$ and $C_{p,n}(T)$ for a certain number of units one can find generic calibration parameters for a particular type of a sensor. Tables 1 and 2 show generic calibration parameters for the KERS sensor $a'_{1\text{-}5}$ and $C'_{p,n}(T)$ obtained by averaging the characteristics of 45 units. One can notice that the torque sensitivity of F_m decreases with temperature due to reduction of the stiffness of the glue bond between the SAW die and the shaft.

III. Various Approaches to Calibration

Let us consider several possible approaches to calibration in the order of increasing complexity and investigate torque measurement errors associated with them.

A. Full Individual Calibration

In this case each unit is calibrated at five temperature points. After calibration, one can find the total measurement errors ΔM characterizing deviation of the sensor model (1), (2) from the real characteristics (calibration errors), hysteresis, non-repeatability, noise (it is greatly suppressed by averaging 400 measurement results) and errors in setting M and T. Fig. 5 shows the total torque errors measured for the sensor No. 46 that does not belong to the group of units used for finding the generic calibration parameters. Different curves in Fig. 5 correspond to different torque values from $-M_{max} = -50$ Nm to $M_{max} = 50$ Nm.

TABLE I. Generic Coefficients for F_t

a'_1	a'_2	a'_3	a'_4	a'_5
15400.56	2843.84	5900.64	-4023.09	3027.67

TABLE II. Generic Reciprocal Sensitivities

T	20	50	70	90	120	150	180
C'_p	29993	30042	30108	30221	30533	31482	33462
C'_n	30019	30060	30115	30215	30508	31413	33283

At $T < 100$°C the errors are dominated by non-linearity, non-repeatability and noise. At $T > 100$°C the errors start increasing mainly due to increased visco-elastic hysteresis in the glue bond. As a result, the maximum torque error is around 1% FS at the highest temperature and 0.1% FS at the lowest temperature.

B. Calibration Only at Room Temperature

In this case the sensor is only calibrated at $T_0 = 20$°C so only $F_{m,t}(M, T_0)$ are known. The complete set of calibration parameters is obtained in the following way. It is assumed that the offset $F_{m0}(T) = F_{m0}(T_0) = const$ and $C_{p,n}(T) = C_{p,n}(T_0) = const$ and $a_{2,4,5}$ equal generic values from Table 1, $a'_{2,4,5}$.

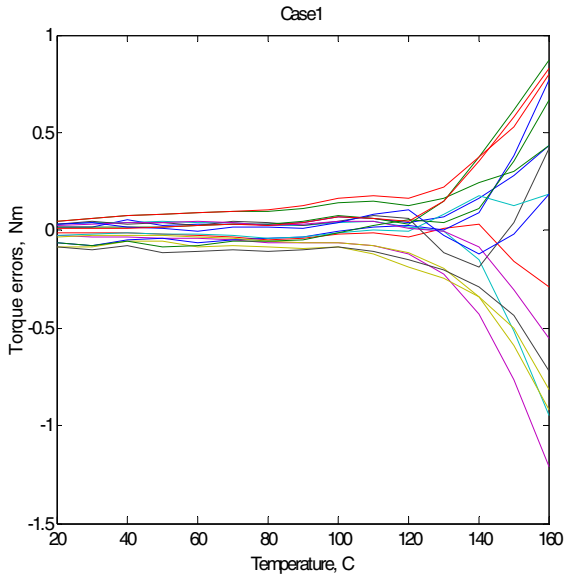


Figure 5. Torque measurement errors in the case A.

The rest of the coefficients are found as follows:

$$a_3 = -3.2[F_t(M_{max}, T_0) - F_t(-M_{max}, T_0)]/ M_{max}, \quad (3)$$

$$a_1 = [F_t(T_0,0) + a'_2 T_0 - \frac{a'_4 T_0^2}{1600} - \frac{a'_5 T_0^3}{102400}]/100. \quad (4)$$

In other words, the thermal variation of F_m is totally ignored and the thermal variation of F_t is assumed to be the same as the generic one. One should bear in mind that there is no point in using a generic variation of the offset F_{m0} with temperature. The spread of the characteristics $F_{m0}(T)$ is so large due to unpredictable modification of the SAW temperature characteristics by the glue bond [10] that this will not improve the accuracy of the calibration model.

The torque measurement errors obtained for the sensor No. 46 in this case are presented in Fig. 6. Clearly, they acquire a large systematic component that increases almost linearly from 20°C to 125°C up to -1.8 Nm. This happens mainly due uncompensated variation of the offset F_{m0} with temperature. Variation of the torque errors with torque at 160° is ±3.8 Nm – both due to the hysteresis and due to the fact that the torque sensitivity variation with temperature is not taken into account and temperature compensation is inaccurate because of using generic coefficients $a'_{2,4,5}$. The total errors are up to 6.5% FS at the highest temperature that is unacceptable for the application under consideration.

C. *Accounting for the Thermal variation of the Offset F_{m0}*

The next step in refining of the calibration method is to take into account the actual characteristic $F_{m0}(T)$ as well as more accurately find the coefficients $a_{1\text{-}5}$. It can only be achieved if the measurement of F_m and F_t is performed for each individual sensor at zero torque in a number of temperature points in addition to calibration at T_0. When doing this, a large batch of transducers can be simultaneously heated and cooled thus reducing the calibration time and cost.

The complete set of calibration parameters is found in this case in the following way. First, it is still assumed that $C_{p,n}(T) = C_{p,n}(T_0) = const$. Second, the complete matrices of raw calibration data

$$F_m(M, T) = F_m(0, T) + 10^3 \times 64 \times 4096 M/C_{p,n}, \quad (5)$$

$$F_t(M, T) = F_t(0, T) - 10 a_3 M/64 \quad (6)$$

are calculated at a number of torque and temperature points, while a_3 is defined by (3). After that, all the calibration parameters are found in a usual way by LMS method.

The torque measurement errors for the sensor No. 46 in this case are shown in Fig. 7. There is no a systematic component of the error varying with temperature any more. This is because variation of the offset F_{m0} with temperature is correctly taken into account and the temperature compensation of torque readings is performed correctly due to using the correct coefficients $a_{1\text{-}5}$ found from $F_t(0, T)$. However, the torque errors grow faster from ±0.1% FS at room temperature to ±1% FS at 125°C compared to the Case A because variation of the torque sensitivity with temperature is ignored. The errors increase up to ±4% FS at 160°C that is still unacceptable for the KERS application.

D. *Accounting for the Thermal Variation of the Sensitivity*

We do not have data for determining the torque sensitivity variation with temperature for each individual unit but we can use the generic thermal characteristic of the sensitivity instead. The question is whether the spread of the thermal variation of the torque sensitivity is small enough to achieve significant error reduction.

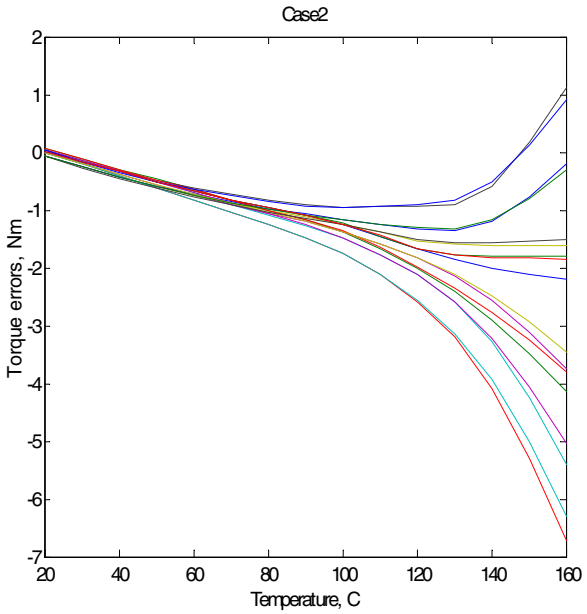


Figure 6. Torque measurement errors in the case B.

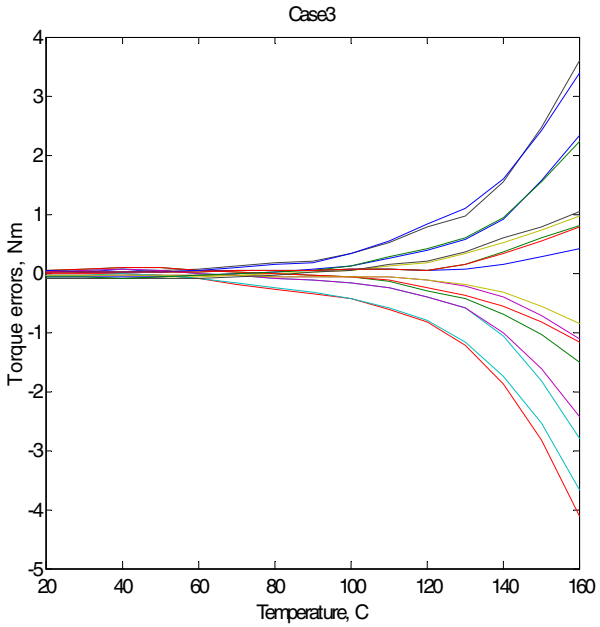


Figure 7. Torque measurement errors in the case C.

At the final stage of refinement of the calibration method it is proposed to calculate the reciprocal sensitivities according to the following rule:

$$C_{p,n}(T) = C_{p,n}(T_0)\frac{C'_{p,n}(T)}{C'_{p,n}(T_0)} \quad (7)$$

where $C_{p,n}(T_0)$ are found from the room temperature calibration and $C'_{p,n}(T)$ are taken from Table 2. In other words, the scaled generic thermal characteristics of the reciprocal sensitivities are used for each individual sensor. The rest of the calibration parameters are found the same way as in the Case C.

Fig. 8 shows the torque measurement errors obtained for the sensor No. 46 in this case. One can see that the errors are considerably smaller at higher temperatures than they were in the Case C. In fact they are almost the same as in the case of the full individual sensor calibration (compare with Fig. 5). Thus, at least for the given unit, the method described in the sub-section D gives an acceptable result in terms of accuracy.

IV. Discussion of the Proposed Calibration Method

To summarize the results presented in the previous section, the proposed method of calibration of the temperature compensated SAW torque sensors in the case of high volume production is as follows.

(1) At the development stage, a sufficiently representative number of the units (20 to 40) are fully calibrated in a number of temperature points (five to nine). Generic calibration parameters $C'_{p,n}(T)$ are found by averaging of the reciprocal sensitivities of the individual sensors.

(2) At the production stage, each individual sensor is calibrated only at room temperature and $C_{p,n}(T_0)$ are determined.

(3) Then the values of $F_m(0, T)$ and $F_t(0, T)$ are measured for a batch of sensors.

(4) The complete set of the raw calibration data is generated for each sensor using (5) and (6) where a_3 is found from (3) and $C_{p,n}(T)$ are found from (7). Then all the calibration parameters are found by the LMS method.

(5) Occasionally, a limited number of the calibrated units are tested over the entire temperature range in order to check stability of the manufacturing process.

It has been shown in Section III that the above method gives small additional calibration errors in the case of one unit, sensor No. 46. It needs to be checked for other units of the same type. Fig. 9 shows the maximum torque measurement errors obtained for 17 units not included in the batch of 45 units used to find the generic calibration parameters. The errors are presented for two temperatures, 125°C and 160°C, and four different calibration methods corresponding to the Cases A-D. One can see that the proposed calibration method (Case D) gives very consistent results for all sensors in the group and the errors, especially at T ≤ 125°C, are very close to those achieved in the case of full calibration (Case A). If the variation of the offset F_{m0} and the sensitivity with temperature is ignored (Case B) then the errors can be as high as 7.5% FS at 160°C.

Similarly, one can plot the maximum temperature measurement errors (see Fig. 10). Again, the proposed calibration method gives consistent results for all sensors with maximum errors not exceeding 2.5°C at T ≤ 125°C which is sufficient for an accurate temperature compensation of the torque readings.

The method under consideration can be applied not just to the KERS torque sensors but also to other types of SAW torque sensors such as flexplate transducers for measuring the engine output torque, driveshaft torque sensors and sensors

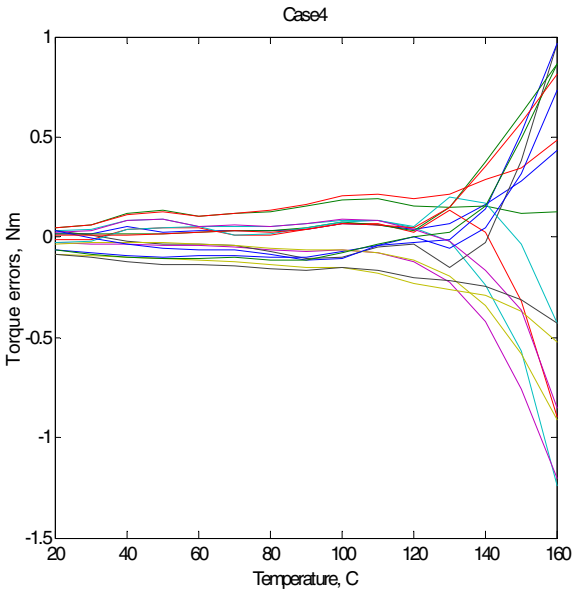


Figure 8. Torque measurement errors in the case D, the final version of the calibration method.

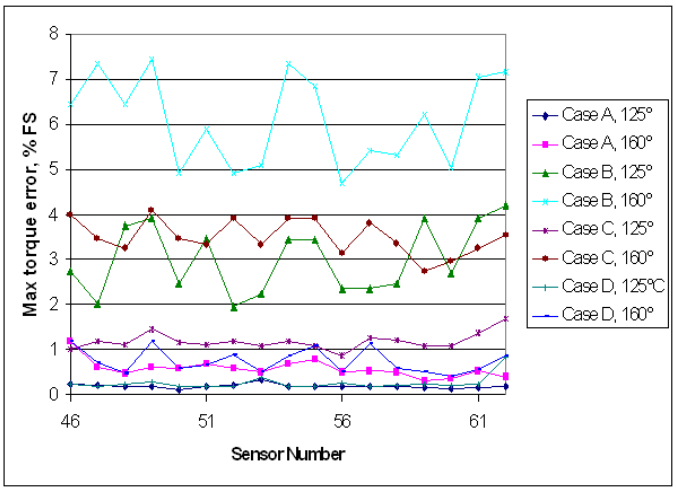


Figure 9. Torque measurement errors for 17 units at different temperatures and for different calibration methods.

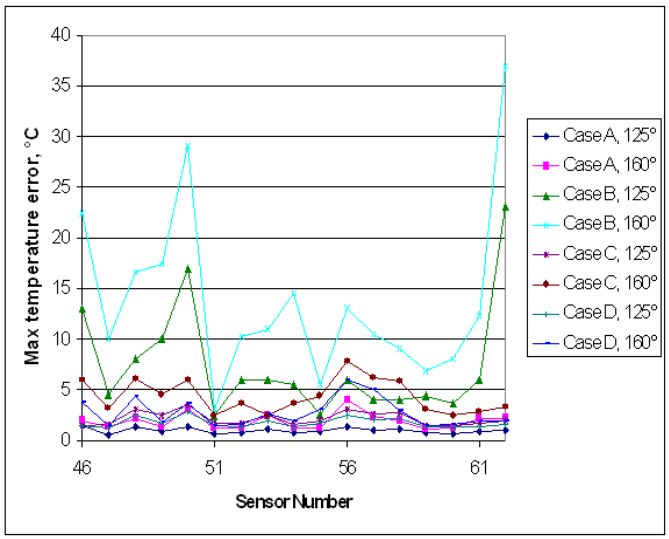


Figure 10. Temperature measurement errors for 17 units at different temperatures and for different calibration methods.

for electric power assisted steering. One should bear in mind that most of the applications do not require operation at temperatures above 125°C and, depending on the specified accuracy, the method described in the Case C may also be suitable. Operation at the negative temperatures should not create any problems because the glue bond stiffness does not drop and the sensitivity variation with temperature is minimal at $-40°C < T < 0°C$.

It is quite possible that a similar method can also be developed for other types of temperature compensated physical SAW sensors and even sensors based on other types of sensing elements. However, a careful analysis of the spread of calibration characteristics and the errors caused by their variation from unit to unit has to be performed in each case.

V. Conclusions

A calibration method for high volume manufacturing of the temperature-compensated resonant SAW torque sensors has been proposed requiring individual calibration of each unit only at room temperature. It is shown that the total torque measurement errors can be as high as 7.5% FS at 160°C and 4.2% FS at 125°C due to the spread of the temperature characteristics from one unit to another one if the individual temperature characteristics of the sensors are ignored. However, if the variation of the difference frequency offsets with temperature is measured on the individual basis then the torque measurement errors can be reduced to 4% FS at 160°C and 1.7% FS at 125°C. The best results are achieved if the thermal variation of the torque sensitivity is also taken into account on the basis of the generic calibration data obtained for a limited number of units at the development stage. In this case, the torque measurement errors can be as low as 1.2% FS at 160°C and 0.4% FS at 125°C. This is very close to the errors that can be achieved in the case of full calibration of each unit in a number of temperature points.

Acknowledgment

The author is grateful to A. Leigh of Transense Technologies plc for performing calibration of the SAW sensors and providing the calibration data.

References

[1] M. Hamsch, R. Hoffmann, W. Buff, M. Binhack, S. Klett, "An interrogation unit for passive wireless SAW sensors based on Fourier transform", IEEE Trans. on Ultrason., Ferroelectrics and Freq. Control, vol. 51, No. 11, pp. 1449-1455, 2004.

[2] http://www.sengenuity.com/prods_wireless_temp.html.

[3] B. Dixon, V. Kalinin, J. Beckley, R. Lohr, "A second generation in-car tire pressure monitoring system based on wireless passive SAW sensors", Proc. 2006 IEEE Int. Freq. Control Symp., p. 374 – 380, 2006.

[4] J. Beckley, V. Kalinin, M. Lee, K. Voliansky, "Non-contact torque sensors based on SAW resonators", Proc. 2002 IEEE Int. Freq. Control Symp., pp. 202-213, May 2002.

[5] V. Kalinin, G. Bown, A. Leigh, "Contactless Torque and Temperature Sensor Based on SAW Resonators", Proc. 2006 IEEE Int. Ultrason. Symp., p. 1490-1493, 2006.

[6] http://www.sensors.co.uk/technology/pages/torqsense.html

[7] V. Kalinin, G. Bown, J. Beckley and R. Lohr, "Pulsed interrogation of the SAW torque sensor for electrical power assisted steering", Proc. IEEE Int. Ultrason., Ferroelectrics, and Freq. Control Joint 50th Anniversary Conf., pp. 1577-1580, 23-27 August 2004.

[8] V. Kalinin, R. Lohr, A. Leigh, G. Bown, "Application of Passive SAW Resonant Sensors to Contactless Measurement of the Output Engine Torque in Passenger Cars", Proc. 2007 IEEE Int. Freq. Control Symp., pp. 499-504, 29 May - 1 June 2007.

[9] V. Kalinin, R. Lohr, A. Leigh, J. Beckey, G. Bown, "High-speed high dynamic range resonant SAW torque sensor for kinetic energy recovery system", Proc. 24th European Freq. and Time Forum, 13-16 April, 2010.

[10] V. Kalinin, R. Lohr, A. Leigh, "Development of a calibration procedure for contactless torque and temperature sensors based on SAW resonators", Proc.2008 IEEE Int. Ultrason. Symp.,

High Sensitivity MEMS Capacitive Tilt Sensor Using Multi-resonant Cylindrical Waveguide Structure

Panwar B. S.
Centre for Applied Research in Electronics, Indian Institute of Technology Delhi, India
E-mail: bspanwar@care.iitd.ernet.in

Siddharth Panwar
Stanford University, Department of Electrical Engineering, California, USA
E-mail: panwar@stanford.edu

Abstract — The present work describes a high sensitivity multi-anti-resonance CMOS process compatible MEMS liquid dielectric capacitive tilt sensor. The dual frequency excitation of four electrode structure enables anti-resonance at equally spaced multiple frequencies and capability of measuring the tilt with enhanced sensitivity along with the direction of tilt.

I. INTRODUCTION

In the past twenty years MEMS technology has matured to provide 3-dimensional structures with multifunctional capabilities obtained through the interface with microelectronic circuits [1, 2]. A noticeable turning point from sensor research toward MEMS research was demonstrated in 1988 though the development of micro-machined movable parts [3], gears and turbines [4] on silicon chip. Since then, the micro-machined devices and system are being identified as the area of thrust for their capabilities of providing miniaturization, multiplicity, and micro-system integration [5, 6]. In the initial stage, MEMS research was directed towards the development of silicon based accelerometers where the change in capacitance caused by the deflection of proof mass is used to measure the angle of inclination by measuring the acceleration due to force of gravity [7]. A novel structure without using the solid proof mass for acceleration and tilt sensing was proposed by Dao et al. [8]. These devices operate on the principle of effect of acceleration on the natural heat convection from heated resistive wires in a gas surrounding the device. A convection based inclinometer with 0.05 degree resolution, an angle range of ±30 degree, and 45 mW of power consumption has been demonstrated by Billat et al. [9]. Leung et al. [10] proposed the first micro-machined accelerometer based on free convection heat transfer of a tiny air bubble in a closed chamber. This was followed by the development of electrolytic tilt sensors [11-13], which are based on the principle of measuring the unbalance voltage in Wheatstone bridge due to the tilt of the sensor. However, these sensors suffer from the draw-back of low sensitivity, and need compensation for hydrophilic nature of oxide layer to reduce the surface tension and optimizing the frequency of operation to avoid the electrolysis process. Identically the sensors based on the concept of convection of heat transfer [7-9] have poor sensitivity and consume more power. In 2006 a MEMS capacitive bending sensor was proposed by Aebersold et al. [14] for monitoring the physiological function (intravascular pressure, intraocular pressure).

The applications of sensors as a medical device, which could be implantable or attached to a body worn instrument, require that the sensor be light weight having smaller power budgets in the form of a single chip. The interrogation and sensing of real time accurate data can be used for therapy for correcting the vestibular balance disorder. An ultra-low-power MEMS tilt sensor using hybrid two chips (MEMS microstructure/CMOS vision chip) solution has been implemented by Constandinou et al [15] for use in health care. The objective of the present study is a step towards developing a new liquid dielectric capacitive tilt sensor which is small, rugged, has low power requirement, and provides very high sensitivity and does not have any moving part. The development of such a device is being aimed for health care for monitoring the movement of a patient in post-operative state and incorporating the sense system for implantable prosthetics and body worn instrument.

II PRINCIPLE OF OPERATION AND TECHNOLOGY

The challenge in the design and development of MEMS based sensor technology is driven by the medical science, where monitoring of physiological functions plays a vital role in the recovery process. An important application is vitro-retinal surgery, where the eye is temporally filled by SF6 and its absorption requires that the head and eye are kept in a position which allows proper recovery [16]. Other very important application is monitoring strain of the instrument in

978-1-61284-110-6/11/$26.00 ©2011 IEEE

the spinal fusion process by an orthopedic surgeon [14, 17]. Therefore, the need arises for the development of light weight, low power, highly sensitive tilt sensors, which could work in a noisy environment to provide safe and secure health care aid to patients in the recovery process. This can be accomplished using the non-power dissipative capacitive elements in the senor and using an innovative dual frequency sensing scheme for reducing the noise. In this process the temperature dependence of angle of tilt, viscosity effects of the liquid used and problems caused by electrolysis process in the electrolytic tilt sensor are eliminated. This makes the obvious choice for a capacitive accelerometer / tilt sensor, where the technology and sensing scheme is to be carefully designed.

A. Dual Frequency Operation of Tilt Sensors

It is know that noise in slurry or low conductivity fluid measurement is a function of 1/f noise, therefore lower the frequency greater is the noise. In the conventional Wheatstone bridge measurement technique of convection or electrolytic tilt sensors, the unbalanced voltage caused by the tilt is sensed across the detector. The obvious limitation of such a technique is its inability to provide the information on the axis of tilt. Therefore, the need arises to develop a sensing system which can provide additional information on axis of tilt while using the advantages of dual frequency excitation. This approach is based on using four electrodes which are simultaneously excited by two frequencies which are twice of the other. The four electrodes a, b, c, and d are arranged in such a way that two orthogonal axis are along the diagonals of a→c, and b→d as shown in Fig 1.

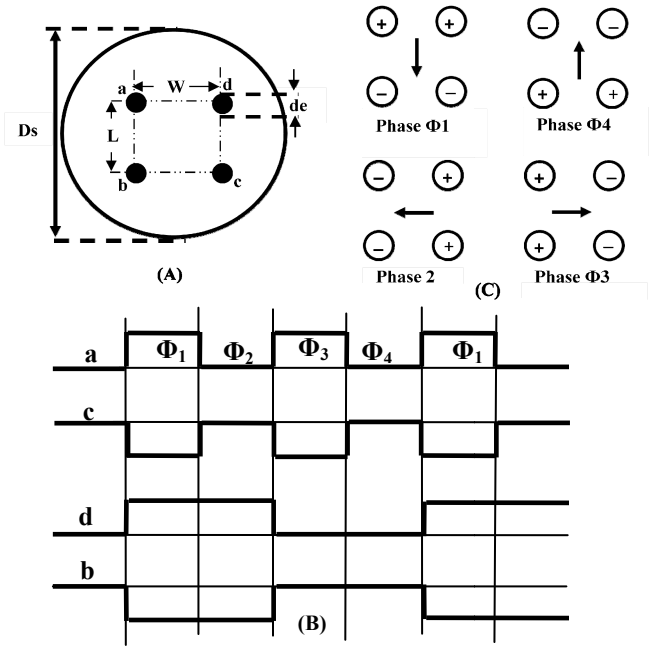


Fig. 1(A) Top view of the capacitive tilt sensor (B) Phase sequenced waveform applied to four electrodes a, b, c, and d (C) Direction of electric field between two electrodes

The frequency of excitation is selected on the basis to prohibit the electrolysis process and reducing the 1/f noise. The top view of the four electrode capacitive tilt sensor is shown in Fig 1(A) and the waveforms applied to the electrodes for dual axis measurement scheme is shown in Fig.1 (B), and the potential difference between four electrode structure used for measuring the tilt is show in Fig 1(C). The potential difference between two electrodes on account of phase sequencing, and choosing one of the frequencies to be twice of other frequency shows that dual axis information can be obtained by such a scheme. This information is contained in the direction of current flow between two posts defining the individual capacitive elements submerged in the liquid dielectric. The potential difference between the electrodes a→b (V_{ab}) and d→c (V_{dc}) or the current flowing between these two set of terminals (I_{ab} or I_{dc}) can be arbitrary chosen to represent –Y axis. Similarly, the voltage difference / current flowing between capacitive sensing electrodes d→a and c→b can be chosen to represent the – X axis. Since, the direction of current flow is in reference with the physical location of the electrodes, therefore these currents starts identifying the direction of tilt. These axes are defined arbitrarily and can be selected on the basis of mount / fixing of the sensor and the electrode structure. In the stable and un-tilted condition the differential current / signal can be sensed using a differential amplifier and the initial adjustment / calibration can be performed. Any tilt along the ±X axis and ± Y axis, synchronized with the clock generated from the exciting signal can be used to identify the axis and measure of tilt along the respective axis.

B. Fabrication Process

The starting wafer is both side polished, <100> n-type, silicon – on – insulator (SOI) wafer. Initially, 1 µm thick oxide layer is grown using wet oxidation process, thereafter 300nm thick silicon nitride layer using low pressure chemical vapor deposition (PCVD) is deposited. Thereafter, photo-resist is coated on the front side of the wafer and mask #1 is used to define the window in nitride and oxide layers. This is followed by etching silicon from the front side of the wafer for a depth of 25µm using Deep Reactive Ion Etching System (DRIE). This is required to have the hermetic sealing on the front side. Remove oxide and nitride layer from both sides of the wafer. Deposit both the sides of the wafer with fresh oxide layer (approx.1 µm thick by wet oxidation), and 300 nm silicon nitride layer by LPCVD technique. Use mask # 2 for opening the windows on the back side of the wafer for etching silicon and the buried oxide layer for taking contact from the back side of the wafer. Etch silicon and buried oxide layer from the back side of the wafer using Deep Reactive ion Etching. Deposit oxide layer on both the sides of the wafer by decomposing tetra ethyoxy silane (TEOS) in oxygen. This oxide on the front side gets deposited on top of oxide and nitride layers deposited earlier. In this process the entire front side of the wafer will get covered by additional oxide layer and groves etched earlier on the back side will also get covered by silicon dioxide layer.

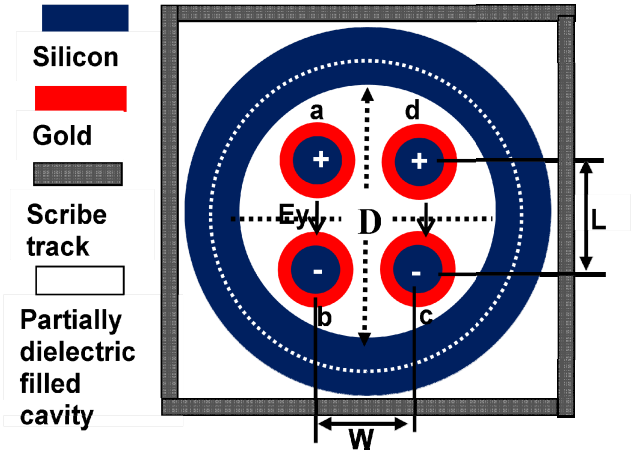


Fig.2. Die of a capacitive tilt sensing structure showing excitation of four electrodes for one of the phase sequenced state

Now use ion milling technique for removing the oxide layer from the front side of the wafer and the bottom of the groove formed by reactive ion etching. After this perform n+ diffusion for making the contact from the back side of the wafer. At this stage the front side of the wafer has 1µm thick oxide layer and 300 nm thick silicon nitride layers, and on top of this we have oxide layer deposited using the LPCVD technique by decomposing TEOS. Coat photo-resist on the front side of the wafer and define pattern in silicon nitride layer for etching deep well for defining the four electrode capacitive structure. This is followed by depositing Cr/Au by sputtering or ALD process. This deposition process is carried out on both sides of the wafer. In this deposition process the entire area of front side along with the electrodes gets covered by Cr/Au layer. This completes the process for taking the contact from the back side of the wafer from four electrodes. On the front side of the wafer nitride layer is removed, while retaining nitride layer on the back side. This is used to open four windows to take contact from four electrodes, which have been defined on the front side. This is followed by etching of 50 µm silicon in KOH using oxide as the mask. This is followed by using second mask on the front side to define the inner and outer rings of the four electrode capacitive sensor structure. A uniform etching of about 50µm of silicon in KOH is carried out. In this process the front side will get etched 50µm whereas the back side will get etched by 100 µm from the side to define the inner and outer rings of the capacitive sensor. This is followed by ion milling of frnt side using nitride as the mask for removing the Cr/Au deposited earlier, which was also existing on the bottom surface. In this process all the four electrodes gets isolated and the device is ready for bonding and packaging. A typical unit cell of the capacitive tilt sensor is shown in Fig. 3.

III MATHEMATICAL MODELLING OF STRUCTURE

An accurate analysis of the unit cell in Fig.3 has been carried out using the finite element tool of COMSOL Multiphysics Ver 4.1, which is elaborated in section IV. However, an insight on the functional behavior of the structure is obtained by considering a rectangular well which houses four stud like structure (electrodes). The inner wall of the well is at ground potential (electrical wall) and it is the partitioning wall between the electrodes that could be either an electrical wall or magnetic wall depending on the charge on the adjacent electrode, which is determined by the phase of excitation. The unit cell shown in Fig. 3 will always have the electrical wall at x = 0 and y = L/2, whereas y = 0 and x = W/2 can have the combination of either an electrical and magnetic wall or magnetic wall and electrical wall depending upon the phase sequence at a particular instant of time of the dual frequency excitation scheme adopted. The assumption of considering the rectangular wall is valid, because does not require any fabrication process change. The structure shown in Fig.3 is analyzed using clockwise rotation of standard Cartesian coordinate system. The structure is analyzed using Green's potential function for converting the differential equation to an integral equation for an unknown quantity of charge density. The analysis is carried out for a unit charge located x = x_0 and y = y_0, and one of three possible combinations of boundary conditions of electrical wall at y = 0 and magnetic wall at x = W/2 are chosen. For a unit charge located at (x_0, y_0) the Green's function, for a media having the dielectric constant ε satisfies the Poisson's equation in (x, y) plane.

$$\nabla^2 G_n (x , y \mid (x_0 , y_0) = - (1 / \varepsilon) \delta (x - x_0) \delta (y - y_0) \quad (1)$$

The Green's function can be expressed as the sum of the products of elementary function with separated variable [18]:

$$G_n (x, y) = \sum_{n=1}^{\infty} G_n (x) G_n (y) \quad (2)$$

Before solving (1) for potential function, we need to obtain the expression for $G_n(x)$ from the Dirichlet and/or Neumann boundary conditions which defines the unit cell. The Neumann type boundary condition at width x=W/2 forces the derivative of the potential function to be zero ($\partial G / \partial n = 0$), whereas the Dirichlet boundary condition at x=0 forces the potential function to be zero (G=0). The potential function $G_n^x(x)$ for electrical wall at x=0 and magnetic all at x = W/2 is expressed as:

$$G_n^x (x) = \sin \frac{(2n+1)\pi x}{w}$$

$$G_n^x (x) = \sin \beta_n \; x \; for \; n = 1, 2 \ldots . \infty$$

Where $\beta_n = (2n+1)\pi / W$

Since the electrode diameter is assumed to be much smaller than the electrode length, therefore the variation of field quantities along Z-axis can be neglected and potential function expressed as sum of the products of elementary functions with separated variables $G_n^x(x)$ and $G_n^y(y)$ using (2) can be expressed as:

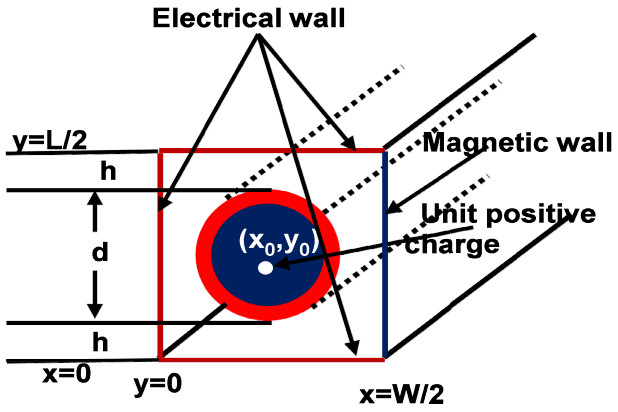


Fig.3. Unit cell of the capacitive sensor showing the formation of electrical and magnetic wall for a specific phase sequence.

$$\frac{d^2G_n(y)}{dy^2} - {\beta_n}^2\, G_n(y) = \frac{4}{\epsilon W} \sin\beta_n\, x_0\, \delta(y - y_0) \qquad (3)$$

The differential equation (3) is solved for the boundary conditions for the "i[th]" layer for the continuity of tangential component of potential function and continuity of the normal component of displacement density.

$$G(x, y_{i-o}) = G(x, y_{i+o})$$

$$\epsilon_i \frac{\partial G(x, y_{i-0})}{\partial y} = \epsilon_{i+1} \frac{\partial G(x, y_{i+0})}{\partial y}$$

After applying these boundary conditions the solution for the Green's potential functions $G(x_o, y \mid x_o, y_o)$ is:

$$G(x_o, y \mid x_o, y_o) = \sum_n \frac{4}{\beta_n W Y|_{y=0}} sin\beta_n x\, sin\beta_n x_o \quad (4)$$

This describes the Green's potential function in the (x, y) plane for a unit charge located at (xo, yo). The admittance $Y|_{yo} = 2\,\varepsilon_o \coth(\beta_n L/4)$ is obtained by using the analogy of permittivity with characteristic admittance of a transmission line. The variational approach adopted by Siddharth [19] for computing the capacitance of unit cell predicts dependence on the electrode width "d", for X-axis separation of "W", and Y-axis separation of "L". An easier expression for computing the capacitance per unit cell has been derived [19] for normalized electrode width "d", and length "L" w. r. t. width "W" of the sensor. The complete expression for the capacitance of a unit cell obtained by this technique is:

$$c = \sum_n \frac{1}{4} \frac{[\beta_n (d/2)]^2}{\cos^2[\beta_n (d/2)]} (2n+1)\,\pi\, Y|_{xo} \quad \text{F/meter}$$

Where $\beta_n = \frac{(2n+1)\pi}{W}$

$Y|_{yo}$ is input admittance of short circuited transmission defined by electrical walls located at L/4 from the unit charge, which is is expressed as: $Y|_{yo} = 2\,\epsilon_o \coth(\beta_n L/4)$.

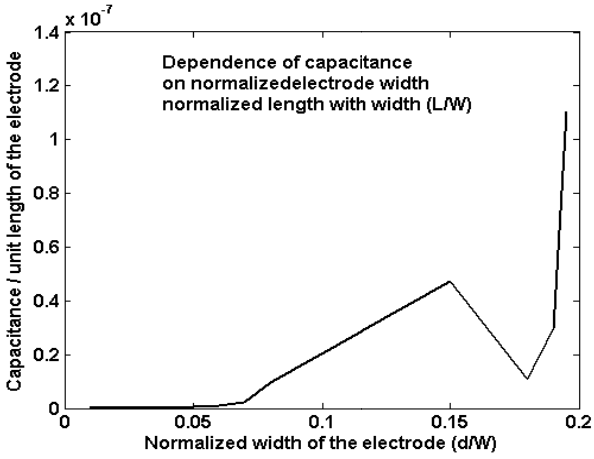


Fig.4. Dependence of the capacitance on the normalized electrode width

The results in Fig. 4 show that the maximum sensitivity is observed for d/W of 0.2. It is apparently the dimension for which the structure is likely to show resonant behavior. This implies that the capacitance has to increase, which will minimize the impedance for the dual frequency measurement scheme thereby maximizing the current. The resonant behavior can be fully exploited for getting very high sensitivity and the differential sensing to reject the common mode signal.

IV. FEM ANALYSIS OF THE STRUCTURE

A capacitive tilt sensor as shown in Fig. 1(A) consists of a hollow metallic cylinder terminated at both ends with insulating material. Four electrodes are placed symmetrically inside the cylinder such that they form four vertices of a square. The electrodes 'a' and 'c' are excited with same frequency 'f_1' but with a phase difference of 180 degrees. Similarly, the electrodes 'b' and'd' are excited with frequency 'f_2' and with a phase difference of 180 degrees. The frequencies f_1 and f_2 are chosen such that $f_2=2f_1$.

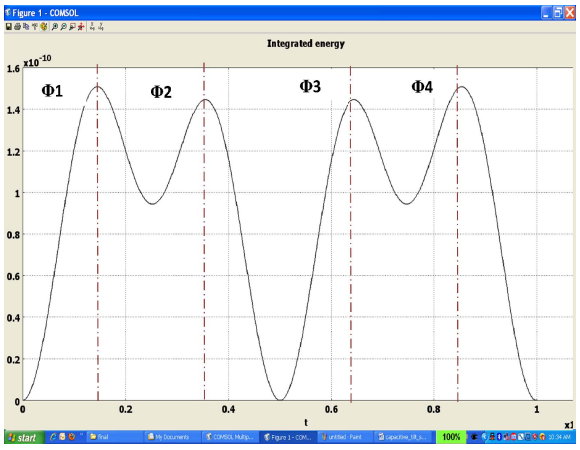


Fig.5. Energy stored in the cavity during different phases for w≠L

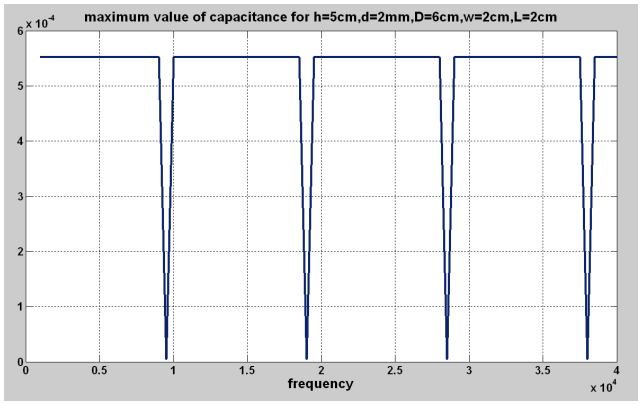


Fig.6. Cylindrical resonant wave guide resonant behavior of the tilt sensor

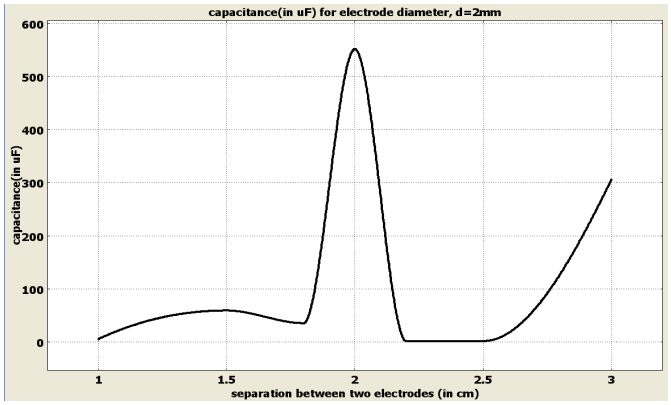


Fig.7. Dependence of resonant cavity capacitance on the electrode separation

The selection of $f_2=2f_1$ forms electric/magnetic walls between the electrodes which keeps changing according the phase sequencing at a particular instant of time. Fig. 1(B) shows the phase sequenced signal applied to four electrode structure and the direction of resultant electric field corresponding to each phase directing towards +X-axis,-X-axis, +Y-axis and –Y-axis is shown in Fig. 1(C). The analysis of the structure is carried out using the FEM Tools COMSOL Ver. 4.1 for analyzing the dependence of the total energy stored in the cavity for W=L and W ≠ L. Fig.5 shows the energy stored in the cavity for the case W≠ L. Since, the phase Φ_1 and Φ_4 have same amplitude of electric field strength therefore the total energy stored in the cavity will be equal for these two phases. A similar behavior is observed for phase Φ_2 and Φ_3, which have the same electric field strength. However the energy stored in phase Φ_1, Φ_4, is higher as compared to Φ_2 and Φ_3. Therefore the structure for maximum capacitance and resonance behavior is considered only for the case of W=L.

The formation of electric wall and magnetic wall by the phase sequencing scheme adopted and the cylindrical cross section of the structure establishes it resemblance with the cylindrical waveguide cavity. Therefore, it was expected to exhibit multi-resonance due the formation of electric and magnetic walls and equally spaced resonant frequencies on account of cylindrical cross-section. The simulation results shown in Fig.6 supplements the anticipated behavior of the structure, which predicts the presence of harmonics of 9 KHz, with a typical 3 dB bandwidth of 100Hz. As a consequence high Q of approx. 30 is obtained at the fundamental frequency, which will increase depending on the harmonic frequency selected for operating the device. Since, the excitation signals are 180 degrees out of phase, there will be anti-resonance, which will reduce the energy stored in the cavity at the resonant frequencies. This phenomenon will lead to almost zero energy stored in the capacitor at the resonance frequency, which corresponds to open circuit behavior of the cavity and a maximum value of $|S_{11}|$ = 0 dB, is expected at the anti-resonant frequencies. Since the energy stored in the cavity for each phase will be equal for the case of W=L.

The maximum value of capacitance is approximated to be 600 µF for the optimized parameters of the geometry of W=L = 2cm, for an electrode diameter of 2mm and internal diameter of the resonant cavity of 6 cm. This is the macro-model of the actual device to be implemented in silicon for which the simulation has been carried out. This macro model device will be tested for validating the concept before implementing the structure on silicon. The dependence of the capacitance on the separation between the electrodes is shown in Fig. 7 for the cavity height of 5cm. However, the behavior of the capacitance of the cavity will be independent of height of the cavity. The macro-model used in the evaluating the performance and the basic concept will be scaled down using the process steps outlined in section II for realizing the structure on the silicon wafer.

V. DEVICE CHARACTERIZATION RESULTS AND DISCUSSIONS

The theoretical analysis and FEM simulations are carried out for the clocking phase sequence φ_1, where electrodes “a”, “d” assumes the positive charge and electrodes “b”, “c” assumes negative charge. This defines an electrical wall between the electrodes “a” and “b” and electrodes “d” and “c” and magnetic wall between electrodes “a” and “d” and electrodes “b” and “c”. In this phase sequencing we will get the information on the tilt angle along the – y-axis. The measurements were performed on the macro device (scaled version of the device to be implemented on the silicon wafer) designed for the optimum performance for the electrode diameter of 2mm, inner diameter of the cylindrical wall of 6 cm, width and length defining the separation between two electrodes of 2cm. The analytical results in Fig. 4, showing the dependence of capacitance on normalized diameter of the electrode (d/W), predicts maximum sensitivity for d/W of 0.2. These results are supplemented by the FEM analysis carried out using the COMSOL Ver. 4.1, MEMS design tools, which show the maximum value of capacitance of 580 µF for d/W ratio of 0.2. This supplements the theoretical formulation to be consistent with FEM analysis.

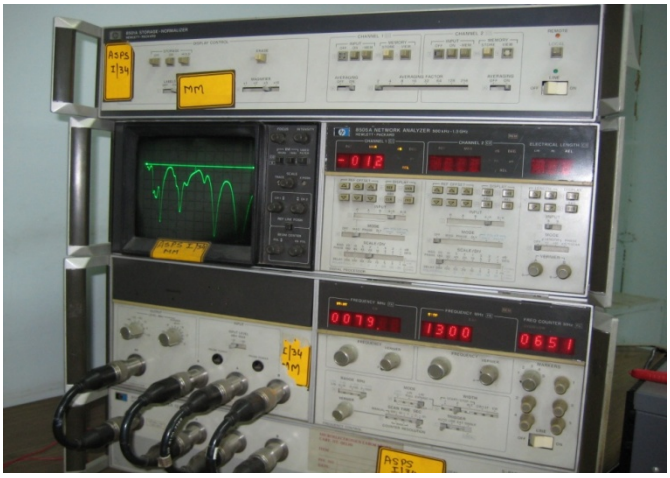


Fig.8: Measured |S11| parameter of the macro device showing formation of multi resonant cavity.

The macro structure is characterized using the 500Khz - 1300 MHz frequency range, HP8505A network analyzer by measuring the $|S_{11}|$ parameter. The measured results in Fig. 8 are consistent with the resonant behavior of forming the multi-resonant cavity predicted using the theoretical formulation and MEMS based simulation tools.

The sensitivity of the structure is primarily determined by the Q of the resonant cavity. An enlarged view of the resonant of |S11| plot shown in Fig.9 measured at the centre frequency of 1140 MHz is observed have 3 dB bandwidth of 20 KHz, predicting the resonant Q of 57, which is close to the value of Q=90 predicted by the FEM tools. The macro device was the first prototype developed for such as study and has considerable scope of improvement, which will bring closer agreement between the measured performance and the simulated results.

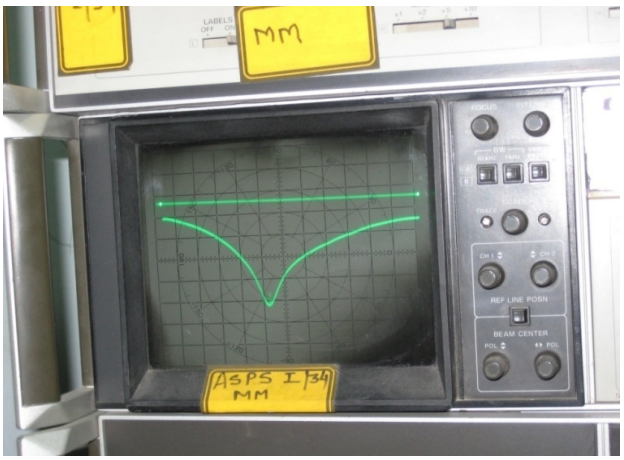


Fig 9. Enlarged peak of $|S_{11}|$ predicting the resonant Q of 57

ACKNOWLEDGMENT

The authors will like to thank Indo French Collaborative Program for Advanced Research (IFCPAR), India Habitat Centre, Lodhi Road, New-Delhi, India for providing the financial support to carry out the investigations reported in the paper. The authors are thankful to Prof. Boris Murmann, Department of Elect. Engg, Stanford University USA, for providing the opportunity to work on this problem.

REFERENCES

1. H. Fujita, "Two decades of MEMS – from surprise to enterprise", International Conference on Micro Electro Mechanical Systems, pp. 1-6, 2007.
2. H. Fujita, "A decade of MEMS and its future", IEEE Intl., Workshop, Nagoya, Japan, Jan. 26-30, pp. 1-8, 1997.
3. FAN Long-Sheng, TAI Yu-Ching, R.S. Muller, " Integrated movable micromechanical structures for sensors and actuators", IEEE Trans. on Electron Devices, Vol. 35, pp. 724-730, 1988.
4. M. Mehregany, K.J. Gabriel, W.S.N. Trimmer, "Integrated fabrication of polysilicon mechanisms", IEEE Trans. on Electron Devices, Vol. 35, pp. 719-723, 1988.
5. K. E. Petreson, "Silicon as a mechanical material", Proc. IEEE, Vol. 70, pp. 420 – 427, 1982.
6. K.J. Gabriel, "Engineering microscopic machines", Sci., Amer., Vol.260, pp. 118-121, 1995.
7. D. Lapadatu, S. Habibi, B. Reppen, G. Salomonsen, and Kvisteroy, "Dual axis capacitive inclinometer / low-g accelerometer for automotive applications", Proc. IEEE International Conference on MEMS Switzerland, Jan 21-25, pp. 34-37,2001
8. D. Rao, D. E. Morgan, H. H. Kries, D.M. Bacelder, " Convective accelerometer and inclinometer", United States Patent #5,581,034, REMEC INC. San Diego, CA, 1996
9. S. Billat, et al., "Convection – based micromachined inclinometer using SOI technology", Proc. IEEE Conference on Mirco Electro Mechanical Systems, pp. 159-161, 2001.
10. Albert M. Leung, John Jones, Eva Czyzewska, J. Chen, and M. Pascal, " Micromachined Accelerometer with no Proof of Mass", IEEE Electron Device Meeting, pp. 899-902, 1997.
11. Andrew Glascott Jones - http://www.sensorsmag.com/sensors/Automotive/MEMS-for-Tilt-Measurement-and-Headlight-leveling/Article/Standard/Article/ detail/557036
12. Ik-Su Kang, et al., " Design and fabrication of a MEMS based electrolytic tilt sensor", International Conference on Micro-processes and Nanotechnology, pp. 216-217,2005.
13. Ho Jung, Chang Jin Kim and Seong Ho Kong, "A MEMS-based electrolytic tilt sensor", IEEE conference on Sensors, pp-1199-1202, 2006.
14. J. Aebersold, et al., " Design, modeling, fabrication and testing of a MEMS capacitive bending strain sensor", International Conference on MEMS, pp. 124-129, 2006.
15. T.G. Constandinou, J. Georgiou, and C. Andreou, "An ultra low power micro-optoelectromechanical tilt sensor", IEEE International Symposium on Circuits and Systems, pp. 3158-3161, 2008.
16. J. Dlouthy, M. Cizek, I. Vicha, and J. Rozman," MEMS technology in head tilt monitoring after vitreoretinal surgery", International Conference Radioelektronika, pp.1-4, 2008.
17. L. Zhao, and E. Yeatmann, "Micro capacitive tilt sensor for human body movement detection", Pro. Body Sensor Networks, Vo.39, pp.195-200, 2007.
18. R. E. Colin, "Field Theory of Guided Waves", New York: Mcgraw-Hill,1960.
19. Siddharth Panwar, "Liquid dielectric tilt-sensor", Independent Study, Stanford University, CA, 2007

Experimental Study of Quartz Crystal Gravity Sensor by MEMS Technology

Takeru Mutoh, Mitsuaki Koyama, Takao Aizawa* and Toshifumi Matsuoka**
Nihon Dempa Kogyo Co., Ltd., Tokyo, Japan
* SUNCOH CONSULTANT Co., Ltd., Tokyo, Japan
** Kyoto University, Kyoto, Japan
E-mail: tmuto@ndk.com

Abstract— Quartz sensor is promising as a gravity sensor. This paper describes an improved technique of gravity sensor, based on the VCXO and crystal properties from the principle of the quartz motion sensor. The size of prototype sensor is around 27.0x12.0x1.0 mm. The sensor is used in the experiment, which give the gravity change. The experimental results show the relationship between gravity change and load resonant frequency. Then, from these results, we show that the load resonant frequency changes in direct proportion to gravity change.

I. INTRODUCTION

The Quartz crystal device is one of gravity sensors, which has high sensitivity. The crystal sensor, which is already realized, is pressure, power, gyros, chemistry, biochemistry, environment and etc.

When comparing with other sensors and such a crystal sensors, the crystal sensors are often responsible for the most sensitive range. Then, when assorting functions of the crystal sensors, It works as not only the physical sensors but also chemical ones.
Generally, the crystal blank for physical sensors is Z-cut. As for form, a Tuning-Fork type or a Double-Ended Tuning Fork type is used. The chemical sensor is the AT-cut, which surface is flat.

The AT-cut sensors are used for chemistry, biochemistry, biology and etc. For example, it is QCM (Quartz Crystal Micro-balance). The crystal blank for QCM has two pair of electrodes. Mr. Takeru Mutoh and Dr. Mitsuaki Koyama were authors who proposed the twin sensor. This sensor is one crystal blank and it has two pairs of electrodes [1].

On the other hand, for AT-cut quartz resonator, Mr. Mitoshi Umeki reported that the first application of MEMS technology. We believe in it [2].

Based on this background, the mineral exploration field, MEMS demanding the realization of the gravity sensor crystal technology is applied. Sensitivity of this specification is the gravity change of 0.5mgal, and this sensitivity is while moving the sensor.

Gravity is sensitive user specification that value is 0.5mgal, so to achieve such crystal sensor specifications; user requirement is at least 0.1mgal. To satisfy this requirement, there are four proposals as follow.

(1) Double-Ended Tuning Fork type sensor is one of the choices to achieve such high sensitivity sensors. Then we confirmed that this sensor was not sensitivity of 0.1mgal by simulation.
(2) By the structure of Double-Ended Tuning Fork type, it is difficult to realize purpose sensitivity of sensor alone. Therefore it comes true by the combination with a mechanical device. Some of these sensor configurations, the crystals may be realized. We do not expect the appearance of this crystal, which name is a twin. When presented with the temperature difference between the sensor and the pressure difference, this behavior occurs. Then, this sensor has uneasiness in reliability.
(3) Gravity changes the shape of the crystal blank is slightly deformed. The stress change accompanying this modification is measured as change of resonance frequency. This method has a too small change of resonance frequency. And this method has the influence of electronic noise. For this reason, predetermined sensitivity is not obtained.
(4) The capacitance is changed by motion of a crystal blank into crystal sensor. And, as the crystal sensor, the load resonance frequency change is measured. This method uses the principle of VCXO. And this sensor is a structure, which has a variable capacitor with a crystal oscillator. Furthermore, the crystal blank is changed easily, and is important to realize such a structure. This uses the principle of the motion sensor. This proposal is studied.

In this paper, the above proposals are shown in (4) is applied. MEMS (Micro Electro Mechanical Systems)

978-1-61284-110-6/11/$26.00 ©2011 IEEE

technology applied to the proposed AT-cut crystal structure of the gravity sensor, and the experimental results are shown.

II. TECHNICAL DESCRIPTION

A. Drawing of a crystal sensor

Figure 1 shows the concept drawing of the crystal sensor. This is cantilever of the single-sided support structure of AT-cut crystal blank with electrodes. This cantilever has a drive electrode and sensing electrode. Driving electrodes is interposed between the cantilevers. Sensing electrodes comprise a variable capacitor. Electrode A and B, it used as the sensing part, realize the variable capacitor. Electrode A for sensing is prepared in the end of this cantilever. Electrode B for sensing is formed also in the inside of the enclosure by the side of the contrary of the electrode. After this, Sensor name is Capacitive-type Quartz Sensor (CQS)

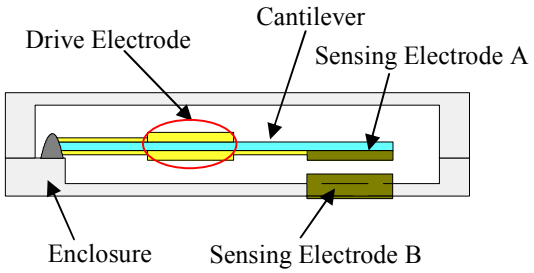


Figure 1. The basic structure drawing of a crystal sensor (cross section)

B. Equivalent circuit

Figure 2 shows the equivalent circuit, and is based on IEC 60122-1 [3].

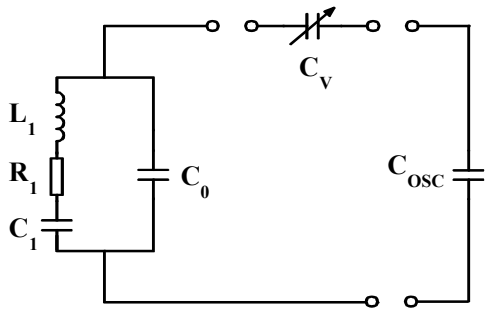


Figure 2. Equivalent Circuit

Figure 2 shows the symbol and symbols related to ; R_1, C_1,L_1 , and C_0: equivalent circuit constant of a crystal unit, C_V: capacitance ingredient of a gravity sensing part, C_{OSC}: capacitance ingredient of an activity circuit, Fr: the electric resonance frequency of a crystal unit, d: distance used as the standard which determines the capacitance of cantilever, δd : distance sensor is displaced by gravity change, S: effective electrode area of a sensor part, ε: dielectric constant. Then, C_L: capacitance of sensor area by the side of a crystal unit, and this combined capacity is the capacitance of the active circuit component C_{OSC} and the sensor capacitance component C_V.

Next, when F_L is oscillation frequency,

$$F_L = F_r \times \left\{1 + \frac{C_1}{2 \times (C_0 + C_L)}\right\} \quad (1)$$

where,

$$C_L = \frac{1}{\frac{1}{C_V} + \frac{1}{C_{OSC}}} \quad (2)$$

and

$$C_V = \frac{S \times \varepsilon}{d} \quad (3)$$

Using from equation (1) through equation (3), the cantilever is changed from d to d +δd, the difference between the load resonant frequencies is

$$\Delta f = F_r \times \frac{C_1}{2} \times (A - B) \quad (4)$$

where,

$$A = \frac{S \times \varepsilon + C_{OSC} \times (d + \delta d)}{C_0 \times C_{OSC} \times (d + \delta d) + S \times \varepsilon \times (C_0 + C_{OSC})} \quad (5)$$

$$B = \frac{S \times \varepsilon + C_{OSC} \times d}{C_0 \times C_{OSC} \times d + S \times \varepsilon \times (C_0 + C_{OSC})} \quad (6)$$

That is, the cantilever "d" from "d +δd" when a change to the difference between the load resonant frequency is larger. This is that sensitivity is high.

Then, figure3 shows in that the AT-cut crystal blank, as a motion sensor is not broken easily. The spring constant of AT- cut crystal blank is about 1800Ns/m. Therefore, no crack or/and no chipping AT-cut crystal blank is not easily broken by bend. This situation photograph shows figure 3, and, this crystal has realized by MEMS technology.

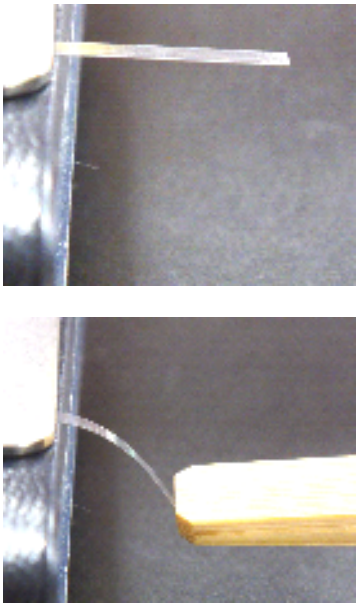


Figure 3. Bent test of AT-cut crystal blank
Crystal blank size : 20mm (L) , 1.6mm (W) , 0.02mm (H)

C. Specification and Trial

Specification of the crystal sensor is as follows.

Frequency	:	73 MHz
Mode	:	Fundamental vibration mode
Culture quartz	:	IEC60758 grade Aa
Surface	:	Polishing
Crystal blank dimensions	:	20.0×1.6 mm
Electrode	:	Cr, Au
Enclosure dimensions	:	27.0 × 12.0 × 1.0 mm

Based on such design specification, the crystal sensor is made as an experiment. Prototype sensor is shown in figure 4. Package materials for the prototype sensor are crystal. Crystal blank, package, and other wet-etching are applied, which technology is one of the MEMS.

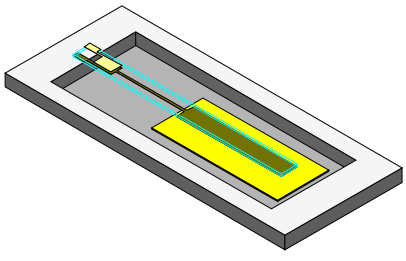


Figure 4. Prototype Sensor

D. Measurement system

Measurement system is constructed with sensor module, , Digital control equipment, and frequency counter. The sensor module is integrated with Crystal Sensor, Oven, and Oscillator Circuit. Figure 5 is the frequency temperature characteristics of the crystal sensor measured by IEC 60444-1 [4].

Temperature range is from +25 degrees C to +60 degrees C, which is a quadratic curve. Then, oven temperature is decided at +54 degrees C with the small amount of frequency change.

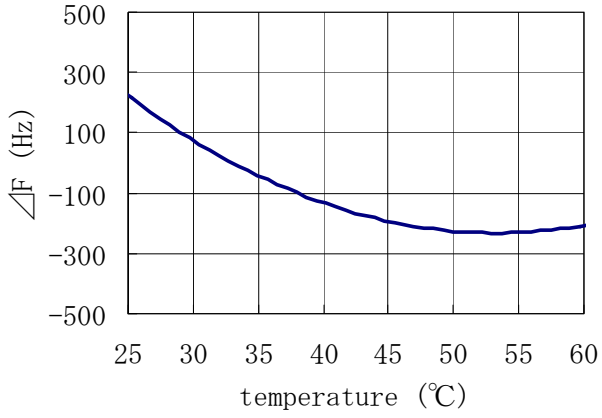


Figure 5. temperature characteristic result of the gravity sensor made as an experiment

Generally in measurement equipment, a cooling fan is used. Vibration of the cooling fan is from 0.05Hz to 5Hz, which is frequency noise of background vibration. Therefore, this cooling fan does not use it.

E. Relation of an inclination angle and gravity change

The structure of the gravity sensor has directivity in sensitivity. Therefore the sensitivity evaluation by the inclination angle is possible. The relation of an inclination angle and gravity change is shown in figure 6.

The gravity of the earth is 980gal. In the case of inclination angle θ=0 degrees, acceleration "a" applied to a sensitivity axis is

$$a = 980(\text{gal}) \times \cos\theta = 980(\text{gal}) \qquad (7)$$

On the other hand, in the case of inclination angle θ=Δθ degrees, acceleration "b" applied to a sensitivity axis is

$$b = 980 \times \cos\Delta\theta(\text{gal}) \qquad (8)$$

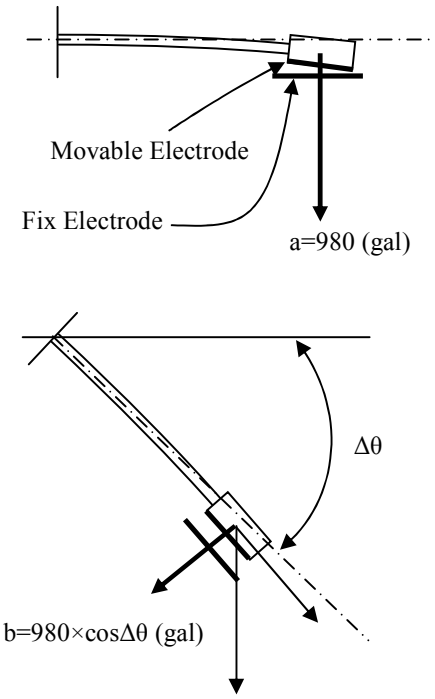


Figure 6. Gravity change by the inclination angle

III. EXPERIMENTAL RESULTS and DISCUSSIONS

A. Relation between gravity and frequency

Gravity change of 20mgal to 230gal is given to the sensor. The relation of gravity and load resonance frequency is shown in figure 7. Then, when gravity change becomes large, load resonant frequency is reduced. That is, load resonance frequency and gravity changes are involved.

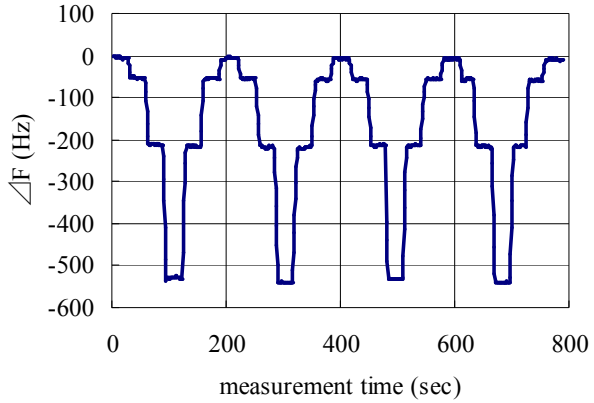


Figure 7. Gravity change and the relation of an amount of frequency variation

B. Relation of inclination angle and frequency

This is a definition, which load resonant frequency is equivalent to the oscillation frequency. The sensor is tilted at an angle from 0 degrees (980gal) to 30 degrees (849gal) to a change. This step change is 10 degrees. The oscillation frequency is measured continuously. Figure 8 is the result.
In Figure 8, the oscillation frequency changes over a maximum of 500Hz. Furthermore, it is reproducible.

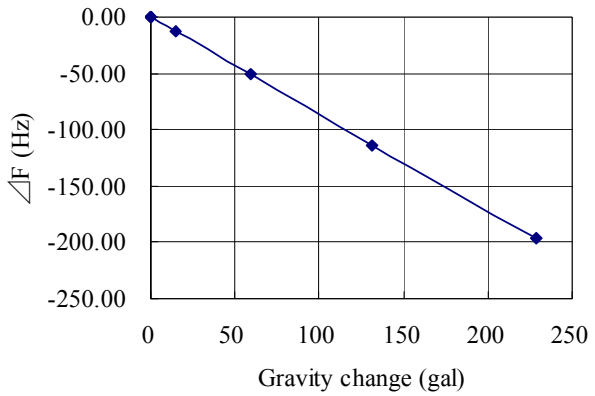


Figure 8.Oscillatory frequency deviation vs. Inclination angle

Gravity changes are given 20mgal sensors. Figure 9 is a waveform measurement. In figure 9, the measurement time 00:52-01:18 and 00:09-00:26 the frequency difference is the same -0.06Hz.

On the other hand, the change in measurement time is 00:52 and 00:09; the frequency drift is 0.02Hz. One of the factors of frequency drift is the temperature. This is not examined.

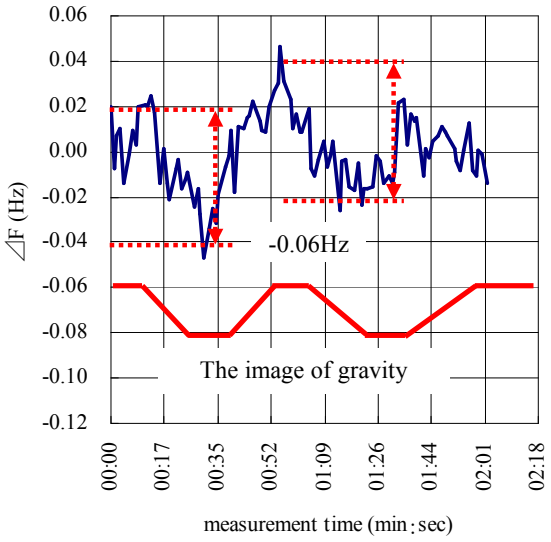


Figure 9. Frequency deviation vs. Time

IV. CONCLUSION

In this paper, AT-cut crystal blank is cantilever. The drive electrode and the sensing electrode are prepared on this cantilever. And the crystal gravity sensor, which changes capacitance, is proposed. Gravity sensitivity 20mgal is detected in the result of the experiment.

In the future, 0.1mgal is a realization of gravity sensitivity. Therefore, the crystal sensor structure and frequency drift is a challenge to improve. And we will study sensor structure and the effective Q value of the oscillation system.

ACKNOWLEDGMENT

The present work is supported by the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study is "Research and development of the moved type gravity investigation system using a small highly efficient gravity array" and is one of the "Basic tool development for seabed hydrothermal deposit" programs.

REFERENCES

[1]T. Muto, S. Watanabe, S. Wakamatsu, and M. Koyama "Examination for Realization of a High Precision Crystal Sensor" IFCS 2008 pp.532-534

[2]Mitoshi Umeki et al.,"622 MHz High Frequency Fundamental Composite Crystal Resonator With an Air-Gapped Electrode" Proceedings of the 2004 IEEE Inter national Frequency Control Symposium and Exposition

[3]IEC 60122-1: Quartz crystal units of assessed quality - Part 1: Generic specification.

[4]IEC 60444-1: Measurement of quartz crystal units parameters by zero phase technique in a π-network - Part 1: Basic method for the measurement

[5]Yasuo Toyosawa, et al, "Specific Research Reports of the National Institute of Occupational Safety and Health, " JNIOSH-SRR- NO.35 (2007)

SAW-Thin-Film Acoustoelectric In-Situ Observation and Measurement

B.H. Fisher and D.C. Malocha
Department of Electrical Engineering and Computer Science
University of Central Florida, Orlando, FL. 32816-2450
brian.fisher@knights.ucf.edu, donald.malocha@ucf.edu

***Abstract*—The thin-film acoustoelectric effect in surface acoustic wave (SAW) devices describes the interaction of electrical energy between a SAW in a piezoelectric medium and a thin-film placed in the wave's propagation path. The real-time observation of the thin-film acoustoelectric interaction is useful in the design and characterization of SAW-based thin-film chemical and physical sensors (i.e. temperature, humidity, viscosity, voltage, current, hall effects, etc.). An in-situ test fixture was designed to be mechanically, thermally and electrically stable. Data has been taken for many SAW devices and over a wide range of frequencies. The results show that the use of the in-situ procedure yielded: good agreement between theoretical predictions and the measured data, characterization a SAW-H_2 gas sensor in real-time and various methods to calibrate the film deposition system and procedure. This paper presents the approach taken in configuring an electron beam evaporation system for ultra-thin-film characterization and the design of test fixtures, data acquisition configuration, and experimental procedures to extract and analyze SAW parameters in real time, and to extract the thin-film properties under test.**

I. INTRODUCTION

The interaction between a surface acoustic wave (SAW) traveling in piezoelectric and a thin-film in the acoustic path may cause a measureable change in the SAW's amplitude and velocity. In the acoustoelectric regime, the magnitude of the change is dependent upon the electrical properties of the film. In-turn, the film's electrical properties may be dependent on a variety of physical parameters such as thickness, temperature, pressure, gas/chemical presence, etc. Hence, if properly implemented the SAW-thin-film acoustoelectric interaction is useful in the analysis of the properties of thin-films as well as thin-film physical sensors [1, 2]. Particularly, this work is a portion of a continuing effort to develop a passive wireless orthogonal frequency coded (OFC) SAW RFID-tag hydrogen (H_2) gas sensor [3-5], a schematic of which is presented in Fig. 1.

Real-time, in-situ observation of the acoustoelectric effect has been reported by previous authors [2]. However, they recommended that measurement of the entire frequency response of the SAW device should be avoided, because the time required for the sweep would be in the order of one to several minutes. Consequently, these authors [2] used SAW delay line sensors in an oscillator configuration. This type of system prevents the simultaneous measurement of multiple harmonics of the SAW device which is advantageous when attempting to characterize the dispersion characteristics of thin-films. Furthermore, the authors [2] omitted the experimental detail for duplicating or enhancing a similar system. The purpose of this paper is to present the approach taken using an electron-beam (e-beam) film evaporation system, test fixture and software, and the experimental procedure used to observe and measure the SAW-thin-film acoustoelectric interaction in real-time during the deposition (in-situ) of a thin palladium (Pd) film. The test thin-film, palladium (Pd), was chosen for the design and characterization of SAW-based hydrogen (H_2) gas sensors. The SAW-Pd-H_2 gas interaction was also observed via with the use of the test fixture and data acquisition system.

The paper to be presented is structured as follows: section 2 discusses the acoustoelectric-effect theory, section 3 details the test fixture, software design, and experimental considerations, section 4 details the experimental procedure, section 5 is a discussion of results and finally section 6 concludes the report and comments on prospective research applications.

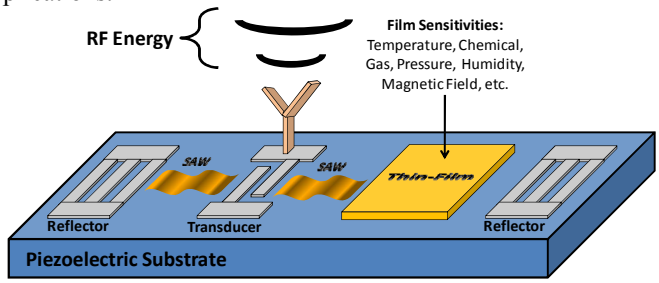


Figure 1 shows a schematic of a passive wireless SAW RFID-tag sensor. If properly designed the SAW propagation parameters are modulated by the presence and the properties of the thin-film.

II. THEORY

The piezoelectric substrate which was used in this work is Y-cut, Z-propagating lithium niobate (YZ-$LiNbO_3$), which was chosen because of its high electromechanical coupling coefficient, and self collimating properties. The acoustoelectric interaction occurs because a SAW traveling in a piezoelectric is accompanied by a traveling electric field (Fig. 2). The moving electric field creates only displacement currents in the piezoelectric (assuming that it is a lossless dielectric) and both conduction and displacement currents in the film (assuming that it is a semiconductor) (Fig. 2). The magnitude of the ratio of the conduction to displacement current is called the loss tangent of the interface. The acoustoelectric equations for SAW propagation in a piezoelectric medium and beneath a thin-film, predicts the wave propagation loss and velocity as a function of the loss tangent of the film-substrate interface.

If the thin-film is a metal, the conductivity is purely real consequently only conduction currents are generated in the

This work was partially supported through NASA-KSC STTR contract NNK04OA28C, in collaboration with Mnemonics Inc., the McKnight Doctoral Fellowship Program and the FHTIC Corridor Matching Research Program.

film and is given by $J_F = \sigma_F E$, where σ_F is the DC conductivity and E is the magnitude of the electric field. Assuming that the substrate is a lossless dielectric then current is a displacement current given by $J_S = j\omega\varepsilon_p E$, where $\omega=2\pi f$ is the angular frequency and ε_p is dielectric permittivity of the piezoelectric substrate. Hence, the loss tangent of the film-substrate interface is given by $\|J_F/J_S\|=\sigma_F/\omega\varepsilon_p$ and $\omega\varepsilon_p=\sigma_S$ may be considered substrate's reactive conductivity.

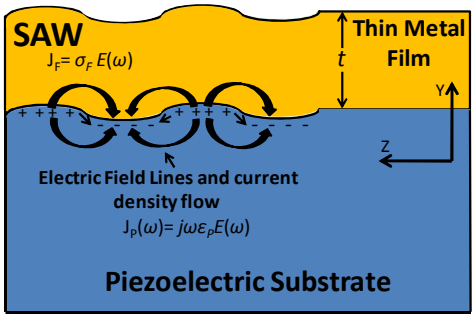


Figure 2 Schematic of SAW propagation in a piezoelectric crystal, beneath a thin-film. This is a cross-section of the film-substrate interface of the SAW device shown in Fig. 1. The subscripts, 'F' and 'P' represent the thin-film and the piezoelectric substrate respectively. σ is the DC electrical conductivity, ε electrical permittivity, ω is the angular frequency of the SAW, and J is the current density.

The acoustoelectric propagation loss, α, is given by (1) and phase velocity, v_p, is given by (2) as

$$\alpha = -\frac{\omega \cdot k^2}{2v_{sc}} \frac{\left\|\frac{\sigma_F}{\sigma_S}\right\|}{\left(1+\left(\left\|\frac{\sigma_F}{\sigma_S}\right\|\right)^2\right)} \quad (1)$$

$$v_p = v_{sc}\left(1+\frac{k^2}{2}\frac{1}{\left(1+\left(\left\|\frac{\sigma_F}{\sigma_S}\right\|\right)^2\right)}\right) \quad (2)$$

where, v_{sc} and k^2 is the short circuit velocity and electromechanical coupling coefficient of the piezoelectric substrate; for YZ-$LiNbO_3$, k^2=4.6%, v_{sc}=3408m/s, and $\varepsilon_P=51\varepsilon_0$. A full derivation of the acoustoelectric equations is provided by [2, 5-9], and will not be revisited here.

A strong interaction between the film and SAW occurs when the magnitude of the conduction to displacement currents at the film-substrate interface are equal (i.e. loss tangent = 1) (Fig. 3). This results in a large propagation loss in the SAW and a shift the velocity to half-way between the open and short-circuit velocities of the substrate. The open-circuit case occurs when the film's conductivity is very small in comparison to the substrate's conductivity hence open circuited and the short-circuit case occurs when the film conductivity is very high in comparison to the substrate's conductivity, consequently short-circuiting the electric field, E, and eliminating the piezoelectric effect at the surface of the piezoelectric.

The DC conductivity, σ_F of most films is dependent on the film's thickness, t, absolute temperature, T, and other physical parameters such as, gases, humidity, pressure, magnetic field, electrostatic potential etc. Thus, the properties of thin-films and their response to various stimuli may be studied through real-time observation of the acoustoelectric effect.

For this study the acoustoelectric interaction was observed by measuring the film's thickness and conductivity, and the SAW's propagation loss and velocity simultaneously. The open to short-circuit transition of the acoustoelectric effect may be observed by depositing a thin-film in the SAW's propagation path and measuring the aforementioned parameters during the film's growth. In metal films, the conductivity transitions from zero to high conductivity (i.e. open to short-circuit) as the film is deposited and the thickness, t, increases. In this experiment, Pd films were deposited in the propagation path on a z-propagating SAW in YZ-$LiNbO_3$ while the propagation loss and velocity were measured at the SAW's 1st and 2nd harmonics (54.5 and 109MHz) simultaneously. The design of the fixture is discussed in the next section.

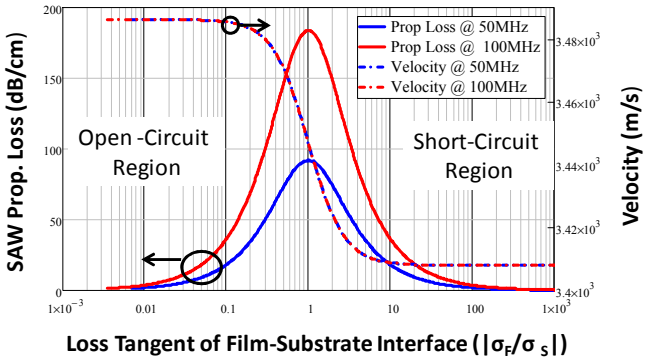


Figure 3 shows a plot of the SAW propagation loss (solid lines) and velocity (dashed lines) as a function of the loss tangent of the film-substrate interface.

III. SYSTEM DESIGN AND CONSIDERATIONS

A. Test Fixture

An in-situ fixture was designed to mask the SAW device so that the film is only deposited in the desired area (Fig. 4). The fixture was placed inside an electron-beam thin film deposition chamber, where RF cables where connected to fixture in order to interrogate the SAW, and measure film resistance (Fig. 5). The RF cables were connected to the measurement equipment via a high vacuum feedthrough which was constructed from an aluminum (Al) nut and bolt, semi-rigid coaxial transmission lines, and high-vacuum epoxy.

The test fixture was designed to be mechanically, thermally and electrically stable by using a (2cm x 7.5 cm x 4cm) brass base (Fig. 6). Sub-miniature version A (SMA) connectors were used to make connection between the RF cables and the probes. The RF probes were designed for stable electrical and mechanical contacts, which eliminates the need for bond wires and packaging. The probes were constructed using small hinges (for vertical motion), PCB, copper prongs (for the probes), and solder (connects PCB and probes) (Fig. 6). Threaded screws were used to apply a vertical force on the probes and hold SAW device die in place in order to

create good electrical and mechanical contact to the SAW device.

A picture of the YZ-$LiNbO_3$ SAW die which was used in these experiments is shown attached to the base in Fig. 6, a detailed schematic of the die is provided in Fig. 7. The schematic of the die shows a one-port differential delay device with large Al probes pads to facilitate easy contact and low electrical contact-resistance. The thin-film in the SAW propagation path makes contact with a set of Al bus-bars which are used to measure the film's sheet resistance, from which, the conductivity may be calculated using the thickness and area of the film. In a similar fashion two-port delay lines were also constructed. There are two RF probes on the fixture which allowed the measurement of a one-port device and film resistivity simultaneously or a two-port device. A secondary fixture was created which allowed both two-port measurement and film resistivity simultaneously.

The shadow mask for the fixture is shown in Fig. 8, and was also constructed from brass. For stability, the shadow mask was attached to the base with threaded screws. The length of the film in the propagation path was controlled using the width of the via-hole, which was controlled using a 56 thread-per-inch screw (i.e. 0.454mm lateral movement per revolution). An electrically controlled shutter (constructed using a DC motor and Al sheet metal) was positioned on the shadow mask to expose or block the evaporating material from the device. This function is important as it allows precise stop and start times for the film deposition on the substrate, and the possibility of deposition automation through computer controls.

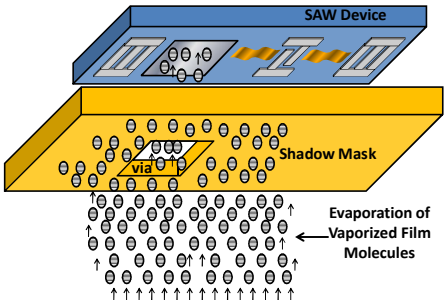


Figure 4 shows a schematic of the shadow masking of a SAW device, which allows deposition of the film precisely on the SAW propagation path.

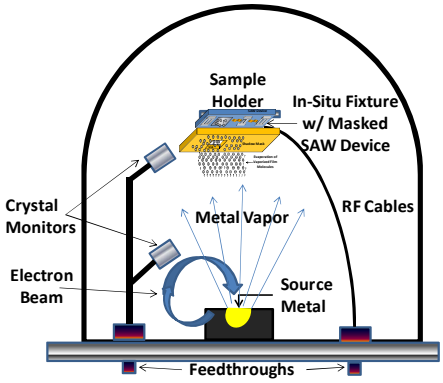


Figure 5 shows a schematic of the shadow mask inside the e-beam film deposition system, with dual quartz crystal monitors, and high-vacuum RF feedthroughs.

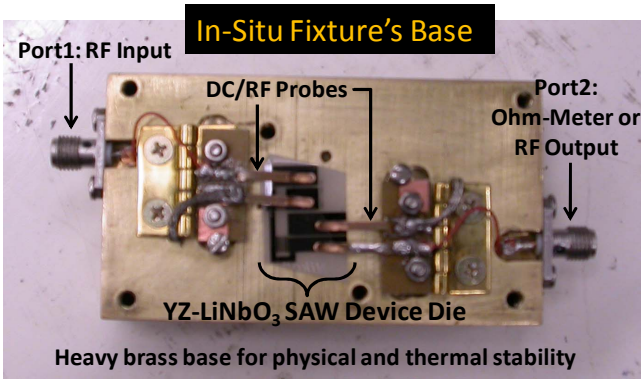


Figure 6 is a picture of the base of the test fixture. The fixture is designed to be mechanically, thermally and electrically stable by using a (2cm x 7.5 cm x 4cm) brass base. SMA connectors were used to make connection between the RF cables and the probes. The RF probes were designed and constructed for stable electrical and mechanical contacts, which eliminates the need for bond wires and packaging. The probes were constructed using small hinges (for vertical motion), PCB, copper prongs (for the probes), and solder (connects PCB and probes).

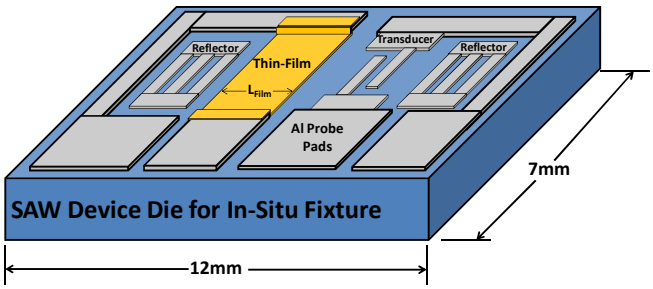


Figure 7 shows a schematic of the SAW die which was design for the in-situ fixture. It is a one-port differential delay device which large Al probes pads to facilitate easy contact and low electrical contact-resistance. The thin-film in the SAW propagation path makes contact with a set of Al bus-bars which are used to measure the film's sheet resistance, from which the conductivity may be calculated using the thickness and area. In a similar fashion two-port delay lines were also constructed.

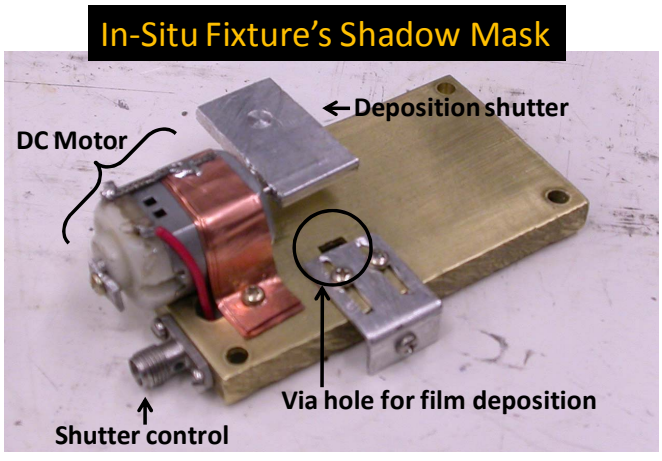


Figure 8 is a picture of the shadow mask for the test fixture. The length of the film in the propagation path was controlled using the width of the via hole, which was controlled using a 56 thread-per-inch screw (i.e. 0.454mm lateral movement per revolution). An electrically controlled shutter (constructed using a DC motor) was positioned on the shadow mask to expose or block the evaporating material from the device.

B. Software Design

The equipment used to acquire data are as follows: a vector network analyzer (VNA) to measure S-parameters, multi-meter (MM) measures film resistance, power supply (PS) which outputs constant voltages for IV sweeps and a quartz crystal monitor (QCM) to measure film thickness. The measurement equipment was controlled with a program created in Visual Basic and MATLAB. Microsoft's Visual Basic was used to control equipment and grab data because, Agilent, the manufacturer of the VNA, MM and PS supplied free sample code for each machine, which was combined and modified to make one central control program. Since Visual Basic is limited in its data analysis capability it was necessary to export the data to MATLAB where it was analyzed in real-time. In the MATLAB toolbox there is a program called "exlink", which allows MATLAB to be controlled via Visual Basic. This functionality was essential in order to achieve the real-time analysis capability.

C. Experiment Considerations

The goal is to measure SAW propagation loss and velocity, thus, it is important to maintain a signal to noise ratio (S/N) which is greater than 0dB. Since the maximum propagation loss of the SAW is controlled by the length of the film in the propagation path, it is necessary to design for a film length that will not bury the SAW response in noise. From (2), the maximum attenuation occurs when the magnitude of the displacement and conduction currents are equal, thus reducing (2) to

$$\alpha_{\max} = -\frac{\pi \cdot f \cdot k^2}{2v_{sc}} \tag{3}$$

The maximum attainable propagation loss in dB is given by

$$PL_{\max\_dB} = -20\log_{10}\left(\exp(\alpha_{\max} \cdot L_{film})\right) \tag{4}$$

where L_{film} is the length of the film in the propagation path. Fig. 9 shows a plot of the maximum theoretically attainable propagation loss as a function of the length the film in the SAW propagation path at 50, 100 and 1000MHz. For the experiments performed in this paper L_{film} = 1mm.

IV. EXPERIMENTATION

Prior to film growth, the lithium niobate (YZ-$LiNbO_3$) substrate was cleaned with an ultra-sonic bath in acetone, rinsed in methanol, then de-ionized water, and finally plasma irradiated with O_2 gas at 150W for a minute. Experimental observation has shown that it is important to keep a consistent cleaning and preparation procedure in order to decrease variability between experiments.

Films were deposited using electron beam evaporation, at a growth rate of 0.1Å/sec, in a bell jar which was evacuated to 10^{-6}Torr. The film thickness was monitored using dual quartz crystal monitors, and the film resistance was monitored during growth (in-situ). A picture of the equipment set-up is shown in Fig. 10; the VNA, MM, PS and control computer are all positioned beside the e-beam chamber.

The VNA was used to sweep a wide frequency range and capture data at multiple SAW harmonic operation points (54.4 and 109MHz). This function is useful when evaluating the dispersion characteristic of the thin film. The SAW propagation loss was extracted by measuring the ratio of reflected wave amplitude with the film to the reflected wave amplitude without film. The group delay was extracted by gating the time domain response and performing a FFT of the time gated response to obtain the frequency response. From which, the group delay is the derivative of the phase with respect to frequency *(-dθ/dω)*, from which, the velocity beneath the film was extracted since L_{film} is known.

A screen shot of the data output from the MATLAB software is shown in Fig. 11. The upper left plot is the time domain response of the SAW, which and transformed to the frequency domain in the upper right plot. This plot is then truncated to the 1dB bandwidth of the SAW response (middle left) from which the SAW propagation loss (middle right), center frequency (lower left) and group delay change (lower right) are extracted. The data was sampled at a rate of 3sec/sample, and at a growth rate of 0.1 Å/sec, resulted in a sample every 0.3 Å. A post processed plot of the SAW propagation loss and velocity as a function of Pd film thickness at 54.5 and 109MHz is given in Fig. 13.

A secondary experiment was performed in order to characterize the response of the SAW-Pd-H_2 gas interaction as a function of time. In this experiment the Pd film deposition was stopped near maximum SAW attenuation, removed from the high vacuum environment, and exposed to 2% H_2 gas (balance N_2) in multiple on/off cycles. The result of the SAW-Pd-H_2 gas interaction is presented in Fig. 14.

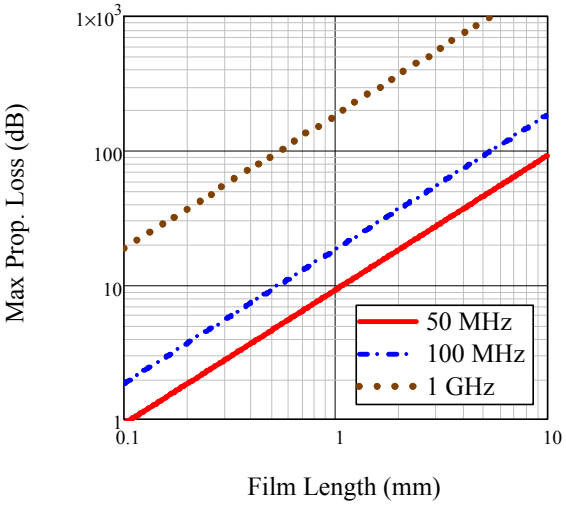


Figure 9 is a plot of the maximum attainable acoustoelectric propagation loss of a z-propagation SAW on YZ-$LiNbO_3$ as a function of the length of the film in the propagation path.

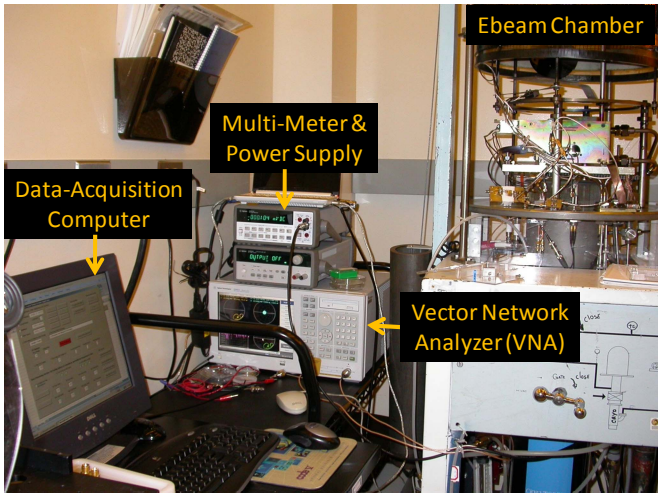


Figure 10 is a picture of the e-beam system, text fixture and data acquisition set-up in class 100 cleanroom.

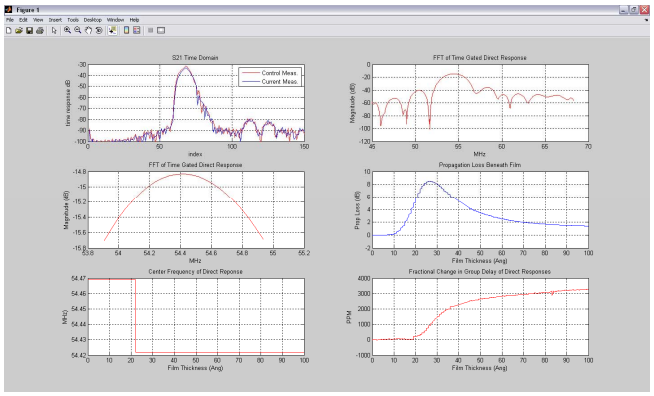


Figure 11 shows the real-time output of data acquisition software; measured S-parameters which are processed to extract propagation loss and delay shift during growth of a Pd film in the SAW propagation path. The plots are updated in real time as film thickness changes.

V. DISCUSSION

A comparison of the predictions (Fig. 3) and the measured data (Fig. 13) shows satisfactory agreement with some differences. One notable difference is the disparity between the predicted and measured maximum attenuation at 54.5 and 109MHz. From (4) the predicted maximum attenuation at 54.5 and 109MHz is approximately 100 and 200dB/cm respectively, whereas, the maximum measured propagation loss is approximately 85 and 165dB/cm respectively. This discrepancy is most likely due to the non-uniform distribution of the film. A light meter was used to infer the film's thickness distribution by measuring the fraction of light which was absorbed by the film as a function of position on the film. Ideally as the film gets thicker a greater percentage of light is absorbed. The film's thickness distribution was found to follow a Gaussian distribution (Fig 12). Another notable difference is between the predicted and measured short-circuit velocity. The short-circuit velocity is predicted to have limit of 3408m/s on YZ-$LiNbO_3$ however the measure velocity gradually decreases to 3320 m/s (Fig 13). The measured data includes both acoustoelectric and acoustoelastic interactions between the film and substrate, whereas, the predictions only account for the acoustoelectric mechanisms. Thus, the decrease of the measured SAW velocity below the theoretical limit is most likely due to acoustoelastic (i.e. in this case mass loading) effect. There may be also be acoustoelastic effects on the SAW propagation loss and other second order effects such as bulk or other generated modes which interfere constructively and de-constructively with the SAW amplitude and velocity at 1st and 2nd harmonic.

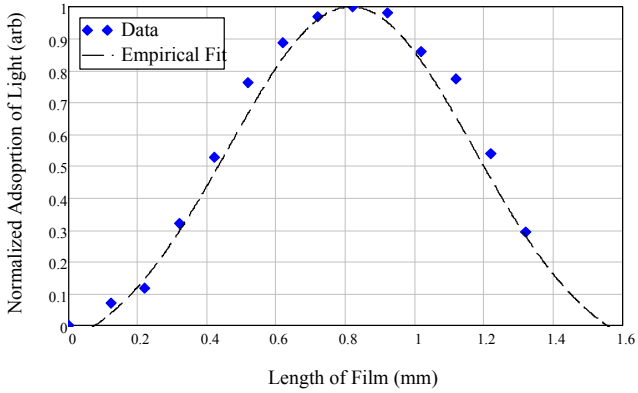


Figure 12 shows the thickness distribution of a 60Ang Pd film deposited through a 1mm aperture. The thickness distribution was extracted from light reflection measurements across the film.

The fixture and software was also used to characterize the SAW-Pd-H_2 gas interaction (Fig. 14). When exposed to 2% H_2 gas, there was a sharp decrease in film resistivity which caused an increase in the SAW propagation loss. The resistivity increased with each exposure to H_2 gas and was accompanied by a gradual increase in resistivity over time. If the increase in resistivity associated with H_2 gas exposure is excluded, there remains a gradual increase in resistance, caused by the aging of the ultra-thin Pd film in ambient atmosphere [10]. Thus, the SAW-Pd-H_2 gas sensor operation appears to be governed by the acoustoelectric effect. Thorough experiments into this interaction may lead to better understanding of the Pd-H_2 gas adsorption and desorption, and conduction mechanisms. Furthermore, this method may be used to evaluate the parameters of delicate ultra-thin films without need for contact from probes (which may damage samples and/or add experimental error).

These results suggest that this software and fixture is suitable tool for SAW sensor research where the knowledge of the real time response is necessary. Furthermore, this system has been used for other purposes such as determining the optimal settings for the VNA (i.e. adjust frequency span, IF BW, averaging, sweep time and/or output power until scatter is minimized). This is necessary when attempting to minimize measurement error. The system was also useful as high vacuum gauge since there is a measureable decrease in SAW propagation loss at high vacuum (10^{-6}Torr) as compared to 1atm [11, 12]. Lastly, the system was used to calibrate the e-beam film deposition procedure that is to develop a procedure to obtain reproducible results.

VI. CONCLUSIONS

The design of a test fixture and data acquisition software for the real-time measurement and analysis of the SAW-thin-film acoustoelectric interaction was demonstrated. The fixture allows measurement of SAW propagation loss, velocity and film resistivity simultaneously. The fixture uses RF and DC probes which eliminate the needs for bonds wires and packaging. Measured results show good agreement with the theoretical predictions. Research applications for the system include: SAW substrate parameter extractions, thin film material characterization, thin film physical and gas sensors, and deposition procedure developement.

ACKNOWLEDGEMENTS

The authors are most grateful to Dr. Robert Youngquist, NASA-KSC and Luis Rodriguez for continuing support discussions and suggestions. The authors acknowledge support through NASA-KSC STTR contract NNK04OA28C, in collaboration with Mnemonics Inc. and the McKnight Doctoral Fellowship Program.

REFERENCES

[1] S. J. Martin and A. J. Ricco, "Effective Utilization of Acoustic Wave Sensor Responses: Simultaneous Measurement of Velocity and Attenuation," presented at Ultrasonics Symposium, 1989. Proceedings, IEEE 1989, 1989, pp. 621-625 vol.1.

[2] D. S. Ballantine, R. M. White, S. J. Martin, A. J. Ricco, E. T. Zellers, G. C. Frye, and H. Wohltjen, *Acoustic Wave Sensors: Theory, Design and Physico-Chemical Applications*. New York: Academic Press, 1997.

[3] B. Fisher and D. C. Malocha, "Room Temperature Hydrogen Gas Sensing Using SAW Devices," in *Proceedings of the Joint Japan-USA International Meeting on Acoustic Wave Devices*. Orlando, FL, 2008, pp. 125-131.

[4] N. Y. Kozlovski and D. C. Malocha, "SAW Passive Wireless Multi Sensor System," presented at Ultrasonics Symposium (IUS), 2009 IEEE International, 2009, pp. 1541-1544.

[5] B. Fisher and D. Malocha, "Study of the Acoustoelectric Effect for SAW sensors," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 57, no. 3, pp. 698-706, 2010.

[6] A. R. Hutson and L. W. Donald, "Elastic Wave Propagation in Piezoelectric Semiconductors," *Journal of Applied Physics*, vol. 33, no. 1, pp. 40-47, 1962.

[7] K. A. Ingebrigtsen, "Linear and Nonlinear Attenuation of Acoustic Surface Waves in a Piezoelectric Coated with a Semiconducting Film," *Journal of Applied Physics*, vol. 41, no. 2, pp. 454-459, 1970.

[8] R. Adler, "Simple Theory of Acoustic Amplification," *IEEE Transactions on Sonics and Ultrasonics*, vol. 18, no. 3, pp. 115-117, 1971.

[9] S. Datta, *Surface Acoustic Wave Devices*. New Jersey: Prentice-Hall, 1986.

[10] B. H. Fisher and D. C. Malocha, "A Study on the Aging of Ultra-Thin Palladium Films on SAW Hydrogen Gas Sensors," presented at IEEE International Frequency Control Symposium (FCS) 2010, pp. 242-247.

[11] A. J. Slobodnik, Jr., "Attenuation of Microwave Acoustic Surface Waves Due to Gas Loading," *Journal of Applied Physics*, vol. 43, no. 6, pp. 2565-2568, 1972.

[12] A. J. Slobodnik, Jr., "Surface Acoustic Waves and SAW Materials," *Proceedings of the IEEE*, vol. 64, no. 5, pp. 581-595, 1976.

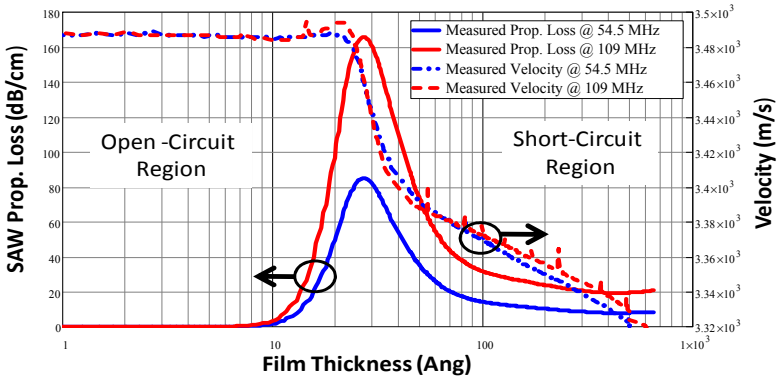


Figure 13 shows a plot of the SAW propagation loss (solid lines) and velocity (dashed lines) at 54.5 and 109MHz as a function of the Pd film thickness. Pd film's conductivity changes due to the film's thickness. Consequently, the SAW propagation loss and velocity is modulated via the acoustoelectric effect.

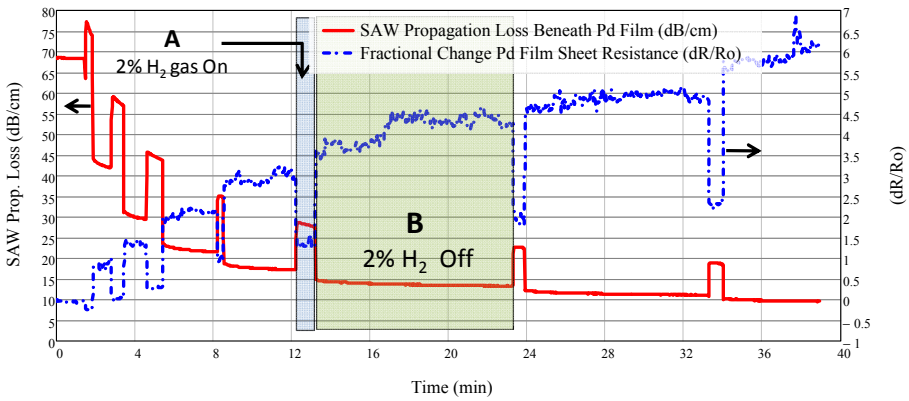


Figure 14 shows the real-time observation and measurement of the modulation of Pd film's conductivity in the presence of H_2 gas. Consequently, the SAW propagation loss is also modulated via the acoustoelectric effect. A: The blue shaded area shows the response of the SAW when the Pd film is exposed to 2% H2 gas. B: Green shaded area shows the response when the film is not exposed to 2% H_2 gas.

Clock Current State Estimation with a Kalman-Like Algorithm Employing Measurement of Time Errors

Yuriy S. Shmaliy and Oscar Ibarra-Manzano
DICIS, Guanajuato University, Salamanca, Gto, 36855, Mexico
Email: shmaliy@salamanca.ugto.mx

***Abstract*—To meet the needs of the IEEE Standard 1139-1999 in part of the estimations of clock state, we propose using a fast recursive p-shift unbiased estimation algorithm employing measurement of the time interval error. A special feature of this algorithm is that it has the Kalman-like form, although ignores both the clock and measurement noise sources and allows for any initial clock state. The algorithm can straightforwardly be applied to provide unbiased filtering ($p = 0$) and smoothing ($p < 0$) of clock state, also to predict ($p > 0$) its future behavior. An application is given for a master clock.**

I. Introduction

The IEEE Standard 1139-1999 [1] suggests that the clock has three states, namely the time interval error (TIE), fractional frequency offset, and linear frequency drift rate. Accurate clock estimation faces two main factors. Firstly, noise in free-run clocks has slow and very slow flicker (colored) and random walk components making TIE nonstationary on a finite interval even with zeroth second and third states. Secondly, ignoring the higher order states may cause uncertainties. We notice that the latter effect is still not well studied. The problem is complicated with both the measurement noise and uncertainties when the reference source of time is removed, as in the Global Navigation Satellite Systems (GNSS).

Under such conditions, the Kalman algorithm derived for white noise and adapted for clocks in [2] may demonstrate divergence [3], become unstable [4], [5], and even noisy [6]. There is one more important obstacle preventing using such an algorithm for accurate estimation in clocks. For the Kalman algorithm to be optimal, the noise covariances must be known exactly. It is known, however, that the phase noise power law is hard to measure close to the carrier and the covariance function is also hard to specify. The clock model thus does not suites the Kalman estimator exactly and the latter typically never gives us estimates better than weighted averaging widely employed in time scales, although not in state space. This fact was neatly demonstrated experimentally in [7], meaning that some more robust solutions are desirable.

Examining the trade-off between the infinite impulse response (IIR) recursive Kalman algorithm and finite impulse response (FIR) ones, Jazwinski summarized in [4] that the limited memory filter, having FIR, appears to be the only device for preventing divergence in the presence of unbounded perturbation in the system. That means that the optimal FIR filter, having an imbedded bounded input/bounded output (BIBO) stability, is more robust against the Kalman one [8]. On the other hand, the Standard [1] states that "an efficient and unbiased estimator is preferred." Referring to and aimed at overcoming the problems with the Kalman filter, transversal FIR filtering has been employed in [5] to design an unbiased estimator for clocks and then investigated in detail in [6], [9]–[11]. As a result, an optimal FIR estimator was proposed for clocks in [12]. It has also been shown in [5] that an unbiased FIR filter allowing noise to have any distribution and covariance becomes virtually optimal in the mean square error (MSE) sense if averaging is provided on a large interval of $N \gg 1$ points [11], as it is in GNSS-based timekeeping. Below we shall show that the unbiased FIR estimator discussed in [12] in a batch form has an efficient Kalman-like algorithm that, in applications to clocks, outperforms the Kalman filter.

II. Unbiased Estimator

In discrete time n, the clock TIE x_n is commonly measured by a precise time interval counter as s_n with a sampling (averaging) time τ for some reference time scale [7]. The measurement noise v_n is typically negligible with direct measurement and must be accounted for if the reference source is removed, as in GNSS.

Following [1], we characterize the clock with three states and model in state space with the time-invariant state and observation equations, respectively, as

$$\mathbf{x}_n = \mathbf{A}\mathbf{x}_{n-1} + \mathbf{w}_n \,, \tag{1}$$

$$s_n = \mathbf{C}\mathbf{x}_n + v_n \,, \tag{2}$$

in which the 3×1 state vector is given by

$$\mathbf{x}_n = [x_n \, y_n \, z_n]^T \,, \tag{3}$$

where x_n is the first state (TIE), y_n is the second state (fractional frequency offset), and z_n is the third state (linear frequency drift rate). The 3×3 matrix $\mathbf{A}$ projects the nearest past state $\mathbf{x}_{n-1}$ to the present state $\mathbf{x}_n$ employing the finite-degree Taylor series expansions,

$$\mathbf{A} = \begin{bmatrix} 1 & \tau & \tau^2/2 \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{bmatrix} \,, \tag{4}$$

and the 1×3 observation matrix is $\mathbf{C} = [1\,0\,0]$, for measurement s_n of x_n. The 3×1 vector of the clock zero-mean noise,

978-1-61284-110-6/11/$26.00 ©2011 IEEE

$$\mathbf{w}_n = [w_{xn}\, w_{yn}\, w_{zn}]^T\,, \tag{5}$$

has the covariance $\mathbf{\Psi} = E\{\mathbf{w}_i \mathbf{w}_j^T\}$ with supposedly arbitrary components formed by different clock noise sources (white and flicker) and high-order states. The measurement noise v_n is commonly zero-mean, $E\{v_n\} = 0$, with the variance $E\{v_n^2\} = \sigma_v^2$. Most generally, v_n can also have arbitrary distribution and covariance.

Because the clock model is time-invariant and noise has flicker and random walk noise components, estimation gives particularly accurate results if provided over all data available, from zero to the current point n, thus on a full averaging horizon. It follows from [5], [10], [11] that the relevant full-horizon p-shift unbiased FIR estimate[1] $\bar{\mathbf{x}}_{n+p|n}$ of $\mathbf{x}_n$ at the point $n+p$ employing measurement s_n has the batch discrete convolution form of

$$\bar{\mathbf{x}}_{n+p|n} = \mathbf{A}^{n+p}(\mathbf{C}_n^T\mathbf{C}_n)^{-1}\mathbf{C}_n^T\mathbf{S}_n\,, \tag{6}$$

where $\mathbf{A}^{n+p}$ is the power of $\mathbf{A}$ specified with (4). Here $\mathbf{C}_n$ and the measurement vector $\mathbf{S}_n$ are given by, respectively,

$$\mathbf{C}_n = \left[(\mathbf{C}\mathbf{A}^n)^T\,(\mathbf{C}\mathbf{A}^{n-1})^T\,\ldots\,(\mathbf{C}\mathbf{A})^T\,\mathbf{C}^T\,\right]^T\,, \tag{7}$$

$$\mathbf{S}_n = [s_n\, s_{n-1}\,\ldots\, s_0]^T\,. \tag{8}$$

As can be seen, the unbiased estimator (6) ignores noise and initial conditions, unlike the Kalman one requiring both of them to be white and known. Although (6) can universally be used for unbiased filtering ($p = 0$), prediction ($p > 0$), and smoothing ($p < 0$) of clock state employing all the data available, the computational problem arises when measurement is provided on a large horizon of $N \gg 1$ points.

III. Kalman-like Unbiased Estimator

To reduce the computational complexity and make possible for the estimate $\bar{\mathbf{x}}_{n+p|n}$ to exist in real time associated with the one pulse per second (1PPS) measurement, the batch algorithm (6) can be substituted with the recursive one justified in Appendix A to have the blocks listed below.

Given the number K of clock states and a time shift p:

1) Set auxiliary matrices :

$$\mathbf{\Xi} = \mathbf{A}^T\mathbf{C}^T\mathbf{C}\mathbf{A}\,, \tag{9}$$

$$\mathbf{P}_{K-1} = (\mathbf{C}_{K-1}^T\mathbf{C}_{K-1})^{-1}\,, \tag{10}$$

$$\mathbf{F}_{K-1} = \mathbf{A}^{K-1}\mathbf{P}_{K-1}(\mathbf{A}^{K-1})^T\,. \tag{11}$$

2) Compute the initial clock state at $K-1+p$:

$$\bar{\mathbf{x}}_{K-1+p|K-1} = \mathbf{A}^{K-1+p}\mathbf{P}_{K-1}\mathbf{C}_{K-1}^T\mathbf{S}_{K-1}\,. \tag{12}$$

3) Update:

$$\mathbf{F}_n = \mathbf{A}\mathbf{F}_{n-1}\mathbf{A}^T - \mathbf{A}\mathbf{F}_{n-1}(\mathbf{I}+\mathbf{\Xi}\mathbf{F}_{n-1})^{-1}\mathbf{\Xi}\mathbf{F}_{n-1}\mathbf{A}^T\,, \tag{13}$$

$$\bar{\mathbf{x}}_{n+p|n} = \mathbf{A}\bar{\mathbf{x}}_{n-1+p|n-1} + \mathbf{A}^p\mathbf{F}_n\mathbf{C}^T \times(\mathbf{s}_n - \mathbf{C}\mathbf{A}^{1-p}\bar{\mathbf{x}}_{n-1+p|n-1})\,. \tag{14}$$

Instantly one indicates that (14) is the Kalman-like estimate, in which $\mathbf{A}^p\mathbf{F}_n\mathbf{C}^T$ plays the role of the Kalman gain. A remarkable engineering property of this algorithm is that it needs only two parameters, K and p, and does not depend on noise and initial error. For filtering the current clock state with $p = 0$, the algorithm (9)–(14) can be applied straightforwardly. For prediction with a step $p > 0$ or smoothing with a lag $p < 0$, p must be considered as a new variable as related to each fixed n. Below, we give an application for this algorithm to a Master Clock.

IV. Current State Estimation of the NIST MC

The NIST has published on the WEB site the UTC – UTC(NIST MC) time differences (285 points) measured each 10 days in 2002–2009, as issued monthly by BIPM [units are in Modified Julian Dates (MJDs) and nanoseconds]. For this measurement, we form the time scale starting with $n = 0$ (52279 MJD) and finishing at $n = 284$ (55129 MJD). The measurement is shown in Fig. 1a (circles).

Supposing that the NIST MC has two states, $K = 2$, the algorithm (9)–(14) was developed for $n \geqslant K$ with $\tau = 8.64\times 10^5$ s to

$$\mathbf{\Xi} = \mathbf{A}^T\mathbf{C}^T\mathbf{C}\mathbf{A}\,, \tag{15}$$

$$\mathbf{F}_1 = \mathbf{A}(\mathbf{C}_1^T\mathbf{C}_1)^{-1}\mathbf{A}^T\,, \tag{16}$$

$$\bar{\mathbf{x}}_{1+p|1} = \mathbf{A}(\mathbf{C}_1^T\mathbf{C}_1)^{-1}\mathbf{C}_1^T\mathbf{S}_1\,, \tag{17}$$

$$\mathbf{F}_n = \mathbf{A}\mathbf{F}_{n-1}\mathbf{A}^T - \mathbf{A}\mathbf{F}_{n-1}(\mathbf{I}+\mathbf{\Xi}\mathbf{F}_{n-1})^{-1}\mathbf{\Xi}\mathbf{F}_{n-1}\mathbf{A}^T\,, \tag{18}$$

$$\bar{\mathbf{x}}_{n+p|n} = \mathbf{A}\bar{\mathbf{x}}_{n-1+p|n-1} + \mathbf{F}_n\mathbf{C}^T(s_n - \mathbf{C}\mathbf{A}\bar{\mathbf{x}}_{n-1+p|n-1})\,, \tag{19}$$

where, by (7), we have

$$\mathbf{C}_1 = \begin{bmatrix} 1 & \tau \\ 1 & 0 \end{bmatrix}.$$

Figure 1 sketches the filtering estimates $\bar{x}_{n|n}$ and $\bar{y}_{n|n}$ of x_n and y_n, respectively, provided by (15)–(19), and we notice that this algorithm produces estimates appreciably faster than the batch one (6) or (11b) discussed in [16].

For a comparison, we also give estimates obtained with the standard 2-state Kalman filter described in [5] (see Appendix B). Because the NIST MC exhibits excellent etalon properties, we set for the Kalman filter zeroth initial conditions, $x_0 = 0$ and $y_0 = 0$. For the resolution of 0.1 ns in the published data, the digitization noise was supposed to be distributed uniformly and its variance calculated as $\sigma_v^2 = 0.05^2/3$ ns^2 $= 8.33\times 10^{-4}$ ns^2.

[1] Here and in the following, $\bar{\mathbf{x}}_{n+p|n}$ means the unbiased estimate at $n+p$ via measurement from the past to n.

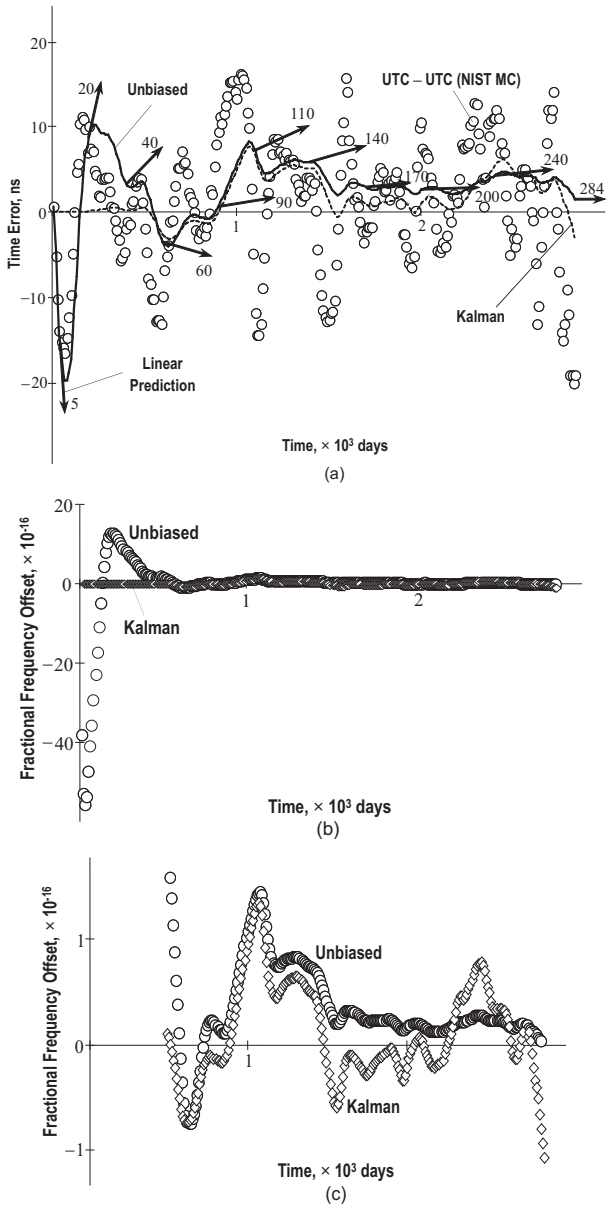


Fig. 1. Estimates of the NIST MC current state via the UTC – UTC(NIST MC) time differences (285 points) measured in 2002–2009 each 10 days: (a) first state x_n, (b) second state y_n, and (c) second state y_n (scaled).

Current estimates of the first state are shown in Fig. 1a along with the linear prediction of time errors depicted with arrows. Since the unbiased estimator gives us the best linear unbiased fit, the first arrow coincides in the direction with several initial measurement points. Increased n, prediction depicts possible behaviors of future errors. Although, the process shown in Fig. 1a looks like rather stationary, prediction reveals some positive average angle resulting in a positive frequency offset diminishing with time (Fig. 1b and Fig. 1c).

Observing these figures, one can arrive at a conclusion that the Kalman filter is lesser suitable for the clock current state estimation. In fact, owing to transients, the Kalman filter does not provide a real picture with a small number of measurements. In the intermediate region (about the point of 1×10^3 days), both filters produce consistent estimates. Finally, the Kalman-like filter (15)–(19) keeps smoothing the data above the intermediate region, whereas the Kalman one exhibits qualitatively the same error picture, being time-invariant.

V. Conclusions

We proposed and examined a Kalman-like p-shift recursive unbiased FIR estimation algorithm intended for filtering ($p = 0$), prediction ($p > 0$), and smoothing ($p < 0$) of the clock state space model via all measurement data available. The algorithm provides the best unbiased fit for the measurement data and is fast, unlike its batch origin. Its remarkable engineering property resides in the independence on noise and initial conditions. As an application, we estimated the current state of the NIST MC employing the UTC – UTC(NIST MC) time differences (285 points) measured each 10 days in 2002–2009, as issued monthly by BIPM. A comparison with the 2-state Kalman filter is also provided and we state that the latter is lesser suitable for clocks than the proposed Kalman-like one.

Appendix A
Full Horizon Recursive Unbiased FIR Estimation Algorithm

Consider the batch unbiased estimate (6). Following the derivation developed for recursive least squares [14], assign $\mathbf{P}_n^{-1} = \mathbf{C}_n^T \mathbf{C}_n$ and expend this matrix to

$$\begin{aligned} \mathbf{P}_n^{-1} &= \sum_{i=0}^{n} \left(\mathbf{C}\mathbf{A}^{n-i}\right)^T \mathbf{C}\mathbf{A}^{n-i} \\ &= \sum_{i=1}^{n} \left(\mathbf{C}\mathbf{A}^{n-i}\right)^T \mathbf{C}\mathbf{A}^{n-i} + \left(\mathbf{C}\mathbf{A}^{n}\right)^T \mathbf{C}\mathbf{A}^{n} \\ &= \mathbf{P}_{n-1}^{-1} + \left(\mathbf{C}\mathbf{A}^{n}\right)^T \mathbf{C}\mathbf{A}^{n} . \end{aligned} \tag{A.1}$$

Using the matrix inversion lemma [15]

$$(\mathbf{A}+\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{I}+\mathbf{C}\mathbf{A}^{-1})^{-1}\mathbf{C}\mathbf{A}^{-1} ,$$

represent the inverse of (A.1) as

$$\begin{aligned} \mathbf{P}_n &= \mathbf{P}_{n-1} - \mathbf{P}_{n-1}(\mathbf{I} + \mathbf{A}^{nT}\mathbf{C}^T\mathbf{C}\mathbf{A}^n\mathbf{P}_{n-1})^{-1} \\ &\quad \times \mathbf{A}^{nT}\mathbf{C}^T\mathbf{C}\mathbf{A}^n\mathbf{P}_{n-1} . \end{aligned} \tag{A.2}$$

Expend the product $\mathbf{C}_n^T\mathbf{S}_n$ in (6) as follows:

$$\begin{aligned} \mathbf{C}_n^T\mathbf{S}_n &= \sum_{i=0}^{n} \left(\mathbf{C}\mathbf{A}^{n-i}\right)^T \mathbf{S}_{n-i} \\ &= \sum_{i=1}^{n} \left(\mathbf{C}\mathbf{A}^{n-i}\right)^T \mathbf{S}_{n-i} + \left(\mathbf{C}\mathbf{A}^{n}\right)^T s_n \\ &= \mathbf{C}_{n-1}^T\mathbf{S}_{n-1} + \left(\mathbf{C}\mathbf{A}^{n}\right)^T s_n . \end{aligned} \tag{A.3}$$

Now rewrite the estimate (6) at $n-1+p$ as

$$\bar{\mathbf{x}}_{n-1+p|n-1} = \mathbf{A}^{n-1+p}(\mathbf{C}_{n-1}^T\mathbf{C}_{n-1})^{-1}\mathbf{C}_{n-1}^T\mathbf{S}_{n-1}\,,$$

use the identities produced by $\mathbf{P}_n^{-1} = \mathbf{C}_n^T\mathbf{C}_n$ and (A.1), respectively,

$$\begin{aligned}\mathbf{P}_{n-1}^{-1} &= \mathbf{C}_{n-1}^T\mathbf{C}_{n-1}\,,\\ \mathbf{P}_{n-1}^{-1} &= \mathbf{P}_n^{-1} - \mathbf{A}^{nT}\mathbf{C}^T\mathbf{C}\mathbf{A}^n\,,\end{aligned}$$

and express $\mathbf{C}_{n-1}^T\mathbf{S}_{n-1}$ as

$$\begin{aligned}\mathbf{C}_{n-1}^T\mathbf{S}_{n-1} &= \mathbf{C}_{n-1}^T\mathbf{C}_{n-1}\mathbf{A}^{-n+1-p}\bar{\mathbf{x}}_{n-1+p|n-1}\\ &= \mathbf{P}_{n-1}^{-1}\mathbf{A}^{-n+1-p}\bar{\mathbf{x}}_{n-1+p|n-1}\\ &= (\mathbf{P}_n^{-1} - \mathbf{A}^{nT}\mathbf{C}^T\mathbf{C}\mathbf{A}^n)\\ &\quad\times\mathbf{A}^{-n+1-p}\bar{\mathbf{x}}_{n-1+p|n-1}\,. \qquad \text{(A.4)}\end{aligned}$$

By $\mathbf{P}_n^{-1} = \mathbf{C}_n^T\mathbf{C}_n$ and (A.3), the estimate (6) can now be rewritten as

$$\begin{aligned}\bar{\mathbf{x}}_{n+p|n} &= \mathbf{A}^{n+p}\mathbf{P}_n\mathbf{C}_n^T\mathbf{S}_n\\ &= \mathbf{A}^{n+p}\mathbf{P}_n[\mathbf{C}_{n-1}^T\mathbf{S}_{n-1} + (\mathbf{C}\mathbf{A}^n)^T s_n]\,. \qquad \text{(A.5)}\end{aligned}$$

Taking $\mathbf{C}_{n-1}^T\mathbf{S}_{n-1}$ from (A.4) and then providing the transformations lead to

$$\begin{aligned}\bar{\mathbf{x}}_{n+p|n} &= \mathbf{A}\bar{\mathbf{x}}_{n-1+p|n-1} + \mathbf{A}^{n+p}\mathbf{P}_n\mathbf{A}^{nT}\mathbf{C}^T\\ &\quad\times\left(s_n - \mathbf{C}\mathbf{A}^{1-p}\bar{\mathbf{x}}_{n-1+p|n-1}\right)\,. \qquad \text{(A.6)}\end{aligned}$$

Further assign

$$\mathbf{F}_n = \mathbf{A}^n\mathbf{P}_n\mathbf{A}^{nT}\,,$$

multiply (A.2) with $\mathbf{A}^{n-1}$ and $(\mathbf{A}^{n-1})^T$ from the left-hand and right-hand sides, respectively, provide the transformations, and finally arrive at the fast recursive Kalman-like algorithm:

$$\boldsymbol{\Xi} = \mathbf{A}^T\mathbf{C}^T\mathbf{C}\mathbf{A}\,, \qquad \text{(A.7)}$$

$$\mathbf{F}_{K-1} = \mathbf{A}^{K-1}(\mathbf{C}_{K-1}^T\mathbf{C}_{K-1})^{-1}(\mathbf{A}^{K-1})^T\,, \qquad \text{(A.8)}$$

$$\mathbf{F}_n = \mathbf{A}\mathbf{F}_{n-1}\mathbf{A}^T - \mathbf{A}\mathbf{F}_{n-1}(\mathbf{I}+\boldsymbol{\Xi}\mathbf{F}_{n-1})^{-1}\boldsymbol{\Xi}\mathbf{F}_{n-1}\mathbf{A}^T\,, \qquad \text{(A.9)}$$

$$\begin{aligned}\bar{\mathbf{x}}_{n+p|n} &= \mathbf{A}\bar{\mathbf{x}}_{n-1+p|n-1} + \mathbf{A}^p\mathbf{F}_n\mathbf{C}^T\\ &\quad\times(\mathbf{s}_n - \mathbf{C}\mathbf{A}^{1-p}\bar{\mathbf{x}}_{n-1+p|n-1})\,, \qquad \text{(A.10)}\end{aligned}$$

in which K is the number of clock states. Update (A.9) and (A.10) for $n \geqslant K-1$, because the inverse in (A.8) does not exists otherwise.

The initial state estimate $\bar{\mathbf{x}}_{K-1+p|n-1}$ for (A.10) can be found using (6) as

$$\bar{\mathbf{x}}_{K-1+p|n-1} = \mathbf{A}^{K-1+p}\mathbf{P}_{K-1}\mathbf{C}_{K-1}^T\mathbf{S}_{K-1}\,. \qquad \text{(A.11)}$$

Note that the batch computation (A.11) does not take much time, because the number of the clock states is limited in [1] with $K = 3$.

REFERENCES

[1] *IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology – Random Instabilities*, IEEE Standard 1139-1999; IEEE: Piscataway, NJ, 1999, pp. 1-36.

[2] P. V. Tryon and R. H. Jones, "Estimation of parameters in models for cesium beam atomic clocks," *J. Res. National Bureau of Standards*, pp 3–16, 1983.

[3] R. J. Fitzgerald, "Divergence of the Kalman filter," *IEEE Trans. Autom. Control*, vol. AC-16, no. 6, pp 736–747, Dec. 1971.

[4] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[5] Y. S. Shmaliy, "An unbiased FIR filter for TIE model of a local clock in applications to GPS-based timekeeping," *IEEE Trans. on Ultrason., Ferroel. and Freq. Contr.*, vol. 53, no. 5, pp. 862–870, May 2006.

[6] Y. S. Shmaliy, "Optimal gains of FIR estimators for a class of discrete-time state-space models," *IEEE Signal Process. Letters*, vol. 15, pp. 517–520, 2008.

[7] Y. S. Shmaliy, *GPS-based Optimal FIR Filtering of Clock Models*, Nova Science Publ., New York, 2009.

[8] W. H. Kwon, P. S. Kim, and P. Park, "A receding horizon Kalman FIR filter for discrete time-invariant systems," *IEEE Trans. Automatic Control* vol. 44, no. 9, pp. 1787–1791, Sep. 1999.

[9] Y. S. Shmaliy, O. Ibarra-Manzano, L. Arceo-Miquel, and J. Muñoz-Diaz, "A thinning algorithm for GPS-based unbiased FIR estimation of a clock TIE model," *Measurement*, vol. 41, no. 5, pp. 538–550, Jun. 2008.

[10] L. Arceo-Miquel, Yu. S. Shmaliy and O. Ibarra-Manzano, "Optimal synchronization of local clocks by GPS 1PPS timing signals using predictive FIR filters," *IEEE Trans. on Instrum. and Measur*, vol. 58, no. 6, pp. 1833–1840, Jun. 2009.

[11] Y. S. Shmaliy and O. Ibarra-Manzano, "Optimal FIR filtering of the clock time errors," *Metrologia*, vol. 45, no. 5, pp. 571–576, Sep. 2008.

[12] Y. S. Shmaliy, "Full-horizon optimal FIR estimator of clock state employing measurement of time errors," *Int. J. Electron. Commun. (AEÜ)*, vol. 65, no. 2, pp. 117–122, Feb. 2011.

[13] J. W. Chaffee, "Relating the Allan variance to the diffusion coefficients of a linear stochastic differential equation model for precision oscillators," *IEEE Trans. on Ultrason., Ferroel. and Freq. Contr.*, vol. 34, no. 6, pp, 655–658, Nov. 1987.

[14] K. M. Passino and S. Yurkovich, *Fuzzy Control*, Addison-Wesley, Amsterdam, 1998.

[15] G. H. Golub and G. F. van Loan, *Matrix Computations,* 3rd Ed., The John Hopkins Univ. Press, Baltimore, 1996.

[16] Y. S. Shmaliy, "Linear unbiased prediction of clock errors," *IEEE Trans. on Ultrason., Ferroel. and Freq. Control*, vol. 56, no. 9, pp. 2027–2029, Sep. 2009.

Summary of the Link Calibration between NIM and PTB Using a Traveling GPS Receiver

Kun Liang[*+], Thorsten Feldmann[+], Andreas Bauch[+], Dirk Piester[+], Aimin Zhang[*], Xiaoxun Gao[*]

[*]National Institute of Metrology (NIM), Bei San Huan Dong Lu 18, 100013 Beijing, P. R. China

[+]Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany

Email: liangk@nim.ac.cn

***Abstract*—The time transfer link between NIM and PTB was for the first time calibrated using a travelling GPS receiver developed by NIM The calibration uncertainty was estimated to about 2 ns. The result obtained could be compared to the result derived from differential calibrations made by BIPM in previous years. Both results agree well within 1 ns.**

I. Foreword

In order to contribute to the realization of TAI, a calibration of the time transfer link between National Institute of Metrology (NIM), P. R. China, and Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany, is needed. A self-developed traveling GPS (Global Positioning System) receiver of NIM, NIM acronym IMEU, which is based on a JAVAD E_GGD receiver has been assembled and operated as traveling GPS receiver (TR) at PTB and NIM in succession for that purpose. After the suitability for the purpose of the IMEU receiver had been verified [1], the method using it for link calibrations has been developed [2]. The operation mode of IMEU differs from that of the GTR50 receiver which was used in previous successful calibration campaigns [3,4]. There are, however, several similarities.

II. Description of the link calibration method

When a GPS time transfer link between two timing labs, designated as timing lab 1 and timing lab 2 has been established, usually neither the antenna delays nor the internal receiver delays of the GPS receivers involved in the link are known. Both quantities are difficult to measure absolutely even though some groups have investigated absolute internal receiver delay calibration successfully [5~7]. Therefore the time transfer results including the difference of internal receiver delays and the difference of the antenna delays do not represent the true time difference between the two time scales or clocks involved. The GPS link thus needs to be calibrated. For this aim, we can use one TR that records GPS measurement data in succession on both sites. This link calibration method is described below. Firstly we do the common clock difference (CCD) experiment with TR and the GPS receiver R1 that is used for time transfer on this link at timing lab 1, and $C1$ is obtained, as illustrated in Figure 1. Then we do the same CCD experiment at timing lab 2 with TR and the GPS receiver R2 that is used for time transfer on this link at timing lab 2 and $C2$ is obtained. At the same time, we should take the reference delay difference RDD of TR at the different labs into account: At timing lab 1 the reference delay of TR is $RD1$, and at timing lab 2 the reference delay of TR is $RD2$, and the reference delay designates the time difference between the on-time reference point of the local time scale and the input signal provided to the receiver. If TR is operated when its time scale is locked to the external reference frequency input, the difference $XO1$ minus $XO2$, each representing the differences between the internal reference and 1PPS input of TR at the respective timing labs, designated XOD, should also be taken care of. Immediately after the operation at lab 1 and lab 2 the same CCD experiment at timing lab 1, called closure experiment, should be repeated to check for any change of the internal receiver delay of TR. Any observed change of $C1$ that is the difference between $C1$ and $C3$ would have to be accounted for in the uncertainty evaluation. By calculating the single difference of two CCD experiments, the time transfer result of the GPS link except the difference of two time references connected to two receivers can be acquired, and this is the calibration value C that can be calculated from (1) ~ (4) and we need indeed. Figure 1 illustrates the whole concept.

$$RDD = RD1 - RD2 \tag{1}$$

$$XOD = XO1 - XO2 \tag{2}$$

$$CorD = RDD + XOD \tag{3}$$

$$C = C1 - C2 + CorD \tag{4}$$

Sponsored by AQSIQ plan project 2010QK387

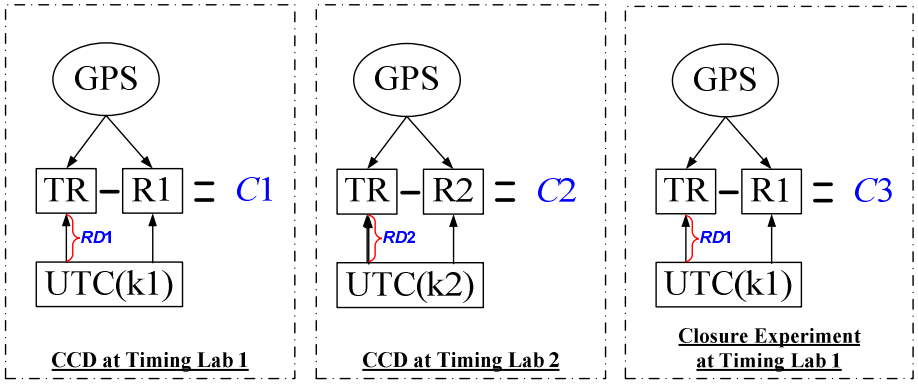


Figure 1. Schematics of a link calibration by a travelling GPS receiver TR

Once *C* has been determined, the time transfer results of the link should be corrected by this value *C* and we speak of a calibrated link. Furthermore, the calibrated GPS link can be used to calibrate a TWSTFT (Two-Way Satellite Time and Frequency Transfer) link or any other existing GPS link using different equipment. This option is particularly attractive because a mobile TWSTFT station is hard to get for making a calibration and some TWSTFT links cannot be calibrated by a mobile TWSTFT station. The latter situation had prevailed for the TWSTFT link between Asian and European institutes established through the IS-4 satellite until spring 2010 [8].

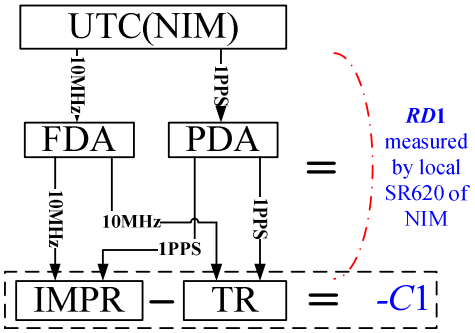
III. Experimental setup of the Link calibration between NIM and PTB

IMEU has been installed and operated as TR at NIM and PTB for time transfer link calibration between NIM and PTB as illustrated in Figure 2. The receivers used in the CCD experiments at NIM and PTB are listed in Table 1.

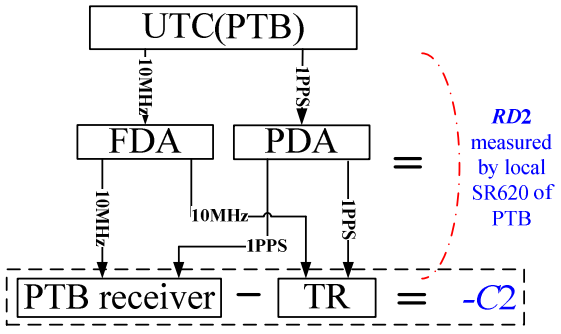
TABLE I. STATIONARY RECEIVERS USED FOR THE NIM-PTB LINK CALIBRATION

Timing lab	Receiver name	Receiver model
NIM	IMPR	Septentrio PolaRx2eTR
PTB	PTBB, PTBG	Ashtech Z12-T
PTB	TTS3	AOS TTS-3

The first CCD was made at PTB from MJD 55265 to MJD 55275. After TR had come back to NIM, we did the CCD experiments at NIM from MJD 55289 to MJD 55302. The schemes of the CCD experiments at NIM and PTB are shown in Figure 2. Due to some reasons, there is not any closure experiment to be reported here. During the link calibration, the 30-s RINEX files were recorded by all the receivers. The P3 code CGGTTS files of TTS3 which are generated and recorded automatically by the receiver and the P3 code CGGTTS files of PTBB and PTBG which are available on the BIPM Time Department FTP server were used in the calibration. In addition we used R2cggtts V4.3 [9,10] to convert the RINEX files of IMEU and IMPR to P3 code CGGTTS files that were used in the calibration.



(a) CCD at NIM



(b) CCD at PTB

Figure 2. Hardware set-up during the link calibration experiments

The reference delays *RD*1 and *RD*2 of TR were measured by SR620 Time Interval Counters (TIC), the difference *RDD* of *RD*1 and *RD*2 that were calculated according to (1) and the corresponding standard deviations of the measurements are given in TABLE II.

TABLE II. REFERENCE DELAYS (*RD*) OF TR

Timing lab	Reference delay	Measurement (ns)	Standard deviation (ns)
NIM	*RD*1	40.294	0.010
PTB	*RD*2	100.637	0.007
NIM-PTB	*RDD*	-60.343	0.013

The internal reference delays *XO*1 and *XO*2 that were measured by SR620 TIC and the difference *XOD* of *XO*1 and *XO*2 that was calculated according to (2) and the corresponding standard deviations of the measurements are listed in TABLE III.

TABLE III. INTERNAL REFERENCE DELAYS (*XO*) OF TR

Timing lab	Internal reference delay	Measurement (ns)	Standard deviation (ns)
NIM	*XO*1	-4.015	0.017
PTB	*XO*2	-4.604	0.012
NIM-PTB	*XOD*	0.589	0.021

So *CorD* according to (3) and the corresponding standard deviation of the measurements are described as Table 4.

TABLE IV. CORRECTION (*CorD*) OF TR

Timing lab	Correction	Measurement (ns)	Standard deviation (ns)
NIM-PTB	*CorD*	-59.754	0.025

It is always considered as an option to use Common-View (CV) instead of All-in View (AV) data analysis to calculate the calibration value for the link as one can expect that most errors cancel in CV. We process measurement data of two kinds, code-based ionosphere-free P3 CV data and code-plus-phase combined RINEX data processed with Precise Point Positioning (PPP) software. GPS P3 CV means that we choose CGGTTS P3 measurement data of the common satellites for two stations at the same epoch to compare and then average the comparisons of all these satellites. PPP results are obtained from NRCan-PPP software [11]. By two methods of data processing, two kinds of link calibration values for the same pair of GPS receivers in operation are obtained. The calibration of the officially operational GPS time transfer link, however, which contributes to the TAI computation, is the more significant one. The main links between NIM and PTB in official operation are listed in TABLE V.

TABLE V. OFFICIALLY OPERATIONAL MAIN LINKS BETWEEN NIM AND PTB

Station	PTBB	PTBG	TTS3
IMPR	P3 CV and PPP	P3 CV	P3 CV

IV. Link calibration value

The results with P3 code and PPP of the CCD experiments using TR and IMPR at NIM are shown in Figures 3 and 4. The results with P3 code and PPP of the CCD experiments using TR and PTBB, PTBG and TTS3 at PTB are shown in Figures 5 and 6. All results of the calibration values are listed in TABLE VI. The calibration values can be computed via (4), and the uncertainty is evaluated as described in Section 4 and tabulated in TABLE VII.

TABLE VI. CCD RESULTS AT NIM AND PTB

Timing lab	Receiver pair	Processing type	CCD results (ns)	Standard deviation (ns)
NIM	IMPR and TR	P3	$C1_{CV}$(IMPR)=35.731	0.767
NIM	IMPR and TR	PPP	$C1_{PPP}$(IMPR)=32.214	0.347
PTB	PTBB and TR	P3	$C2_{CV}$(PTBB)= 111.903	1.092
PTB	PTBG and TR	P3	$C2_{CV}$(PTBG)= -398.090	1.201
PTB	TTS3 and TR	P3	$C2_{CV}$(TTS3)= 126.440	1.396
PTB	PTBB and TR	PPP	$C2_{PPP}$(PTBB)= -409.828	0.291

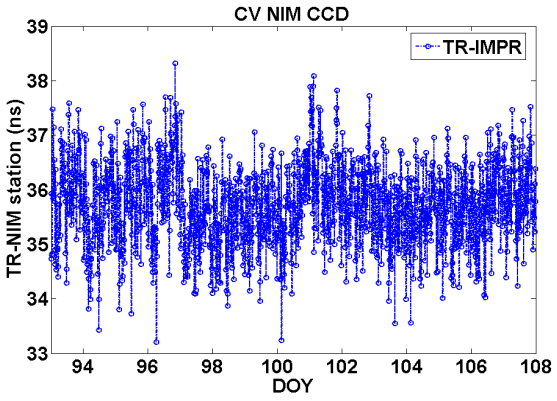


Figure 3. Results of P3 code at NIM

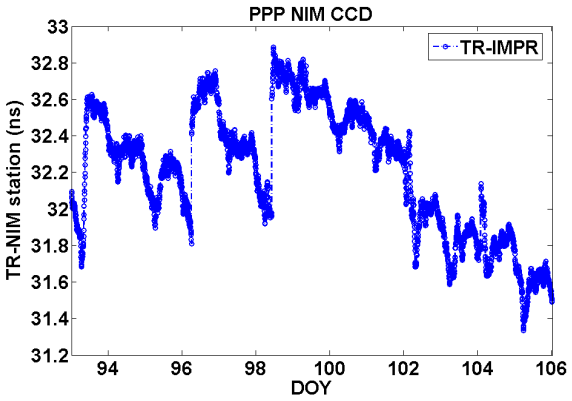


Figure 4. Results of PPP at NIM

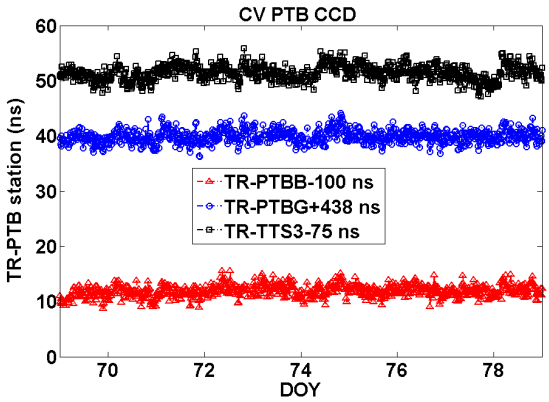


Figure 5. Results of P3 code at PTB

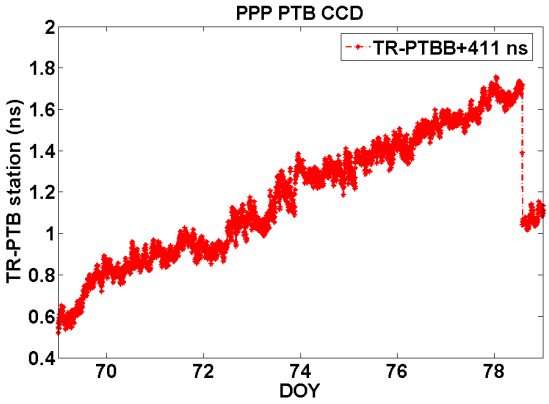


Figure 6. Results of PPP at PTB

The jumps of PPP CCD results might come from the data processing of the PPP software and be related to bad observations (mainly caused by tracking problems) of TR at NIM and PTBB at PTB. And the small drifts seen in the PPP CCD results have not been understood yet.

TABLE VII. CALIBRATION VALUES

Processing type	Calibration values (ns)	Uncertainty (ns)
P3	C_{CV}(IMPR-PTBB) = -135.9	1.8
P3	C_{CV}(IMPR-PTBG) = 374.1	1.9
P3	C_{CV}(IMPR-TTS3) = -150.5	2.0
PPP	C_{PPP}(IMPR-PTBB) = 382.3	1.4

For verification of the IMPR-PTBB P3 link calibration result, we can compare it to the previously obtained link calibration value of 135.8 ns, which is the difference between the total delays for the two stations as derived from the BIPM differential calibration results of IMPR and PTBB. The campaigns were made during MJD 55188~55194 in December 2009 at NIM [12] and during MJD 54566~54576 in April 2008 at PTB [13] covering IMPR and PTBB, respectively. For a meaningful comparison, the link calibration value from the BIPM results has to be corrected by the CGGTTS delay parameters (including the reference delay, the antenna cable delay, and the internal delay and antenna delay of the receiver) of PTBB because of the use of the original CGGTTS data of PTBB from the BIPM Time Department FTP server in our link calibration campaign. The two link calibration values from BIPM results and C_{CV}(IMPR-PTBB) apparently agree well within 1 ns after more than one year elapsed between the campaigns.

We used subsequently the calibration values to calibrate all time transfer links shown in TABLE V during MJD 55260 to 55310 and give a direct impression of the calibrated links depicted in Figure 7. The results show that all the links after calibration can agree well with one another.

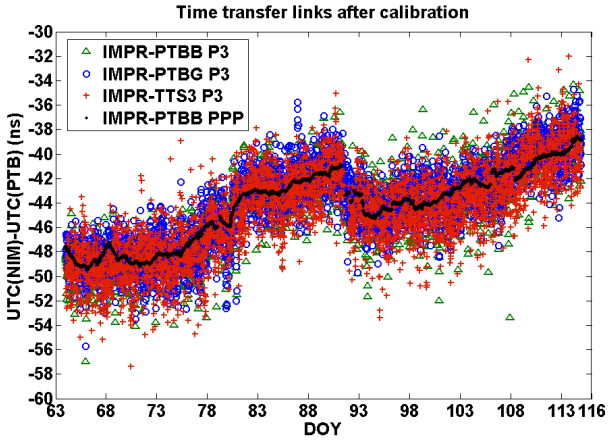


Figure 7.Calibrated time transfer links between NIM and PTB

V. Uncertainty Evaluation

For the GPS link calibration we can evaluate the uncertainty u_C as

$$u_C = \sqrt{u_A^{\,2} + u_B^{\,2}} \tag{5}$$

u_A is the evaluated uncertainty of type A, and u_B is the evaluated uncertainty of type B. They should be evaluated as

$$u_A = \sqrt{u_{A,1}^{\,2} + u_{A,2}^{\,2}} \tag{6}$$

and

$$u_B = \sqrt{\sum_{k=1}^{23} u_{B,k}^{\,2}} \tag{7}$$

$u_{A,1}$ and $u_{A,2}$ are the statistical uncertainties of the determination of the CCD results at NIM and PTB separately. $u_{B,1}$ and $u_{B,2}$ are the uncertainties due to the instability of the connections of the TR to UTC(NIM) and UTC(PTB), respectively. $u_{B,3} \sim u_{B,10}$ are the uncertainties from the measurements of the reference delay difference *RDD* of TR. $u_{B,11} \sim u_{B,16}$ are the uncertainties from the measurements of the internal reference delay difference *XOD* of TR (there is no contribution from systematic error of TIC because each *XO* would be acquired by the difference between two types of time interval measurements *X*1 and *X*2, i.e. *XO*=*X*2-*X*1, using the same TIC, by which the systematic error of the TIC cancels out). $u_{B,17}$ is the uncertainty due to the difference between the determinations of the reference points of UTC(NIM) and UTC(PTB). $u_{B,18}$ is the uncertainty due to the propagation (multipath). Atmospherical effects cancel out in a quasi zero baseline setup. $u_{B,19}$ is the uncertainty related to the instability of the antenna and its cable of TR. $u_{B,20}$ is the uncertainty due to the ambiguity estimation (only for PPP) for each receiver [14]. $u_{B,21}$ is the uncertainty due to the variation of the internal delay of TR that could be evaluated by the difference between the first CCD results and the closure CCD results. No closure measurement could be performed this time. Here we estimate instead as default value 0.80 ns which is more than the maximum corresponding value noted during the PTB-METAS link calibration [4]. $u_{B,22}$ and $u_{B,23}$ are the uncertainties due to the temperature effects (0.20 ns/degree and 0.24 ns/degree) on 1PPS output and GPS measurements separately of TR [1] at NIM and PTB, respectively. While the experiment was performed at NIM and PTB, the temperature variations at NIM and PTB were 0.8 degrees and 0.2 degrees, respectively. From that $u_{B,22}$ and $u_{B,23}$ was obtained by using (8).

$$u = \sqrt{(0.20 \times temperature\_stability)^2 + (0.24 \times temperature\_stability)^2} \tag{8}$$

The uncertainties of type B in this calibration campaign have been evaluated as listed in TABLE VIII.

TABLE VIII. UNCERTAINTY CONTRIBUTIONS OF TYPE B.

Uncertainty	Value (ns)	Description
$u_{B,1}$	0.10	Instability of the connection to UTC(NIM)
$u_{B,2}$	0.10	Instability of the connection to UTC(PTB)
$u_{B,3}$*	0.50	Systematic error of the local TIC for *RD*1 measurement at NIM
$u_{B,4}$*	0.50	Systematic error of the local TIC for *RD*2 measurement at PTB
$u_{B,5}$	0.06	TIC trigger level timing error for *RD*1 measurement at NIM
$u_{B,6}$	0.06	TIC trigger level timing error for *RD*2 measurement at PTB
$u_{B,7}$	0.01	Jitter of the TIC after 100 *RD*1 measurements at NIM
$u_{B,8}$	0.01	Jitter of the TIC after 100 *RD*2 measurements at PTB
$u_{B,9}$	0.10	TIC nonlinearities for *RD*1 measurement at NIM
$u_{B,10}$	0.10	TIC nonlinearities for *RD*2 measurement at PTB
$u_{B,11}$*	0.06	TIC trigger level timing error for *XO*1 measurement at NIM
$u_{B,12}$*	0.06	TIC trigger level timing error for *XO*2 measurement at PTB
$u_{B,13}$*	0.02	Jitter of the TIC after 100 *XO*1 measurements at NIM
$u_{B,14}$*	0.01	Jitter of the TIC after 100 *XO*2 measurements at PTB
$u_{B,15}$*	0.10	TIC nonlinearities for *XO*1 measurement at NIM
$u_{B,16}$*	0.10	TIC nonlinearities for *XO*1 measurement at PTB

$u_{B,17}$	0.02	Determination of the UTC reference point at PTB
$u_{B,18}$	0.30	Multipath
$u_{B,19}$	0.18	Antenna cable and antenna
$u_{B,20}$	0.30	Uncertainty of the ambiguity estimation(only for the PPP link)
$u_{B,21}$*	0.80	Instability of the internal receiver delay due to the loss of the closure experiment
$u_{B,22}$#	0.26	Temperature effects of TR at NIM
$u_{B,23}$#	0.07	Temperature effects of TR at PTB

Most evaluated uncertainties have a similarity with those described in [14]. The items with an asterisk do not exist in [14] due to the use of the local TICs at NIM and PTB and the TR in a different operation mode. The items with a pound have different values using the different evaluation methods from the ones in [14] because of the use of a different type of TR.

VI. Summary

The calibration of the GPS links between NIM and PTB with self-developed TR was successfully performed. The calibration uncertainty was estimated as about 2 ns. After that the result for the IMPR-PTBB P3 link in this campaign and the result for the same link by the differential calibration provided by BIPM have been compared. Both results agree well within 1 ns. The construction of the new clock room in the new campus of NIM will be finished soon, and the realization of UTC(NIM) will be transferred to the new campus. Then the TWSTFT links with NIM will be put into routine operation and calibrated based on the results presented in this contribution.

References

[1] K. Liang, T. Feldmann, A. Bauch, D. Piester, A. Zhang, X. Gao, "Performance evaluation of NIM GPS receivers inuse for time transfer with PTB," Proc the 24th European Frequency and Time Forum (EFTF), Noordwijk, The Netherlands. 13-16 April 2010.

[2] K. Liang, A. Zhang, X. Gao, "Improved Calibration Method using NIM GPS Traveling Receivers for the Link between NIM and PTB." Proc AP'RASC&ATF 2010. Oct 2010.

[3] H. Esteban, J. Palacio, F. J. Galindo, T. Feldmann, A. Bauch and D. Piester, "Improved GPS-based Time Link Calibration involving ROA and PTB," IEEE Trans. UFFC, vol. 3, 2010, pp. 714–720.

[4] T. Feldmann, A. Bauch, D. Piester, A. Stefanov, L.-G. Bernier, C. Schlunegger, K. Liang, "On improved GPS-based calibration of the time links between METAS and PTB," Proc. of the 24th European Frequency and Time Forum (EFTF), Noordwijk, The Netherlands. 13-16 April 2010.

[5] E. Powers, P. Wheeler, D. Judge and D. Matsakis, "Hardware Delay Measurements and Sensitivities in Carrier Phase Time Transfer," Proc the 27th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, VA, December 1998, pp.293-305.

[6] G. Petit, Z. Jiang, J. White, R. Beard, and E. Powers, "Absolute Calibration of an Ashtech Z12-T GPS Receiver," GPS Solutions, vol. 4, 2001, pp. 41-46.

[7] A. Proia, G. Cibiel, L. Yaigre, "Time stability and electrical delay comparison of dual frequency GPS receivers", Proc. 41st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Santa Ana Pueblo, USA, November 2009, pp. 293-301.

[8] M. Fujieda et al., "Impact of the transponder configuration on the Asia-Europe TWSTFT network", to be published in these proceedings

[9] P. Defraigne, G. Petit, C. Bruyninx, "Use of geodetic receivers for TAI," Proceedings of the 40th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Long Beach, California, December 2001.

[10] L. Tisserand, "R2CGGTTS Documentation Version 2.4.3," 2009.

[11] J. Kouba and P. Héroux, "Precise Point Positioning Using IGS Orbit and Clock Products," GPS Solutions, vol. 5 n. 2, 2001, pp. 12-28.

[12] BIPM. CalibGeo_NIM-2009. ftp://tai.bipm.org/TFG/CALIB_GEO/NIM/CalibGeo_NIM-2009.pdf [EB]. 2010.

[13] BIPM. CalibGeo_PTB-2008. ftp://tai.bipm.org/TFG/CALIB_GEO/PTB/CalibGeo_PTB-2008.pdf [EB]. 2008.

[14] T. Feldmann, A. Bauch, D. Piester, Michael Rost1,Elizabeth Goldberg, Stephen Mitchell, and Blair Fonville, "Advanced GPS based Time Link Calibration with PTB´s New GPS Calibration Set-up," Proceedings of the 39th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, VA, Octomber 2010.

.

Ovenless, Small-size, Low-power, IMU/Quartz-Oscillator Ensemble Performs like Atomic Clocks

Gus R. German and David W. Allan
Allan Space-Time Solutions, LLC
Fountain Green, UT 84632

John W. Cline and Gregory T. Rogers
Bliley Technologies, Inc.
Erie, PA 16506

Abstract— **Heretofore, all precision clocks and oscillators have required one or more ovens to control the temperature of frequency dependent components. Ovens always add significantly to the power budget and limit the upper operating temperature. In the 1970s, Jack Kusters, et al. developed the dual-mode oscillator, which allows one mode to be used to measure temperature [1]. Then the other mode is made as frequency independent with temperature as can reasonably be done. The modes used are the B- and C-modes. With knowledge of the temperature behavior for these two modes, the B-mode measurements can be used to compensate the much smaller frequency dependent changes in the C-mode so that the output is temperature independent [2].**

In 1968, David W. Allan developed the time-scale algorithm for the atomic-clock ensemble providing time for the USA at the National Bureau of Standards in Boulder, Colorado. German and Allan have modified that algorithm to operate with a quartz-crystal oscillator ensemble. In addition, they have developed an algorithm that uses the ensemble to determine the temperature coefficients for each of the oscillators (clocks) in the ensemble – called automatic temperature compensation (ATC). The output of the ATC feeds the time-scale ensemble algorithm, which has the property of providing an output better than the best clock, and even the worst clock enhances the performance. With this configuration, flicker-floors of parts in 10^{12} have been obtained with only a gradual deterioration in stability in the long-term. The ensemble output also compensates for long-term frequency drift. To date, the EQUATE timing stability has demonstrated timing errors of less than 100 microseconds for prediction intervals out to 10 days. The ensemble operates without an oven, with micro-degree temperature measurement precision, has been tested over a temperature range of -40 to +85° C and can reasonably operate at even higher temperatures. This R&D system is now in its bread-board phase, and efforts are in progress to turn it into a small, low-powered, timing product [3].

This quartz-crystal oscillator timing ensemble is an essential part of the EQUATE technology package, which provides both precise timing as well as six-axis inertial sensing. We have demonstrated excellent frequency stability during large dynamic motion and during large temperature excursions. The whole timing and navigation package should be small and require very little power – well suited for a GPS denied environment.

I. OVERVIEW

There is a broad strategic need for high-accuracy INS (Inertial Navigation System) and for a high-accuracy clock system for GPS-Denied Navigation. The EQUATE technology (Ensemble of Quartz-clocks Adapting To the Environment) is a fully integrated ensemble of multiple (typically 6 or 7) precision, small, low-power quartz-crystal ***velocity-sensing*** and ***timing*** oscillators. The associated algorithms and measurement system compensate for temperature, shock/acceleration effects, and long-term drift. A lab proof-of-concept demonstration has been achieved for the EQUATE technology.

The highly stable timing portion of the technology is the main message of this paper. Since the other aspects of the EQUATE technology are relevant to successful timing, they will be discussed as they tie to the timing portion. One of the significant features of EQUATE timing is that excellent quality timing can be obtained in a dynamic environment both as to movement as well as to temperature.

The salient features of the EQUATE technology are: (1) atomic-clock like timing performance without a temperature controlling oven. The piezoelectric property of quartz makes it a good candidate for an (inertial measurement unit) IMU sensor as the oscillation frequency shifts with acceleration. One can configure the most sensitive directions for an ensemble of oscillators and achieve six-axes inertial sensitivity – thus providing (2) IMU capability. (3) In a shock or vibration scenario the clocks are configured to minimize these effects. Since gravity is an acceleration vector, a properly configured ensemble (4) can determine the local vertical. Because time is the integral of frequency and velocity is the integral of acceleration, (5) time or phase variations correspond to velocity variations in the EQUATE package; thus eliminating one level of integration errors as is common in all other inertial navigation system (INS) packages. These relative-velocity measurements can be made in all six-axes with high precision – (6) including the sensing of zero velocity in some scenarios. Because the local vertical can be determined and with an azimuthal reference (7) EQUATE has no coordinate drift within the small angles of measurement uncertainty. (8) EQUATE is being designed to operate over the full range of military environmental conditions. And (9) EQUATE has very high reliability using ensemble technology,

which includes a self-calibration capability to remove frequency drift errors and mal-performing clocks, – giving it a significant level of robustness.

The EQUATE technology brings together many proven technologies in a unique way – giving it several unique dimensions; i.e. ovenless operation, determining the local vertical, adaptively improving its performance as it learns from environmental changes, and combining precise timing and inertial sensing into one package.

II. TECHNICAL RATIONALE AND APPROACH

The time of the clocks are measured synchronously every second with sub-picosecond resolution – thus reducing the measurement noise to a negligible level [4]. This allows the measurement of systematic effects quickly and accurately. Each crystal's *in situ* temperature is transduced by measuring the frequency of the B-mode with typical sensitivity of ~ 26 ppm/°C. This technique greatly reduces any hysteresis effects. These measurements are then used to remove the temperature effects on the C-mode frequency. An algorithm has been developed that provides an automatic and adaptive estimates of the temperature coefficients for each of the clocks. The net of this is a temperature independent output from each of the clocks -- providing the input for the clock ensemble algorithm, which produces EQUATE's timing output. The crystal gamma (acceleration-sensitivity) vectors can be accurately determined for each EQUATE ensemble member and these are used for the IMU calculations and for the removal of motional effects in order to maintain good timing performance. The oscillators 1–6 are arrayed, as an example, with sensitivity directions +X, –X, +Y, –Y, +Z, and –Z on each face of a cube. This configuration also has the advantage, that for shock and vibration, opposite oscillators have first order cancelation benefits.

Clock-ensemble algorithms have been around for decades. One of the authors wrote the algorithm that generated time for the National Bureau of Standards in 1968 (NBS was renamed NIST). This algorithm (denoted AT1) with improvements is still used to generate time for the civil-sector of the nation's community. This algorithm is particularly adaptable to be used with clocks which have flicker noise, which is commonly observed as a good model in quartz-crystal oscillators. There are several papers that discuss the NBS/NIST AT1 algorithm. [5]– [9].

The EQUATE ensemble uses the virtues of the AT1 algorithm, with additional unique features added for quartz-crystal oscillators. The virtues of the EQUATE algorithm include: (1) the software computed time is more stable than the best clock in the ensemble; (2) even the worst clock enhances the computed time; (3) the time or phase of every clock is continually known with respect to the ensemble time and to every other clock; (4) ensemble robustness and reliability are inherent since if any one of the clocks mal-performs, it is rejected and its time and/or frequency is recalibrated; (5) the ensemble is adaptive as each of the clocks may improve or degrade with time (adapts to the environmental situation); (6) is optimized for both short-term and long-term performance; (7) is designed to work with environmental changes without an oven and at low power; (8) it is designed to ameliorate the effects of shock and vibration

Figure 1. 6 Member Dual-Mode Oscillator Ensemble in Thermal Chamber with Aluminum Plate for Thermal Substrate and RF Shielding for Isolation.

while measuring acceleration; (9) because the phase or time residuals are known for each of the ensemble members, both translational and rotational velocities are transduced; (10) having the ability to sense zero acceleration and zero velocity in some scenarios, it can use this information to recalibrate bias drift – resulting in a $t^{1/2}$ position error dispersion rate; and (11) both the random and the systematic parameters for each of the clocks are known and adaptively updated with time so that the EQUATE algorithms can minimize perturbing effects and maximize its performance in providing accurate position, orientation -- avoiding coordinate drift -- and providing precise short and long-term timing.

III. EQUATE TIMING EXPERIMENTS

EQUATE's ovenless timing has been tested over the military specification range of -40° to +85° C with rubidium clock like performance. In this experiment, the clocks were all placed on the same thermal substrate (see Figures 1 and 2). The temperature was held at each of the extremes for two hours, and then it was ramped up/down at a 2° C/minute rate. Figures 3 and 4 show the B-mode and C-mode fractional frequencies, respectively, over an interval of 3 days. The temperature is effectively measured for each clock in the ensemble to a precision of about a microdegree. The frequencies are measured every second with a precision of

Figure 2. Detail showing oscillator packages, aluminum foil used to provide shielding and plastic used to improve thermal uniformity & reduce transients.

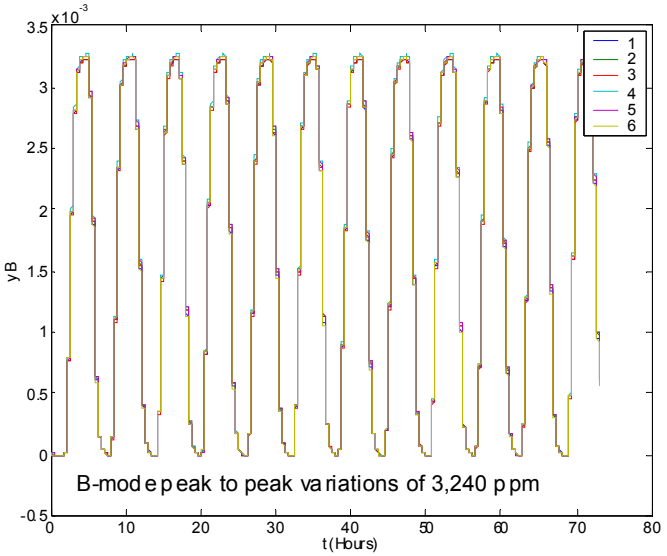


Figure 3. B-mode frequencies of the six clocks in the EQUATE ensemble.

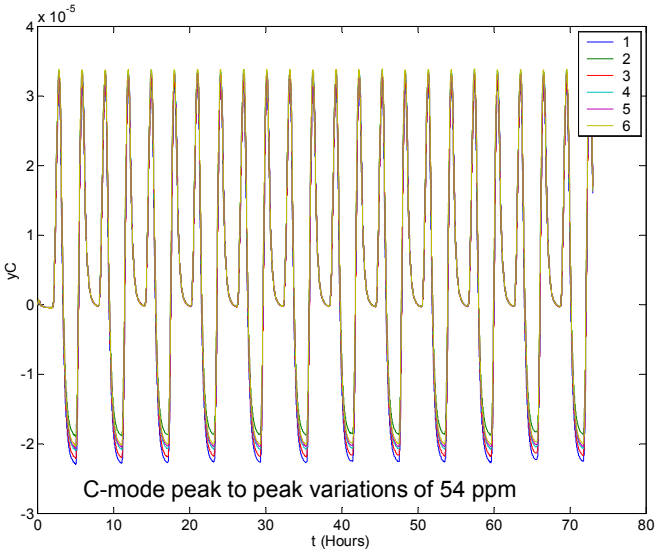


Figure 4. C-mode frequencies of the six clocks in the EQUATE ensemble.

about 1 x 10^{-12}. Figure 5 shows a plot of the fractional frequency output of the EQUATE timing algorithm. Figure 6 shows the EQUATE timing instabilities along with other well known frequency standards.

The Allan deviation instabilities (ADEV) shown in Figure 6 extend out for averaging times, τ, of the order of a day. The lower the value on the chart, the more stable the clock is. For example, if a clock has $\sigma_y(\tau) = 1 \text{ X } 10^{-11}$, then it has a time predictability of about 1 microsecond over a day. The stability comparisons in Figure 6 are like comparing apples, to oranges, to bananas because the EQUATE timing package is temperature modulated every six hours from -40° to +85° C, and the rubidium clock is in an ambient temperature environment and is phase-locked to GPS; the OCXO is free running and is in an ambient temperature environment.

So, Figure 6 demonstrates the excellent performance of the EQUATE timing package under very adverse temperature conditions, and EQUATE has the extremely valuable asset of needing no oven – in contrast to the rubidium and OCXO which both shield the temperature sensitive components in controlled ovens that limit their upper operating temperatures.

For these data, the EQUATE timing package was using an external reference to determine the temperature coefficients of the ensemble member clocks. The frequency drift has been subtracted from the EQUATE's timing package performance, as would happen in an application scenario in anticipation of moving into a self-calibration mode, when the reference is no longer available. The reference for all of the clocks in these stability plots is an HP 58503 GPS receiver – a highly stable ovenized quartz-crystal oscillator locked to the GPS atomic clock constellation.

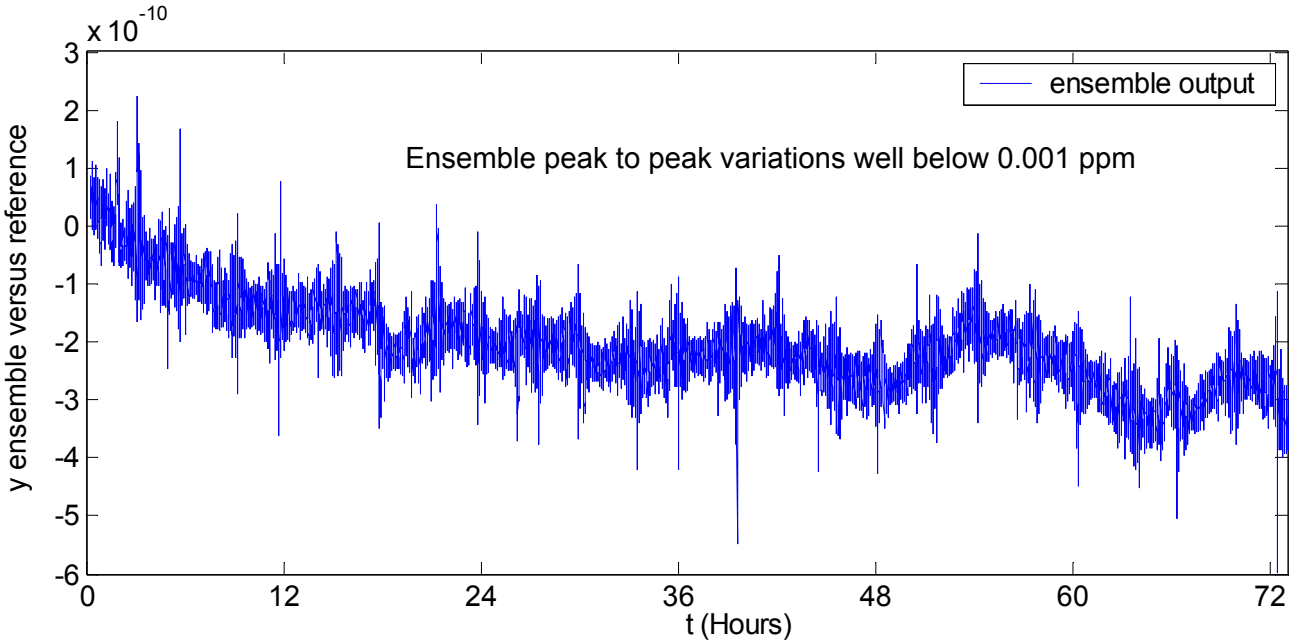


Figure 5. Fractional frequency output of the EQUATE timing ensemble.

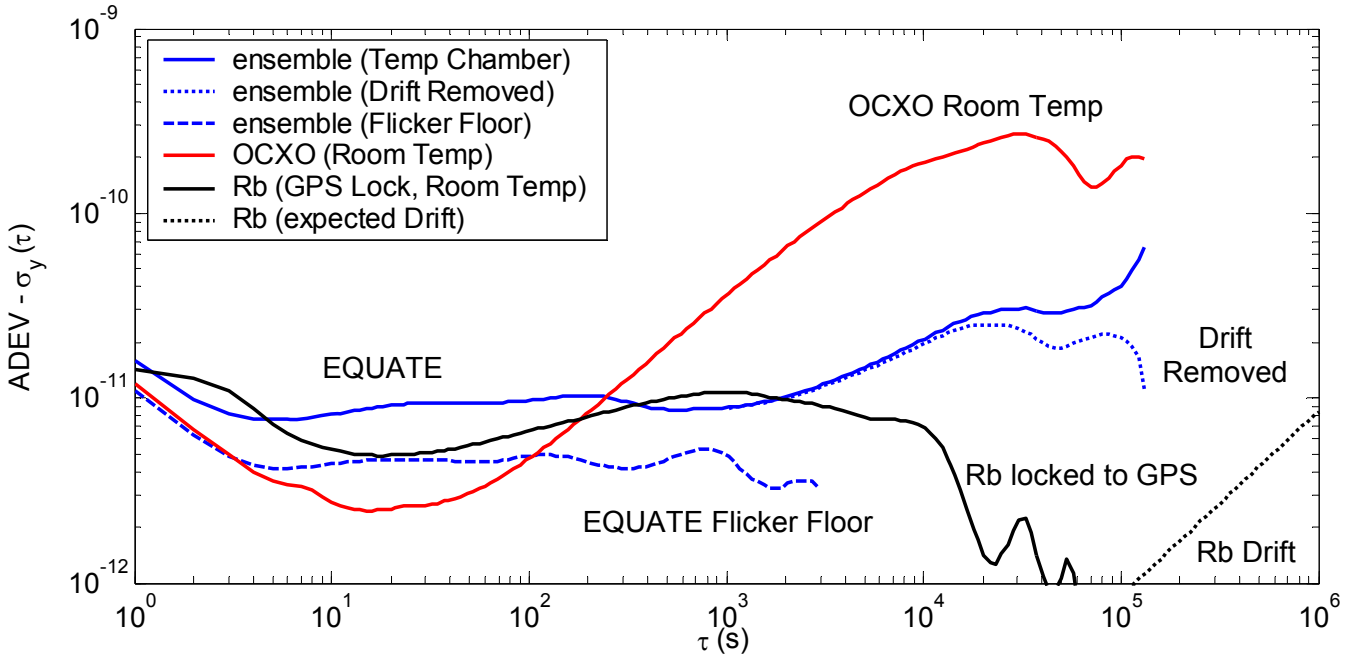


Figure 6. Frequency Stability plot for Ovenless EQUATE ensemble undergoing peak-to-peak temperature variations of 125° C (235° F) while the other clocks being compared are at ambient temperatures and the rubidium clock is phase locked to GPS.

OCXO Room Temp – The rise in the instabilities of the ovenized OCXO for larger averaging times is primarily due to diurnal temperature variations, which can be as large as 10 degrees C peak-to-peak; in comparison, the EQUATE timing package is going through a 125 degree C peak-to-peak oven cycle four times per day.

Rb locked to GPS (Room Temp) – The rise in the instabilities of the rubidium atomic clock – peaking at about 1,000 seconds averaging time – are the effects of the servo locking the rubidium to the GPS satellites' signals; then the fall off for larger averaging times is as its stabilities approach the GPS signal to which it is locking. The dotted line sloping up to the right is the typical instabilities introduced by a rubidium clock's frequency drift, when not locked to GPS.

EQUATE – The blue curves show the stabilities obtained for the EQUATE timing package. The flicker-floor dashed line which we believe from the theory will continue at this level of a few parts in 10^{12} for longer averaging times. Hence, we see that the EQUATE timing package can be as good or better than high quality commercial atomic clocks while consuming less than 1/1000th of the power and while going through a temperature range that is more than 10 times as large. The solid blue line instabilities are from EQUATE timing package data where some frequency drift was present along with frequency pops and jumps. The dotted blue line shows the effects of removing the frequency drift from these data in anticipation of self-calibrated drift.

EQUATE requires initial calibration. Once calibrated properly it can fly-wheel forward from its last known position, orientation and time. After calibration and when no external timing referencing is available, then EQUATE becomes its own source of precise and accurate time. Its atomic-clock timing stability is then available for the IMU portion of the EQUATE package. The blue curve in Figure 7 is an example of the long-term stability of the EQUATE timing package when it is self referenced while being modulated in temperature every six hours over 125 degrees C range. The black curve shows the EQUATE frequency stability when an external reference is available and the temperature is being modulated in the same way. In an application scenario, the black curve is the frequency stability obtained during the time when the EQUATE package is being calibrated. The blue curve is the long-term frequency stability when the external reference is removed and EQUATE becomes the primary reference. It is apparent that the self-referencing capability of the EQUATE timing package is a very important dimension of the EQUATE technology. These are the first results for self referencing. Theoretical indications are that the future holds improvements even over these excellent results.

Also, in Figure 7, for comparison in a similar temperature environment, the stability is plotted of an OCXO and of a commercial rubidium frequency standard, but over a reduced temperature range of -20° C to +65° C to conform with the specifications of the manufacturer of the rubidium clock. In this case, the OCXO and the rubidium clocks were subject to 40 C degrees less temperature variations than the EQUATE ensemble. In all three cases the temperature was held at the extremes for two hours and the ramp rates were two degrees per minute. The peak-to-peak phase or time modulation (x_{pp}) in an ADEV stability diagram is given by:

$$\sigma_y(\tau) = \frac{x_{pp}}{\tau}\, sin^2(\pi f_m \tau),$$

where f_m is the frequency of modulation. In these experiments, $1/f_m$ was about six hours. Some residual

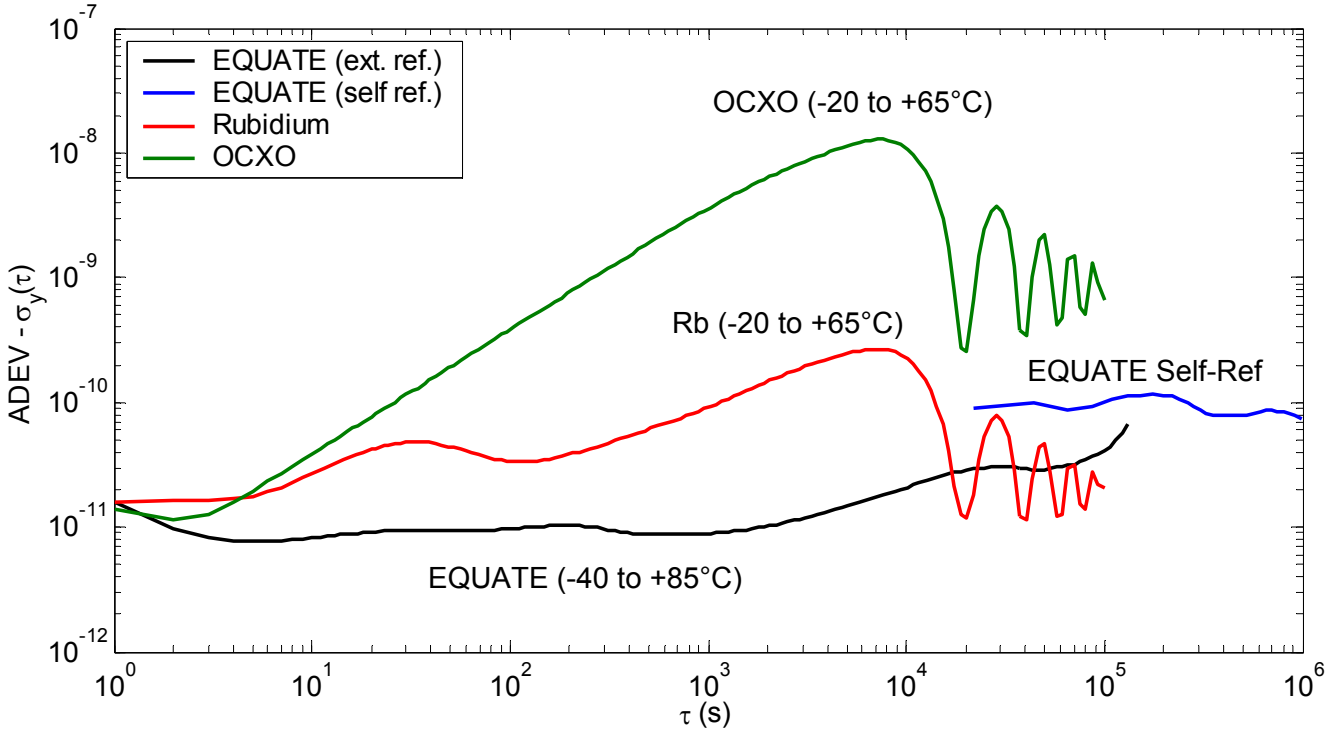


Figure 7. Frequency Stability plot with wide temperature modulation with about six hour periods: for ovenless EQUATE (-40° to +85° C), for a commercial rubidium, and an OCXO (-20° to +65° C)

temperature effects were present in the EQUATE output as can be seen by the hump at τ = 10,980 s (~3 hours) in the ADEV stability diagram shown in Figure 6. This equation assumes sine wave modulation. Using the data shown in Figures 6 and 7, x_{pp} was calculated for each of the three clocks, and under the assumption of sine wave modulation the estimated frequency dependence per degree C was 480, 9.2, and 0.63 parts in 10^{12}, respectively.

The time prediction error of a clock for flicker noise is given by $x_{pred}(\tau_p) = 1.2\ \tau_p\ \sigma_y(\tau_p)$, where τ_p is the prediction interval. The GPS Wing's desired prediction error in a GPS denied scenario is 100µs. Using this equation and the flicker floor level shown for the EQUATE timing ensemble indicates that the prediction error for EQUATE would be less than 100µs for prediction times out to ten days.

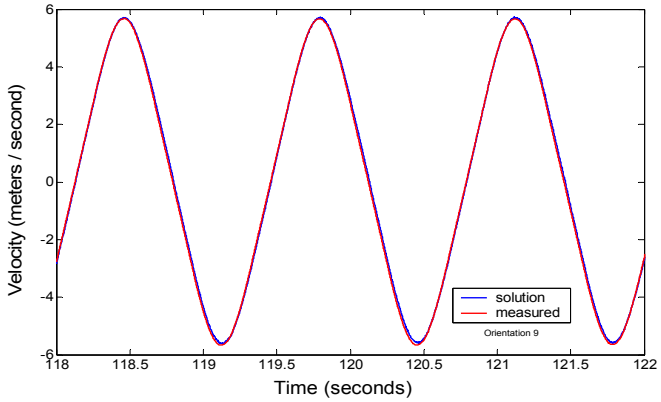


Figure 8. EQUATE measures velocity in six-axis pendulum configuration.

IV. EQUATE INERTIAL SENSING EXPERIMENTS

EQUATE is unique, as mentioned above, because it is both a precise timing package as well as providing inertial sensing capability. Most clocks shield against environmental temperature changes and against shock and vibration to get the best timing performance. Similarly, most INS packages may be adversely affected by these same environmental conditions. The EQUATE technology approaches these environmental perturbations very differently. Taking advantage of the piezoelectric effect and the high Q inherent in quartz – providing excellent short-term frequency stability – changes in temperature and motional effects are not shielded but utilized. The EQUATE algorithms measure the temperature with micro-degree precision; this precision is one reason for the success of the ovenless very stable timing approach. The piezoelectric effect in quartz with certain crystal cuts provides about a milli-g sensitivity to acceleration. Because time deviations correspond to velocity deviations, this gives significant advantage to this technology. So by precisely measuring and removing the effects of these environmental parameters gives the EQUATE technology the capability of having long-term frequency stability (excellent time predictability) as well as IMU capability. Hence, we see that temperature effects and motional effects are all integrated into the timing of the EQUATE technology.

The frequency stability of EQUATE has been measured under dynamic operation with a six-axis test bed. Six clocks were optimally configured around the perimeter of a light weight cube -- 20 cm on a side. This cube was placed on the

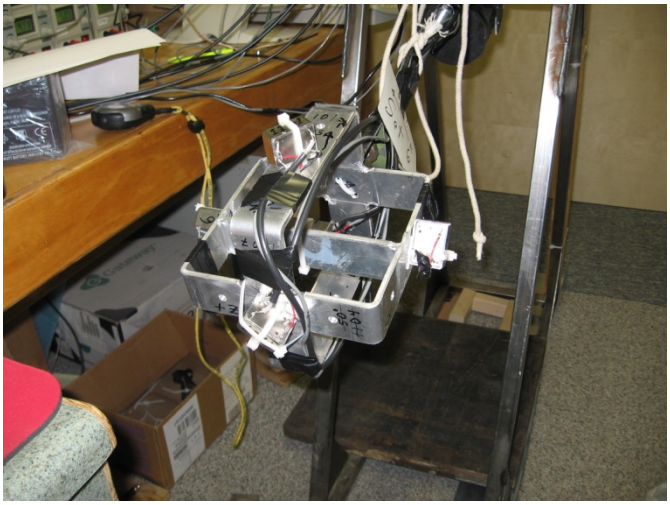


Figure 9. Cube /Pendulum Test Fixture for EQUATE Inertial Experiments.

end of a 60 cm rigid shaft. One end of the shaft was attached to a fulcrum and the other to the cube. Over 20 kilograms of lead were placed near the top of the shaft to give a whipping action at the end of the pendulum. The cube traversed about 5 meters during its period of 1.3 seconds and obtained accelerations of about 5 g's. The pivot point of the shaft was attached to a resolver to accurately measure the shaft angle. Our goal was to know the position of the quartz resonator in each clock in the cube to within 1 mm. The cube was made so that it can be rotated in any position on the end of the shaft to produce translation and rotation in all six axes. Clock frequency stabilities of parts in 10^{12} during these high dynamic conditions were measured, and the velocity was measured accurately as shown in Figure 8, and Figure 9 shows the cube/pendulum configuration.

The self-calibration capability of EQUATE leads to an *in situ* determination of each individual clocks temperature coefficients and a theoretical position-error dispersion rate going as $t^{1/2}$ in some scenarios. This dispersion rate was demonstrated for a linear tracking mode at a level of σ_s = 0.00078 meters $\tau^{1/2}$, where σ_s is the predicted error of position over a prediction time τ in seconds. Excellent timing stability was also demonstrated during this experiment when the effects of linear motion were subtracted. We are yet to perform the experiment combining high dynamics and large temperature variations. We have done each of them individually and obtained excellent timing stability, and the theory predicts that the EQUATE package should work well under these combined conditions also. That is yet to be proven.

V. CONCLUSIONS

The many unique features of quartz are the reason why over 2 billion quartz resonators are made per year. Yet its extremely high-Q properties have not been fully utilized. The EQUATE technology utilizes several of these unique properties to obtain frequency stabilities like an atomic clock, and to demonstrate the potential for inertial sensing as well. One of those properties is to remove coordinate drift, which plagues every other INS unit. EQUATE achieves the atomic-clock like, long-term frequency-stability timing performance by utilizing the outstanding high-Q, which readily yields excellent short-term stability, and then sensing systematic instabilities in an ensemble mode to remove these effects– thus obtaining excellent long-term stability as well.

Rather than shielding against environmental changes, EQUATE measures these environmental effects and removes them – adaptively improving its knowledge of these effects with time. The ovenless performance is one of the major benefits of this approach for lower system power. In the ensemble mode, EQUATE automatically determines the temperature coefficients for each of its member clocks and subtracts the effects due to temperature, which provides it with the unique characteristic of having temperature independence over extremely large temperature variations without an oven.

These temperature independent outputs from each of the clocks in the EQUATE ensemble are then fed to an NIST time-scale algorithm, uniquely modified to work with the features of quartz – rather than with atomic clocks – to generate outstanding long-term frequency stability performance. This translates to very small time prediction errors for long-term precise timing needs as well as providing an excellent timing reference for the IMU part of the EQUATE package. EQUATE's ovenless operation allows for reduced system size and power. The EQUATE technology approach has led to an improvement in frequency stability that is more than a hundred times better than the previous state-of-the-art, while also demonstrating the potential for INS capabilities with a non-drifting coordinate system. The EQUATE timing and INS package could be of major benefit in GPS denied environments as well as in a large number of other applications.

REFERENCES

[1] J.A. Kusters, M.C. Fischer, and J.G. Leach; Dual Mode Operation of Temperature and Stress Compensated Crystals; Proc. 1978 Frequency Control Symposium, 389-397.

[2] J.R. Vig; Temperature-insensitive dual-mode resonant sensors - a review; Sensors Journal, IEEE, June 2001, 62-68.

[3] D.W. Allan, J.A. Kusters, G.R. German; Adaptive Multi-Axis Sensor Array; U.S. Patent 7,472,032, Dec. 30, 2008.

[4] D.W. Allan and H. Daams; Picosecond Time Difference Measurement System; Proc. 1975 Frequency Control Symposium, 404-411.

[5] D.W. Allan, J.E. Gray, and H.E. Machlan; The National Bureau of Standards Atomic Time Scale: Generation, Stability, Accuracy and Accessibility; NBS Monograph 140; Ch. 9; 27 p.

[6] D.W. Allan, J.E. Gray, and H.E. Machlan; The National Bureau of Standards Atomic Time Scales: Generation, Dissemination, Stability, and Accuracy; IEEE Trans. Instrum. Meas. 21; 388-391.

[7] M.A. Weiss, D.W. Allan, and T.K. Peppler; A Study of the NBS Time Scale Algorithm; IEEE Trans. Instrum. Meas. 38; 631-635.

[8] D.W. Allan, M.A. Weiss, and T.K. Peppler; In Search of the Best Clock; IEEE Trans. Instrum. Meas. 38; 624-630.

[9] D.W. Allan; In Search of the Best Clock--An Update; Proc. 1988 Freq. Stand. Metrology Symp. 29-36.

Formation of a Paper Neural-fuzzy Time Scale in the Eastern Asia

Chia-Shu Liao, Fand-Dar Chu, Huang-Tien Lin, Yi-Jiun Huang
National Time and Frequency Standard Laboratory
Telecommunication Laboratories, Chunghwa Telecom
Taoyuan 326, Taiwan
csliao@cht.com.tw

Kun-Yuan Tu (2nd author), Yen-Wen Chung
Department of Computer Science and Information Engineering
Vanung University
Taoyuan 320, Taiwan
kytu@mail.vnu.edu.tw

Wei-Chih Hsu
Department of Computer and Communication Engineering
National Kaohsiung First University of Science and Technology
Kaohsiung 811, Taiwan.
weichih@nkfust.edu.tw

***Abstract*—This work presents a novel scheme to simulate a paper time scale by utilizing ensemble clock data of time laboratories in the eastern Asia-Pacific region. Based on integration of the above data with a neural-fuzzy prediction approach, the time difference with a local coordinated universal time is analyzed by a reference time scale devised in this work. Validity of the proposed scheme is investigated by undertaking a 250-day long neural-fuzzy predictor that is based on the calculated clock ensembles from Asian time laboratories. In the Asia-Pacific rim region, TWSTFT links from four laboratories, including the National Institute of Information and Communications Technology of Japan (NICT), the National Time Service Center of China (NTSC), the Korea Research Institute of Standards and Science of Korea (KRISS), and the Telecommunication Laboratories of Taiwan (TL), have formed a time transfer network based on use of multi-channel TWSTFT time transfer modems. Thus, a paper neural-fuzzy reference time scale by using the ensemble clock data among NICT, NTSC, KRISS, and TL in the Asia-Pacific region is simulated. Moreover, the performance of the proposed time scale is evaluated.**

I. INTRODUCTION

In the eastern Asia-Pacific rim region, TWSTFT links from four laboratories, including National Institute of Information and Communications Technology of Japan (NICT), National Time Service Center of China (NTSC), Korea Research Institute of Standards and Science of Korea (KRISS), and Telecommunication Laboratories of Taiwan (TL), have formed a time transfer network based on use of multi-channel time transfer modems as shown in Fig. 1. These four laboratories keep national standard time and are represented as local realizations of *UTC(lab)* in their countries. The Bureau International des Poids et Mesures (BIPM) Time Section coordinates the international time scale and publishes time difference results between *UTC* and *UTC(lab)* at the middle of the following month in Circular T [1]-[3]. However, to predict the real-time time differences for specific applications is necessary from the view of a particular region [4], [5] or time laboratories [6]-[8].

It is potentially to calculate a paper time scale in this region based on the clock difference information obtained from the aforementioned TWSTFT network and the BIPM ftp site. To provide a precise and reliable reference time source, this work attempts to devise a paper time scale in the eastern Asia-Pacific region with improved frequency stabilities by utilizing the neural-fuzzy prediction technique [9], [10].

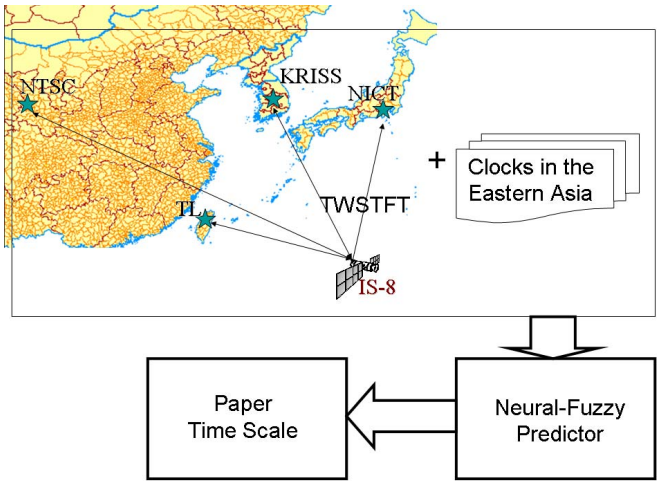


Fig. 1. Conceptual diagram of the proposed scheme.

This work was supported in part by the Bureau of Standards, Metrology & Inspection, Ministry of Economic Affairs, Taiwan under Grant 100-1403-05-05-13.

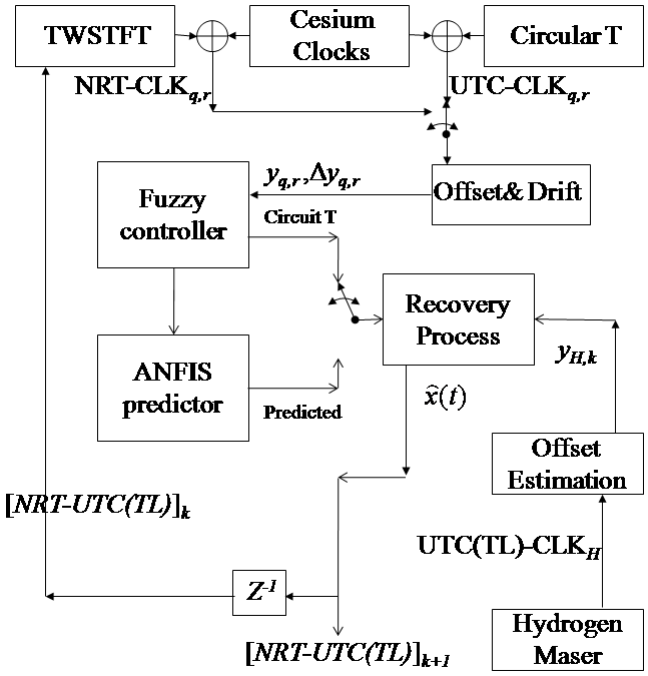


Fig. 2. System architecture of the proposed scheme.

II. PROPOSED SCHEME

Fig. 1 shows the conceptual block diagram of the proposed scheme. It consists of the TWSTFT network among the laboratories in the eastern Asia-Pacific rim region, atomic clock comparison data with respect to *UTC(lab)*, and neural-fuzzy predictor. Fig. 2 shows the system architecture diagram of the neural-fuzzy predictor applied in the proposed scheme. Furthermore, the frequency offset (y_k) and its change with respect to time (Δy_k) are fed into a fuzzy controller (as shown in Fig. 3) and an adaptive network-based fuzzy inference system (ANFIS) predictor (as shown in Fig. 4), respectively. It is worth noting that if the proposed scheme accompanies an active hydrogen maser and a microstepper as depicted in Fig. 2, it can be used as actualization of the simulated paper time scale to a physical clock.

A. *Fuzzy Controller*

In fuzzy controller, triangular-shaped membership functions (mfs) are chosen for the control input variables benefiting from their simple structure and computation. In this work, the input space is divided into five sets, i.e. negative big (NB), negative small (NS), zero (ZE), positive small (PS), and positive big (PB), for a frequency offset or its change. The sets are designed to be closer to the desired value of zero in order to provide precise control. The inputs to an if-then rule are the frequency offset (y_k) and its change (Δy_k). These rules can be expressed as

$$\textit{If } \{ y_k \textit{ is } \boldsymbol{Ar} \textit{ and } \Delta y_k \textit{ is } \boldsymbol{Bs}\}, \textit{ then } \{\boldsymbol{u} \textit{ is } \boldsymbol{Crs}\}, \quad (1)$$

where ***Ar, Bs***, and ***Crs*** denote the linguistic values defined by fuzzy sets on the universes of discourse.

The fuzzy-controller output is determined from the center-average method formulated as

$$\hat{y} = K_{out} \frac{\sum_{i=1}^{n} w_i u_i}{\sum_{i=1}^{n} u_i}, \quad (2)$$

where

K_{out} : output control gain;

$\hat{y}$: predicted output frequency offset;

w_i : grade of the *i*th output MF;

u_i : output label for the value contributed by the *i*th MF;

n : number of contributions from the rules.

B. *ANFIS predictor*

The ANFIS architecture used in this study is shown in Fig. 4. The number of membership functions (mf_i) assigned to each input of the ANFIS was arbitrarily set to 2, and so the rule number is 16. The ANFIS used here contains a total 104 fitting parameters, of which 24 are premise parameters and 80 are consequent parameters.

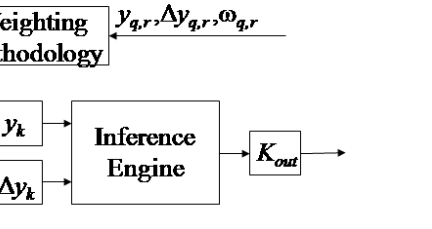


Fig. 3. Fuzzy controller architecture employed in this work.

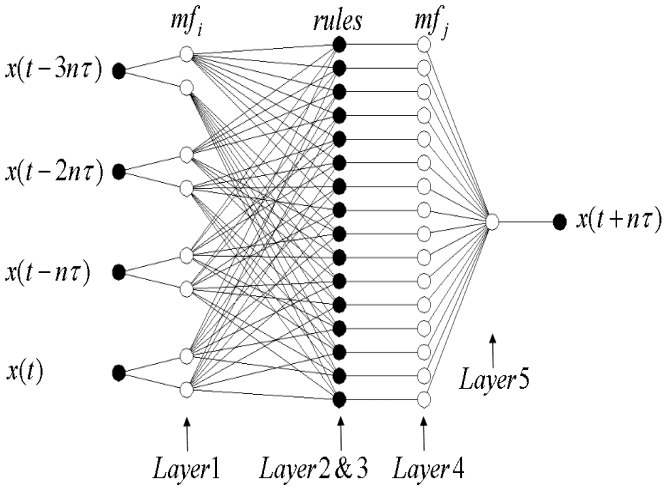


Fig. 4. ANFIS architecture employed in this work.

A Type-3 ANFIS employs a first-order, defined by the following rule base structure:

$$\text{if x is A and y is B, then } z = px + qy + r, \tag{3}$$

where A and B are fuzzy sets for the input and p, q, and r are constants. The node functions in Layer 1 are all bell-shaped with maximum equal to 1 and minimum equal to 0, such as

$$\mu_{A_i}(x) = \frac{1}{1 + [\left(\frac{x - c_i}{a_i}\right)^2]^{b_i}}, \tag{4}$$

where $\{a_i, b_i, c_i\}$ is the parameter set. Every node in Layer 2 multiplies the incoming signals and sends the product out. The output of this layer (w_i) is given by

$$\begin{aligned} w_i &= \mu_{A_i}(x(t-3n\tau)) \times \mu_{B_i}(x(t-2n\tau)) \\ &\times \mu_{C_i}(x(t-n\tau)) \times \mu_{D_i}(x(t)) \end{aligned}. \tag{5}$$

In Layer 2, each node output is then fed to Layer 3 to determine the individual rule's firing strengths as a ratio of all the rules firing strengths. For the nth node in this layer, the output of this layer (normalized firing strength) is given by

$$w_n = \frac{w_n}{w_1 + w_2 + \ldots w_i}. \tag{6}$$

Every node in Layer 4 multiple the normalized firing strengths by the consequent parameters:

$$\text{output} = w_i f_i \text{ , and } f_i = p_i x + q_i y + r_i, \tag{7}$$

where $\{p_i, q_i, r_i\}$ is the parameter set. The final layer (layer 5) which computes the weighted average of all the outputs to generate an overall output signal:

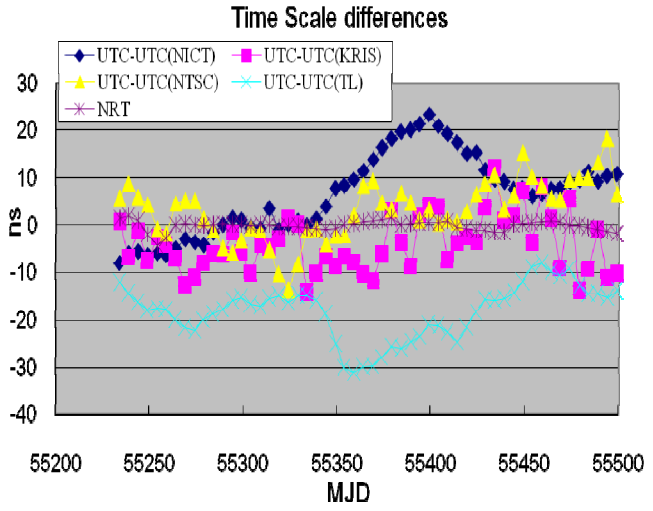


Fig. 5. Time scale differences of *UTC-UTC(TL), UTC-UTC(NICT), UTC-UTC(NTSC), UTC-UTC(KRIS)*, and *NRT*.

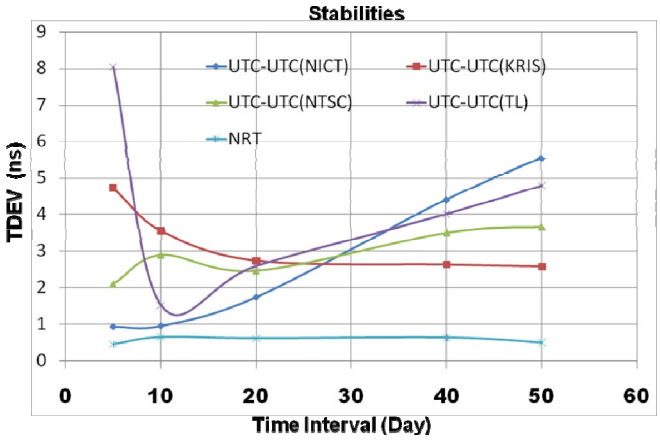


Fig. 6. Time stabilities of *UTC-UTC(TL), UTC-UTC(NICT), UTC-UTC(NTSC), UTC-UTC(KRIS)*, and *NRT*.

$$\text{Overall output} = \frac{\sum w_i f_i}{w_i} \tag{8}$$

III. Simulation Results

The neural-fuzzy reference time scale (*NRT*) was computed using the atomic clock and TWSTFT data from MJD=54959 to 55499 of time laboratories in the eastern Asia-Pacific region. For simplicity, 18 atomic clocks at NICT, 16 clocks at NTSC, 3 clocks at KRISS, and 11 clocks at TL, totally, 48 clocks were used for computation of our simulation. A 250-day long training set based on the aforementioned clock ensembles were used in the proposed neural-fuzzy predictor. In this work, we chose about five-month input-output data pairs of the following format:

$$\begin{aligned} &[x(t-3n\tau-m\tau), x(t-2n\tau-m\tau), x(t-n\tau-m\tau), x(t-m\tau) \\ &; \hat{x}(t+n\tau-m\tau)] \qquad \textit{and} \quad m = 0,1,2,3,4,5 \end{aligned}, \tag{9}$$

where

τ is equal to 5days;

n is equal to 6 for one month prediction;

$\hat{x}(t+n\tau-m\tau)$ is the predicted value.

Whenever the Circular T is published, equation (9) can predict NRT of the next month again.

Fig. 5 shows the time scale differences of the *UTC-UTC(k)*, and the predicted *NRT*, where *k* represents one of following laboratories: NICT, KRISS, NTSC, and TL. This figure shows a *UTC-UTC(TL)* time step jump from MJD=55339 to 55359, but our simulated *NRT* does not have such jump during the same period. This figure also shows a *UTC-UTC(NICT)* frequency change at MJD=55399, whereas our simulated *NRT* does not appear such frequency change at the same time. Besides, Fig. 5 shows the simulated time scale *NRT* is less than 5 ns.

Fig. 6 shows the time stabilities of *UTC-UTC(NICT), UTC-UTC(NTSC), UTC-UTC(KRIS), UTC-UTC(TL), and*

NRT. According to these results, the stabilities of *NRT* were the lowest (best). Quantitatively, the TDEV of *NRT* was lower than 1 ns for the whole time interval in Fig. 6.

IV. SUMMARY

This work has presented a novel scheme to generate a paper reference time scale by using the ensemble clock data among the TWSTFT laboratories such as NICT, NTSC, KRISS and TL in the eastern Asia-Pacific region. Experimental results show the performance of the simulated time scale *NRT* is less than 5 ns and the TDEV of *NRT* is lower than 1 ns.

ACKNOWLEDGMENT

Special thanks to NICT for the great job for the operation of the Asia-Pacific TWSTFT network. We would like to express our gratitude to NICT, NTSC, and KRISS colleagues for using their data and TL colleagues for their assistance.

REFERENCES

[1] BIPM, Circular-T. [Online]. Available: http://www.bipm.org/jsp/en/TimeFtp.jsp?TypePub=publication

[2] C. S. Liao, H. T. Lin, F. D. Chu, Y. J. Huang, K. Y. Tu and W. H. Tseng, "Forming a Real-time Time Scale with Asia-Pacific TWSTFT Network Data," in *Digest 2010 Conference on Precision Electromagnetic Measurements (CPEM)*, 2010, pp. 436-437, Daejeon, Korea, June 13-18, 2010.

[3] F. Arias, Z. Jiang, W. Lewandowski and G. Petitt "BIPM Comparison of Time Transfer Techniques," in *Proc. of 2005 joint IEEE IFCS and PTTI,* Vancouver, Canada, pp.312-315, August 2005.

[4] S. W. Lee, "Real-time Formation of a Time Scale Using GPS Carrier-phase Time Transfer Network," *Merologia*, vol. 46, pp. 693–703, 2009.

[5] K. Senior, P. Koppang and J. Ray, "Developing an IGS Time Scale," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* ,vol. 50, pp. 585-593, 2003.

[6] D. Matsakis, "Time and frequency activities at the U.S. naval observatory," in *Proc. IEEE Freq. Control Symp. Expo.*, pp. 217-224, Aug. 29-31, 2005.

[7] A. Bauch, D. Piester, and E. Staliuniene, "A new realization strategy for the time scale UTC(PTB)," in *Proc. IEEE Int. Ultrason. Ferroelectr. Freq. Control Joint 50th Anniv. Conf*, pp. 518-523, 2004.

[8] S. W. Lee, C. B. Lee, and S. H. Yang, "A New Approach for Steering UTC(KRIS)," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 4, pp.1247-1251, 2009.

[9] J. S. R. Jang, "ANFIS: Adaptive Network-based Fuzzy Inference System," *IEEE transactions on Systems*, *Man, and Cybernetics*, vol. 23, no. 3, pp. 665-685, June 1993.

[10] W. H. Hsu, K. Y. Tu, J. S. Wu, and C. S. Liao, "Frequency Calibration Based on the Adaptive Neural-Fuzzy Inference System," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 4, pp.1229-1233, 2009.

Satellite Orbit Determination and Time Transfer Based on TWSTFT

Wu Wenjun[1,2], Li Zhigang[1], Yang Xuhai[1]，Lei Hui[1,2]，Chen Xuan[1,2]，Feng Chugang[3]

1 National Time Service Center，Lintong, China
2 Graduate University of Chinese Academy of Sciences，Beijing, China
3 Shanghai Astronomical Observatory，Shanghai, China
Email：wuwj@ntsc.ac.cn

***Abstract*—Two-way satellite time and frequency transfer (TWSTFT) is one of the most precise and accurate long-distance time transfer techniques nowadays. A C band multi-station satellite tracking system is developed based on TWSTFT in China. This system can not only measure the distance between a satellite with transponders and earth stations, but also realize the time synchronization between all earth stations. The satellite orbit can be determined by the distances. Five earth stations have been constructed in China. The master station is placed at the National Time Service Center (NTSC) in Lintong. There are a very small antenna terminal (VSAT) of 3.7 meters, up/down converters, a modem and an atomic clock in each station. The up and down link frequencies are 6GHz and 4GHz .The spread spectrum signal generated by the modem is 20MHz chip rate .The transmitter power is less than 1W. The results show that the accuracy of ranging, orbit determination and time synchronization between all earth stations is about 1cm，10cm and 0.1ns respectively.**

I. INTRODUCTION

A GEO satellite is placed at an altitude of 36000km and it can be kept in the fixed position above the equator, so it is the important part of the Chinese satellite navigation system. But it is more difficult to determine the orbit of GEO navigation satellites with high accuracy, because the geometry of the satellite and ground tracking stations is poor and the satellite's movement with respect to the earth station is very small[1]. Two-way satellite time and frequency transfer has been proven to be the most appropriate means of comparing time-scales and atomic frequency standards nowadays[2]. A C band multi-station satellite tracking system, i.e., a satellite orbit determination system by transfer, is developed based on TWSTFT by the NTSC[3]. There are five earth stations in China, the major station is at Lintong. The modem with multi-channel receiver is used in this system. It can not only range the distances between the satellite and earth stations, but also make the time comparison between every two earth stations by TWSTFT.

II. RANGING AND ORBIT DETERMINATION

A. Ranging

Ranging principle for multi-station satellite tracking system is that precise time signals from atomic clocks of earth stations are modulated with different pseudo noise(PN) code sequences, then the signals are transmitted to the satellite through an up-converter and a 3.7 meters VAST at *Tt*. When the satellite receives the signals from earth stations, the transponder processes and retransmits the signals. The earth stations receive the signals from the satellite through the 3.7 meters VAST and down-converter at *Tr*. Then the modem demodulates the signals and

978-1-61284-110-6/11/$26.00 ©2011 IEEE

computes the signal transmission time delay *t*. Fig. 1 is the principle block diagram of satellite ranging.

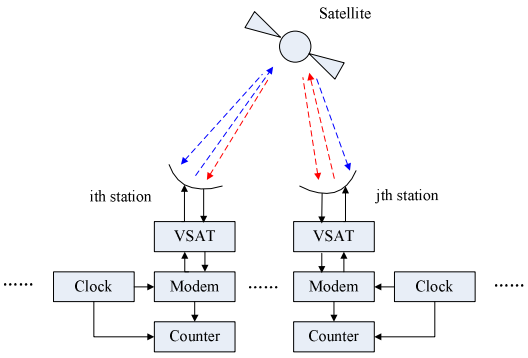


Figure 1. Principle of time delay of signals for both stations

The distances between the satellite and earth stations are determined by (1):

$$s = (c \times t) / 2 \qquad (1)$$

where: c is the speed of light

In the actual measurement, the time delay t includes ionosphere delay, troposphere delay, Sagnac effect, equipment delay and so on. These error delays must be subtracted from the delay t.

$$t=(T_r-T_t)/2+\Delta t_{ionosphere}+\Delta t_{troposphere}+\Delta t_{sagnac}+\Delta t_{transponder}+\Delta t_{equipment}+\varepsilon \qquad (2)$$

The ionosphere delay can be removed by ionosphere TEC models or dual-frequency method. The troposphere delay can be removed by troposphere empirical models[4]. The Sagnac effect can be canceled out by the Sagnac correction in the up/down links. The transponder delay can be measured before the satellite launched or solved as a parameter in the orbit determination. The equipment delay can be measured in real-time. ε is the measurement noise. As shown in Table I, the accuracy of raging is about 1 cm for the multi-station satellite tracking system.

Table Ⅰ. RANGE ACCURACY(1σ) FOR 5 STATIONS

station	accuracy(RMS)
1st station	6.0mm
2nd station	12.9mm
3rd station	6.1mm
4th station	8.3mm
5th station	12.8mm
Average of 5 stations	9.2mm

B. Satellite Orbit Determination

In the duration from 07 June 2005 to 09 June 2005, Sino-1 satellite was observed continually. The 12 sets of parameters, including 6 sets for satellite orbit parameters, 1 set for light pressure, 2 sets for experiential acceleration, 3 sets for transponder delays, were determined. The results show that the accuracy of orbit determination about 10cm. The Residuals (o-c) of 5 stations for orbit determination of Sino-1 satellite are as follows:

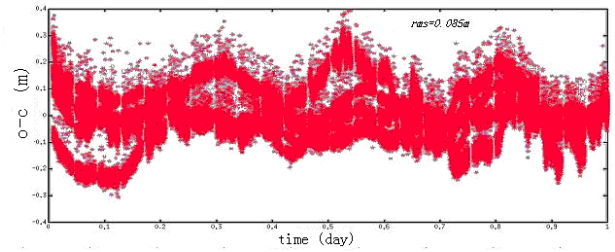


Figure 2. Residuals (o-c) of 5 stations for orbit determination of Sino-1 satellite in the duration of 7 June 2005, RMS=0.085m

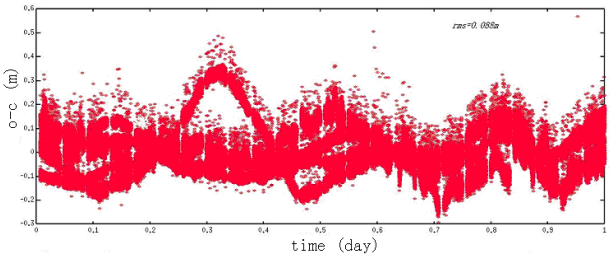


Figure 3. Residuals (o-c) of 5 stations for orbit determination of Sino-1 satellite in the duration of 8 June 2005, RMS=0.088m

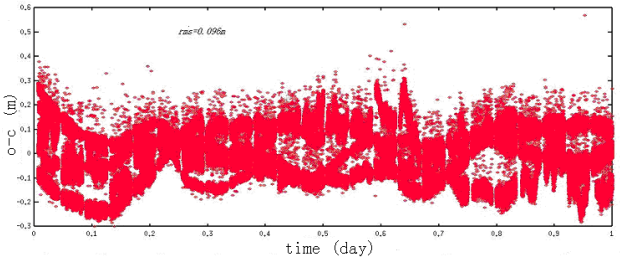


Figure 4. Residuals (o-c) of 5 stations for orbit determination of Sino-1 satellite in the duration of 9 June 2005, RMS=0.096m

III. TIME COMPARISON

The satellite tracking system is developed based on

TWSTFT. The multi-channel receiver is used in this system. It can not only range the distances between the satellite and earth stations, but also make the time comparison between every two earth stations. The time comparison is based on the exchange of timing signals through GEO between two earth stations. The time interval counter reading at station *i* is:

$$R_{ij}=T_j-T_i+\tau_i^U+\tau_j^D+\tau_s+\tau_i^T+\tau_j^R+\tau_j^{Sagnac} \quad (3)$$

$$R_{ji}=T_i-T_j+\tau_j^U+\tau_i^D+\tau_s+\tau_j^T+\tau_i^R+\tau_i^{Sagnac} \quad (4)$$

Where:

R_{ij}: the reading of the time interval counter at station *i* receiving signals from station *j*

T_i: local time-scale for station *i*

τ_i^U:signal path uplink delay for station *i*

τ_i^D: signal path downlink delay for station *i*

τ_s: satellite path delay through the transponder

τ_i^T: the equipment transmitter channel delay for the station *i*

τ_i^R: the equipment receiver channel delay for the station *i*

τ_i^{Sagnac}: Sagnac correction in the downlink

By subtracting (4) from (3), we get:

$$\Delta T_{ji}=1/2(R_{ji}-R_{ij})+1/2[(\tau_i^U+\tau_j^D)-(\tau_j^U+\tau_i^D)]+1/2[(\tau_j^R+\tau_i^T)-(\tau_j^T+\tau_i^R)]+1/2(\tau_i^{Sagnac}-\tau_j^{Sagnac}) \quad (5)$$

As shown in Fig. 5-7, the results of clock offset between the master station and slave stations by TWSTFT are given. The raw TWSTFT data had not been calibrated. The rubidium atomic clock is used in the 4th slave station. This station's clock offset is not shown.

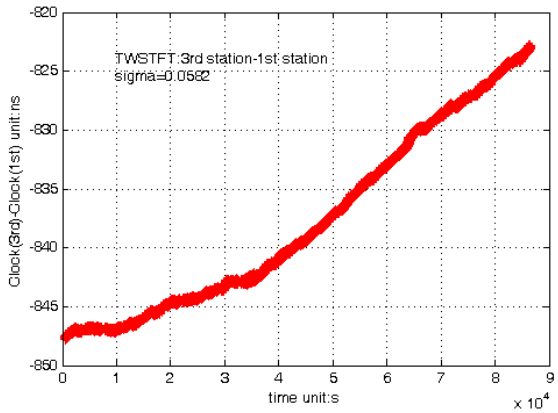


Figure 5. Results of clock(3rd)- clock(1st) in the duration of 7 June 2005

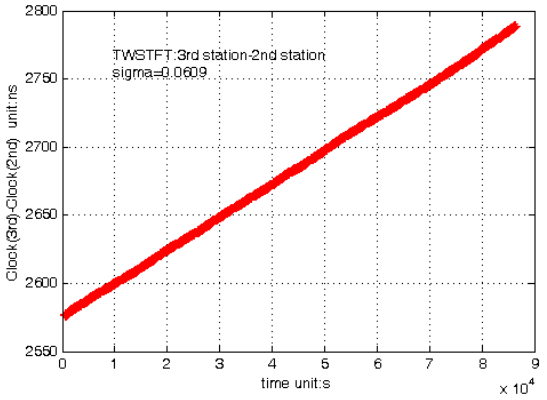


Figure 6. Results of clock(3rd)- clock(2nd) in the duration of 7 June 2005

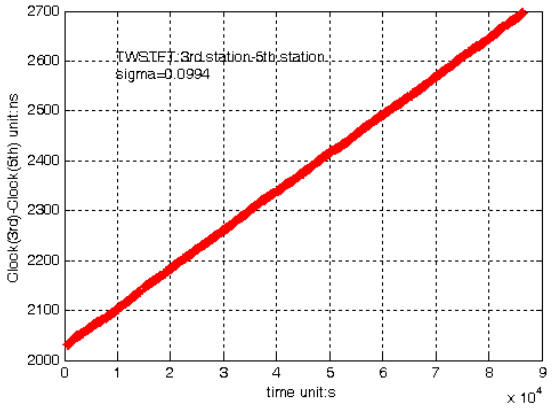


Figure 7. Results of clock(3rd)- clock(5th) in the duration of 7 June 2005

IV. CONCLUSION

The multi-station satellite tracking system based on TWSTFT can provide us with highly precise data of time and distance of GEO satellites with transponders. The orbit accuracy is improved greatly than other methods, e.g. The CCD and the United system for S Band (USB) technologies. This technology has been successfully used to the Chinese Area Positioning System, a satellite navigation system developed by China.

ACKNOWLEDGMENT

Supported by the National Basic Research and Development Program of China (Grant No.2007CB 815503), Chinese Academy of sciences national defense program(NO.CXJJ-09-M34) and CAS graduate science innovative funding research projects.

REFERENCES

[1] Scobal P R., Methods of Orbit Determination, Krieger Pulishing Company,USA,1976

[2] ITU Radiocommunication Sector, “The Operational Use of Two-Way Satellite Time and Frequency Transfer Employing PRN Code,”ITU-R TF. 1153-3 (Geneva), 2010

[3] Li ZhiGang,Yang XuHai,Ai GuoXiang,Shi HuLi,Qiao RongChuan, Feng ChuGang, “A New Method for Determination of Satellite Orbit by Transfer”, Sci. China Ser G-Phys Mech Astron, Vol. 52, No.3, 2009,pp.384-392

[4] Elliott D.Kaplan and C.J.Hegarty, Understanding GPS Principles and Applications, Second Edition, ARTECH HOUSE, INC, 2006, pp.308 -319

Time Transfer Using Frame Detection in Fiber-Optical Communication Networks: New Hardware

Kenneth Jaldehag, Sven-Christian Ebenhag, Carsten Rieck, and Per Olof Hedekvist
Measurement Technology, Communications Lab, Time and Frequency
SP Technical Research Institute of Sweden
Borås, Sweden
kenneth.jaldehag@sp.se

Abstract— **In this paper, a new, recently developed hardware for time transfer using passive listening and detection of SDH frame headers in fiber-optical networks is described. The method has been presented earlier, and results, using prototype equipment and an experimental fiber-link, have shown that time transfer with a precision of a few nanoseconds is possible over links with network distances exceeding 500 km. In order to further develop the method and make it available to regular users of time keeping equipment, it has been essential to minimize the space requirements of needed hardware and to make the implementation and installation more easily and straight forward.**

I. INTRODUCTION

The motivation of this work is to develop an alternative and complementary time transfer method using existing infrastructure, and to make it accessible to regular time and frequency users with precision and accuracy comparable to the commonly used satellite-based methods. It is well known that time transfer methods using global positioning systems, such as GPS (see e.g. [1] and references therein), are used in a variety of time and frequency applications. For applications requiring extended robustness and reliability, it is often necessary that complementary, backup methods are available with similar precision and accuracy. For instance, in international time metrology, two-way satellite time and frequency transfer (TWSTFT) (see e.g. [1] and references therein) using geostationary satellites, has been used for many years. Also, in recent years, time transfer methods using optical networks have been extensively studied and developed; see for example [2-4]. The fact that optical and radio-based transmission differ in signal frequency and infrastructure, make the combination of both types of methods much less vulnerable to intentional or unintentional disturbances.

As a contribution to the robustness and reliability of time and frequency transfer, SP Technical Research Institute of Sweden has developed a novel time and frequency transfer method based on passive listening and detection of SDH frame headers in fiber-optical networks. The method has been studied since 2005, see e.g. [4-6], and results, using several months of data and prototype hardware implemented in experimental fiber-links, have shown that time transfer with a precision of a few nanoseconds is achievable over links exceeding 500 km.

This paper describes a new hardware, which replaces the prototype equipment, which is essential for extended implementation and straight forward user installation. Section II will briefly review the method and its principles of operation. Section III goes through the details of the new hardware and its different components. Finally, Section IV lists some intended future work with discussions.

II. PRINCIPLES OF METHOD

The method is based on passive listening and detection of SDH frame headers, presently using an OC-192/STM-64 connection between core IP-routers at a nominal bit rate of 9953 Mbit/s, but is in practice with minor adjustments applicable to any STM line rate (or packet-based data transmission network).

SDH [7] defines the transmission of packets of data in nominally 125 µs long frames, where each frame starts with a well-defined sequence of A1 and A2 bytes that defines the beginning of a new frame, followed by frame information and finally the payload. The A1A2 sequence is chosen since it is extremely improbable that it occurs anywhere else in the bit stream and can therefore be used as a reference marker for the detection of the start of a new frame. At STM-64, this sequence is 192 A1 bytes followed by 192 A2 bytes.

In the time transfer setup, the reference marker is an electrical pulse which is generated at detections of a full A1A2 sequence. To succeed in an accurate time transfer, this operation must be performed both at the bit stream leaving the node, as well as the bit stream arriving at the node, i.e., in a two-way sense (see Fig. 1). The fundamental concept is to

This work is sponsored by the Swedish Post and Telecom Agency (PTS) under contract 04-9438.

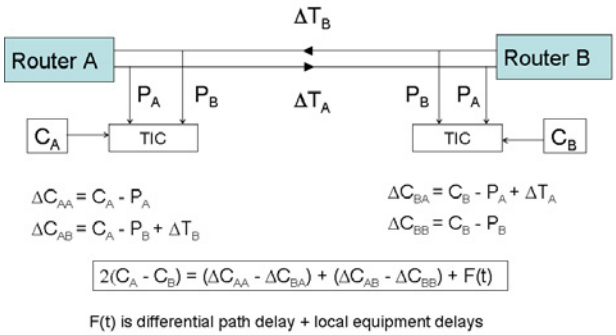


Figure 1. Principles of the two-way time transfer method. Optical signals are split-off to an O/E receiver. Frame headers are detected and pulses (P_A and P_B), based on the local oscillator in respective router, are generated. Time interval counters are used to relate pulses to local clocks (C_A and C_B) at both ends of a link. The two-way relation cancels most of the one-way delay variations leaving residual effects and a constant offset in F(t).

detect the time when this frame-start sequence is transmitted from a node in the network, in combination with the time when the same sequence arrives at the receiving node, where both time stamps are measured relative to the local clock at each node.

The transmission from a router is synchronized to a local oscillator (OCXO). The frequency stability of this clock source for the SDH framing do affect the performance of the time transfer by its jitter specification of less than 30 ps. This jitter is not notable in the present measurements; however, router oscillators could become a limiting component in future systems.

To implement the system, each fiber is equipped with two passive fiber-optic power splitters. At the transmitter, where the power level is high, 1% of the light is directed to the time analysis circuits, and at the receiving end, where the power level is low, 10% is split off to the circuits. The 11% added loss to the fiber transmission will decrease the power margin of the system, but it is anticipated that all systems are implemented with a far higher margin. All previous experiments have also validated this assumption.

III. Despcription of New Hardware

The recently developed hardware called Time Transfer Unit (TTU) is shown in Fig. 2. Each TTU contains custom-made printed circuit boards (PCBs) with different tasks and a single-board computer. The unit is 2U high, which is approximately 90 mm. The unit is equipped with two redundant fans attached to the back panel for necessary cooling. It will also work without the fans if a high enough air flow is available through the ventilation holes located in the front of the top- and bottom covers. Furthermore, with references to Fig. 3, which shows the modular design of the TTU, and Fig. 4, which shows more details of the internal functionalities, the following sub-sections describe each part of the new hardware.

A. Backplane

The cards are connected to slots on a backplane which supplies each card with necessary power. It distributes 1-pps

Figure 2. New, recently developed 2U (90 mm) hardware. Front panel and top view (without cover).

and 5/10-MHz signals from the oscillator card to the channel cards. It also serves the serial communication between each card and the single-board computer, necessary for example for measurement data, status bits, and Field Programmable Gate Array (FPGA) configuration updates.

B. Channel card

Up to 4 channel cards with 2 input channels each may be installed in the TTU. The incoming optical signal is converted to an electrical signal using an Avalanche Photo Diode (APD) receiver and is parallelized to 16 channels using a 1:16 DEMUX and then feed into an FPGA for analysis and frame header recognition (HR). The signal is thereafter serialized again in a 16:1 MUX and frame header pulses with a nominally frequency of 8-kHz are eventually generated based on the frame detection. The reason for parallelizing the signal is that the implemented FPGA is at its best working at 622 Mbit/s and the incoming signal is at 9953 Mbit/s. The alternative would have been to use an FPGA that could handle the fast bit stream without parallelizing, but at the time of development a faster FPGA was much more expensive and its jitter performance not yet analyzed.

Three PLLs are used in the setup, one for reconditioning of the incoming signal after the APD, which introduces jitter, and two PLLs in the FPGA on the Rx- and Tx-side, respectively. The two PLLs in the FPGA are essential in order to compensate for temperature variations in the FPGA.

The frame header pulses are used to stop a chip-scale 50-ps rms time interval counter (TIC) which is started by the 1-pps signals from the oscillator card. As explained later in sub-section G, several types of measurements are forwarded from the channel card to the serial bus. Each channel card contains two temperature sensors, one per channel. The front panel of the TTU contains one diode-lamp per channel to indicate

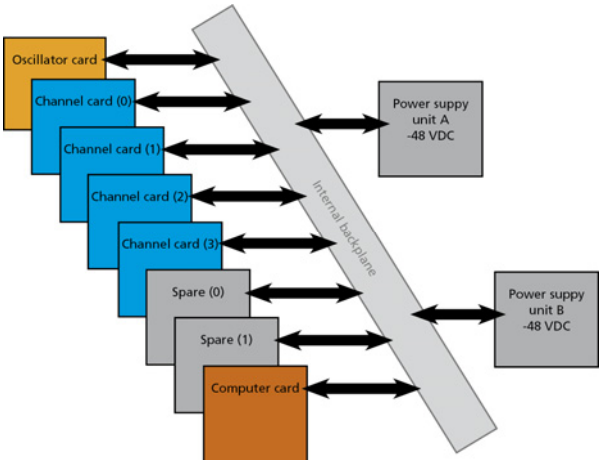


Figure 3. Principles of the modular TTU-design. A backplane connects each card powered from redundant -48 VDC power supply units.

synchronization to the incoming optical signal. It also contains outputs of the 8-kHz frame header pulses for testing and verification purposes.

C. *Oscillator card*

The oscillator card contains an OCXO which may serve as an intermediate clock bridging the asymmetry over multiple router links. At (end-user) sites, containing clocks or time scales to be compared, the connectors for external inputs located on the front panel are used to feed external 1-pps and 5/10-MHz signals to the TTU. The external signals are selected manually using jumpers located on the card.

Daisy-chain outputs (1-pps and 5/10-MHz) are available on the front panel, either, depending on the jumper settings, from the OCXO or the external clock. These may be used to connect TTUs at sites requiring more than 8 optical input channels. Two outputs (1-pps and 10-MHz) are also available on the front panel coming directly from the OCXO. Diode-lamps on the front panel indicate lock to the external clock signals. The OCXO may be steered in frequency and synchronized to the external 1-pps from the singe-board computer. A temperature sensor is available for monitoring.

D. *Single-board computer*

An Advanced RISC Machine (ARM) based single-board computer with a minimal GNU/Linux operating system is used to handle data collection, status requests, oscillator status and steering, as well as FPGA-updates. Application Programming Interfaces (APIs) have been written for these tasks. SSH (Secure SHell) data connections are available for remote login and data download. Ethernet, USB, and RS232 connectors are located on the front panel of the TTU. A Solid State Drive (SSD) is implemented large enough to store measurement data for a couple of weeks in case the Internet connection to the TTU temporarily goes down.

E. *Spare slots*

Two spare slots are available on the backplane for future

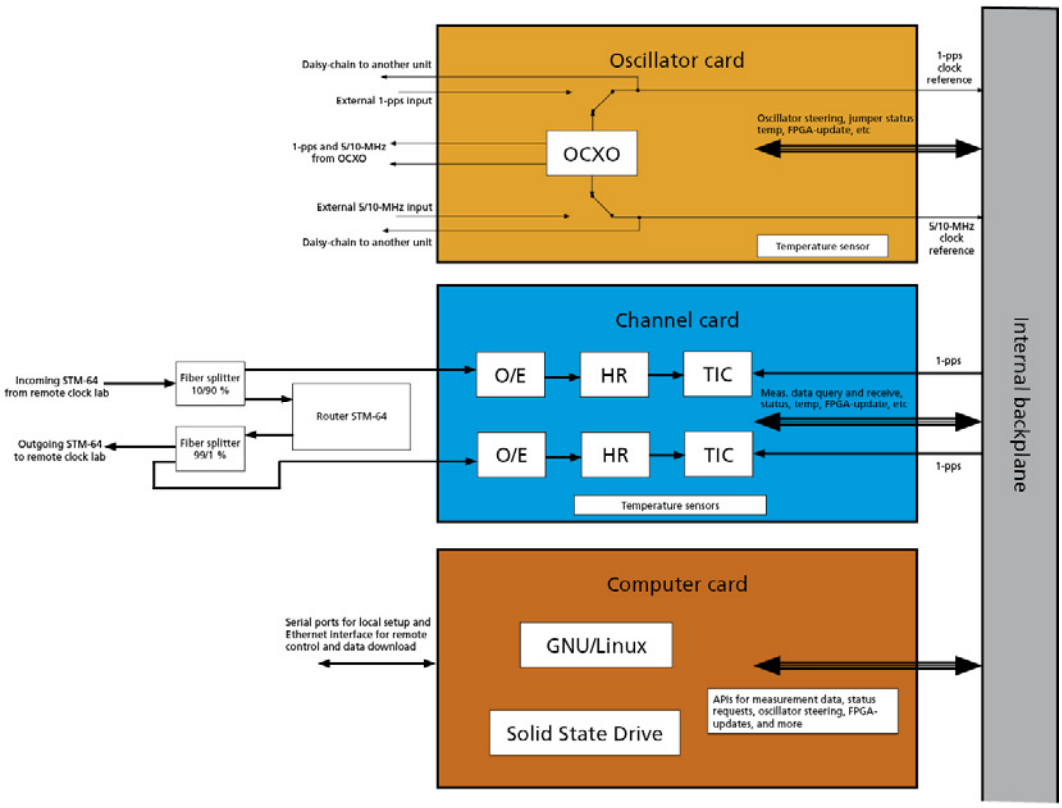


Figure 4. Details of the custom-made cards and the single-board computer. The SDH frame Header Recognition (HR) in the channel board contains DEMUX, FPGA, and a MUX.

implementation, such as for example bit-sequence generators and GNSS-receivers. Three out- or inputs are prepared on the front-panel to be used for this purpose.

F. Redundant -48 VDC power supply card

A power supply card containing two redundant -48 VDC supplies is installed in the TTU. The choice of -48 VDC is supported by the fact that many of the units will be installed in telecommunication facilities where -48 VDC is standard. Two diode-lamps on the front panel indicate the proper connection of power.

G. Measurement approach

To determine the time difference between two clocks (A and B) located at the two user ends of a two-way time transfer link, four time interval measurements are needed as shown in Fig. 1. In the prototype system [4], these measurements contained time intervals in the range of 0 to 125 μs as the TIC started on a 1-pps and stopped on a (SDH frame header) pulse train with a rate of 8 kHz.

In the new system, the approach is quite different and contains three measures. The first measure is a time interval measurement started with the 1-pps from the local clock and stopped by a pulse related to the detection of a frame header in the FPGA. In this way the principle is the same as the prototype system. However, the chip-scale TIC installed on the channel card can only measure time intervals in the range of up to 1.8 μs. The solution is to shift the frame header pulse with an integer number of a known time period (hereafter called HR-period) related to the optical bit-length so as the measured time interval is in the range of the TIC. The HR-period was chosen to be 12.86 ns which is presently 128 times the nominal bit-length (a choice also related to the capacity of the FPGA).

It is necessary to keep track of the number of HR-periods that has been added or subtracted. The second measure is thus the accumulation of HR-periods. The third measure that completes the measurement approach is a delay in the FPGA due to the parallelization of the incoming signal as mentioned above. This delay has a resolution of the nominal bit-length, which is roughly 100 ps in the present design.

IV. Future Work and Discussion

The largest error source for the method is presently asymmetric delay variations due to temperature variations in the two fiber-paths between routers. The size of the error is in the order of 2-4 ns peak-to-peak during a day. The different length of the Dispersion Compensation Modules (DSM) housed at intermediate amplifier sites in between routers, are the main cause for this [5]. In order to reduce this effect the temperature variations needs to be locally minimized or modeled.

As long as the mechanical lengths of the fibers in between routers are not changed, no phase change due to equipment will appear in the measurements described in Section III.G, even after restart or reboot of the TTU. This needs to be verified but means in principle that a constant asymmetry in the fiber path and local instrumental delays may be calibrated at installation, and the method could be used as a time transfer method as intended. A proper type of calibration method needs to be developed, for instance using a temporary or continuous GNSS-link.

One effect that will cause the need for a re-calibration, or otherwise compensated for, on certain fiber links, apart from mechanical fiber changes, is the loss of optical signal in between router paths. The receiving end could at these moments receive an optical signal not locked to the transmitting router oscillator which will cause a (slow) phase shift. Once the transmission is back to normal again the accumulated phase shift will remain and cause a phase jump in the link. This unwanted effect has not so far been investigated in detail.

In the near future, the intention is to establish time links between SP and different users. These links include both national and international users such as national metrology institutes, and these will first serve as test beds for the method and instrumentation, with the intention of becoming continuous time links.

The calculation of the time links using the prototype equipment has so far been in a post-processing mode for convenience. With the new TTU, a near real-time operation would be feasible in which measurement data is casted over the Internet to one or even several processing sites for redundancy.

Acknowledgment

The authors would like to acknowledge the continuous support from the Swedish University Computer Network (SUNET) and Peter Löthberg at STUPI AB. The authors would also like to thank InformAsic AB, Gothenburg, Sweden who made the theoretical ideas of the TTU realized in practice.

References

[1] J. Levine. “A review of time and frequency transfer methods”, Metrologia, vol. 45, pp. 162 – 174, 2008.

[2] M. Kihara, A. Imaoka, M. Imae, and K. Imamura, “Two-Way Time Transfer through 2.4 Gb/s Optical SDH Systems”, IEEE Trans. Instr. Meas., vol. 50, pp. 709 – 715, 2001.

[3] H. Schnatz, G. Grosche, O. Terra, T. Legero, B. Lipphardt, and S. Weyers, “A 900 km long optical fiber link for remote comparison of frequency standards” in this proceedings, 2011.

[4] R. Emardson, P. O. Hedekvist, M. Nilsson, S. C. Ebenhag, K. Jaldehag, P. Jarlemark, C. Rieck, J. Johansson, L. Pendrill, P. Löthberg and H. Nilsson, "Time Transfer by Passive Listening over a 10 Gb/s Optical Fiber", IEEE Trans. Instr. Meas., vol. 57, pp. 2495 – 2501, 2008.

[5] S. C. Ebenhag, P. O. Hedekvist, P. Jarlemark, R. Emardson, K. Jaldehag, C. Rieck, and P. Löthberg, "Meaurements and Error Sources in Time Transfer Using Asynchronous Fiber Network", IEEE Trans. Instr. Meas., vol. 59, pp. 1918 – 1924, 2009.

[6] K. Jaldehag, S. C. Ebenhag, P. O. Hedekvist, C. Rieck, and P. Löthberg, ”Time and Frequency Transfer Using Asynchronous Fiber-Optical Networks: Progress Report,” in Proc. of the 41st Annual Precise Time and Time Interval (PTTI) Meeting, Santa Ana Pueblo, New Mexico, USA, pp. 231 – 252, 2009.

[7] S. Bregni, “Clock Staibility Characterization and Measurement in Telecommunications”, IEEE Trans. Instr. Meas., vol. 46, pp. 1284 – 1294, 2008.

Implementation of Real-Time MTIE Assessment Method

Andrzej Dobrogowsk, Michal Kasznia
Chair of Telecommunication Systems and Optoelectronics
Poznan University of Technology
Poznan, Poland
dobrog@et.put.poznan.pl, mkasznia@et.put.poznan.pl

***Abstract*—In this paper the implementation of the method enabling computation of Maximum Time Interval Error in the real time is described. The results of implementation tests are presented and described.**

I. INTRODUCTION

The Maximum Time Interval Error (MTIE) is an important characteristic of timing signals in the telecommunication network [1, 2, 3]. MTIE recognizes maximal phase differences within considered observation intervals and therefore it is well suited for dimensioning of buffers (elastic stores) at the boundaries of the time scales. The point estimate of MTIE is commonly computed for a series of observation intervals using the sequence of time error samples measured at some network interface. The computation usually is performed off-line, after time error data measurement. However, in order to reduce the evaluation time of the timing signal, we can compute the MTIE in the real time, during the time error measurement process. Some methods enabling real-time MTIE computation were developed and tested by the authors of this paper [6, 7].

Computational ability of the methods proposed by the authors – DSDR method and EF method was analyzed and compared in some computational experiments performed off-line and simulating real-time performance [7, 8]. Both methods have presented very good performance for short observation intervals or time error data representing random behavior (e.g. WPM phase noise). However, for longer observation intervals and data series showing monotonic change of time error process the time of calculation was longer than sampling interval considered. Application of additional procedure controlling the computation progress allows to suspend the computation for some observation intervals and to make some toilsome operations later (for the successive sampling intervals), when smaller number of operations is executed, or at the latest, when the measurement is finished [9]. With this procedure applied the EF method has presented better performance during tests and therefore it was implemented for a real measurement and computation process.

In the paper the results of the implementation of the EF method enabling MTIE assessment in the real time are presented. To test this method in a real measurement process, a special time error measurement test module using FPGA technology was made. The results of the processes of measurement and calculation performed for different timing signals are presented.

II. MAXIMUM TIME INTERVAL ERROR

According to the international standards, the point estimate of the Maximum Time Interval Error is defined as the maximum peak-to-peak time error variation of a given timing signal, with respect to an ideal timing signal within a particular time period [1, 2, 3]. MTIE can be estimated from the formula

$$\hat{MTIE}(n\tau_0) = \max_{1\le k\le N-n}\left(\max_{k\le i\le k+n} x_i - \min_{k\le i\le k+n} x_i\right) \qquad (1)$$

where $\{x_i\}$ is a sequence of N samples of time error function $x(t)$ taken with sampling interval τ_0, $\tau=n\tau_0$ is an observation interval, and n can change from 1 till N–1 depending on the considered values of observation intervals.

MTIE point estimate is usually computed off-line, after the time error measurement process is completed. Following directly the formula (1) in order to find the estimate of MTIE for the observation interval τ, all intervals having the width of τ, existing within the sequence of N time error samples, must be reviewed. The window having the width of $\tau=n\tau_0$ and including n+1 samples is set at the beginning of the data sequence $\{x_i\}$ and then it is shifted with the step of τ_0 (one sample step) along the items of the sequence. For each window's location the peak-to-peak value of time error in the window is found. The maximum peak-to-peak value found for all existing locations of the window is the value of MTIE(τ) estimate. The process of the window reviewing does not depend on the data value. The complexity of calculation grows with n and therefore the direct method applied for

This work was supported by the Ministry of Science and Higher Education (project no. N N517 1645 33) and Poznan University of Technology (project no. 83-038/11 DS-PB).

computation of the parameter for the series of observation intervals is really time-consuming. The idea of direct search (plain computation) of MTIE is presented in Fig. 1.

Several time effective methods of MTIE assessment were proposed in order to avoid the time-consuming plain computation [4, 5]. Some different mechanisms were applied to reduce the complexity of computation. In the process of the MTIE search using the extreme fix (EF) method some window's locations are excluded from inspection if the peak-to-peak value for each of these locations is not greater than the value found until now, or if a greater peak-to-peak value may be found for the successive window's locations [5]. The EF method is based on fixing the positions of minimum and maximum samples for a given window's location. The general rule is that the next window's location is originated at the first extreme previously found. The reviewing process is performed only for the new samples that come into the window's location. The review of the whole window at its new location is performed only in the case of one-sample shift (the first extreme is first sample in the window's location), when the new sample appearing in the window is not a new extreme. The idea of EF method is presented in Fig. 2.

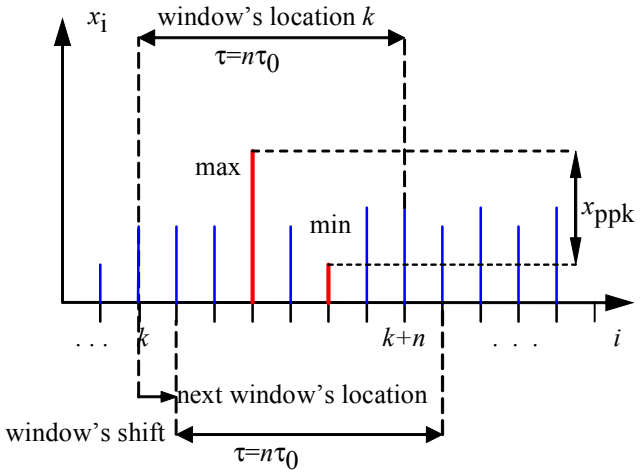


Figure 1. The idea of direct search for MTIE

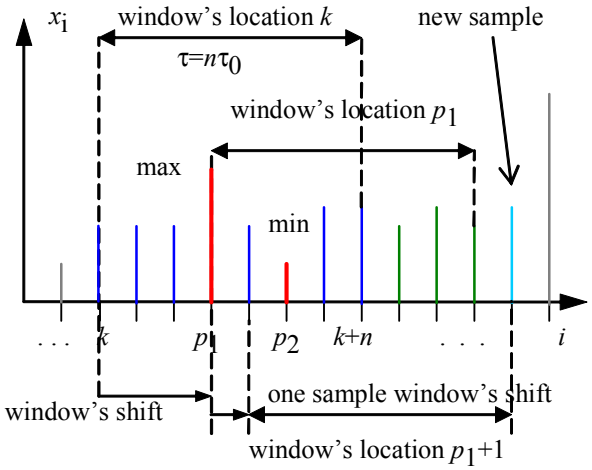


Figure 2. The windows shift in the extreme fix method

Another class of time effective MTIE computation methods contains the methods with sequential data reduction. The methods follow the suggestion according to which during the MTIE search process for some observation interval τ_i we find the extreme samples for some window's location from the set of extreme samples found previously during the MTIE search for the smaller observation interval τ_{i-1} ($\tau_i > \tau_{i-1}$). Therefore, we can reduce the number of time error samples used for the MTIE calculation. One of these methods, proposed by the authors of this paper – direct search with sequential data reduction (DSDR) – uses plain computation at each level of the procedure (for raw and reduced data) [6]. Other method, proposed by Bregni and Maccabruni [4], uses the rule of binary decomposition for reduction of time error data set used for computation.

III. Real-Time Methods of MTIE Assessment

Performing the MTIE computation in the real time, during time error measurement process, we can observe the current value of the parameter. Any possible wrong behavior of the analyzed signal (recognized if MTIE exceeds the limit) may enable applying proper activity of the network maintenance team.

In order to calculate the MTIE estimate simultaneously for several observation intervals in the real time, all necessary operations should be performed in the time period between two sampling instants, i.e. during the sampling interval τ_0. The ability of real time assessment depends on the several conditions, (e.g. number of observation intervals), but the computation method applied is very important.

The time effective methods of MTIE computation were adopted for the real-time computation by the authors of this paper. The results the experimental tests of the real-time MTIE computation using DSDR, EF, and binary decomposition methods were presented in [7, 8]. The experiments were performed in order to verify the maximum time spent for necessary operations within one sampling interval. The results of the experiments have shown very good performance of the method with binary decomposition, but this method has one obvious disadvantage – the number of samples in the observation intervals considered must be a power of 2. The performance of two other methods – EF and DSDR – was good for limited range of observation intervals – small number of intervals having short length, e.g. for the range from 0.1 s till 10 s. The maximum computation time for wider ranges of observation intervals (including longer observation intervals, e.g. between 100 s and 1000 s) has exceed the length of considered sampling interval for the EF and DSDR methods. Such situations result from the accumulation of computational procedures within one sampling interval caused by specific data arrangement.

In order to avoid these troublesome situations during the measurement and computation process performed in the real time, some procedures controlling the computation progress were introduced [9]. Using these procedures, the computation for a few longer observation intervals (belonging to the considered set of intervals) may be suspended in order to reduce the number of operations carried out within one

sampling interval. This analysis is suspended, if the number of samples reviewed in the course of the computation process performed within one sampling interval exceeds (or may exceed in the next step) some threshold value. The analysis of suspended intervals may be performed later (for the successive sampling intervals, in the case of DSDR method), when smaller number of operations is executed, or at the latest, when the measurement is finished (in the case of EF method). Such suspension causes a delay of the computation process, because the computations for the suspended observation intervals must be concluded after the measurement is finished.

The results of the experimental tests of the real-time MTIE computation using DSDR and EF methods with additional procedures were presented in [9, 11]. Application of the modifications proposed – control of computation progress with threshold – has enabled effective protection against exceeding the limit. The results of the tests have proved the ability of the real-time MTIE computation using both methods for the range of 21 observation intervals from 0.1 s till 1000 s performed during time error measurement with sampling interval τ_0=1/30 s. The comparison of the delay time (time of computation performed when the measurement is finished) for both methods is very interesting. Surprisingly enough, the best results were obtained for the EF method, which previously revealed difficulties in the real-time computation [7]. Introduction of the progress control causes the delay of obtaining the final result from some seconds (max 10 s) for modern computer with multi-core microprocessor up to about half a minute for older computer with single-core microprocessor. Application of DSDR method with progress control cause longer delay – from about 20 s using modern computer till almost 12 minutes using older computer. The delay obtained for the EF method is acceptable in comparison with considered time of measurement equal to T=4000 s. Therefore the EF method was implemented for the real measurement and computation process.

In the case of real-time calculation using EF method [7], the computation procedures for each observation interval run independently. Windows' locations (representing the observations intervals) are analyzed after filling out by the samples without waiting for the analysis by the preceding shorter windows. All windows are activated after filling out their first locations by the samples. The number of samples reviewed in the course of the computation process performed within one sampling interval is tracked. If the review of the whole window's location have to be performed for some observation interval (critical case for the EF method), and the number of samples reviewed within this sampling interval will exceed assumed threshold, then the computation for this observation interval is suspended until the end of measurement. The computation for the suspended observation intervals will be completed after the end of measurement. The example of this procedure is presented in Fig. 3 [9]. The extreme sample (8 at the position no. 10) leaves the current location of 10-sample window and the one-sample shift has to be done. Because the new sample coming to the window's location is not "more extreme" than the leaving sample – the review of the whole window at its new location is necessary. Therefore the analysis of the 10-sample window is suspended, because time of this review accumulated with the time of the operations done until now may exceed the length of sampling interval. The analysis of this window will be concluded after the end of time error measurement (Fig. 4) [9]. The algorithm of the real-time MTIE computation using EF methods is presented in Fig. 5.

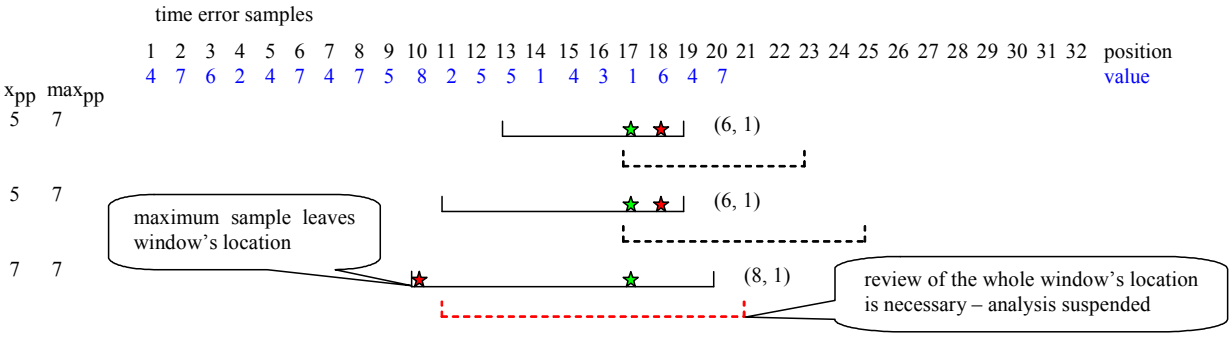


Figure 3. Suspension of the analysis for the real-time EF method

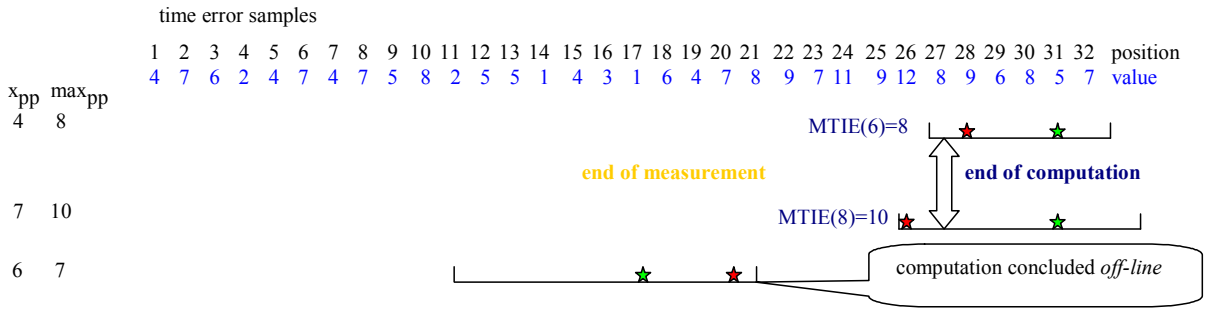


Figure 4. Continuation of the suspended analysis for the real-time EF method

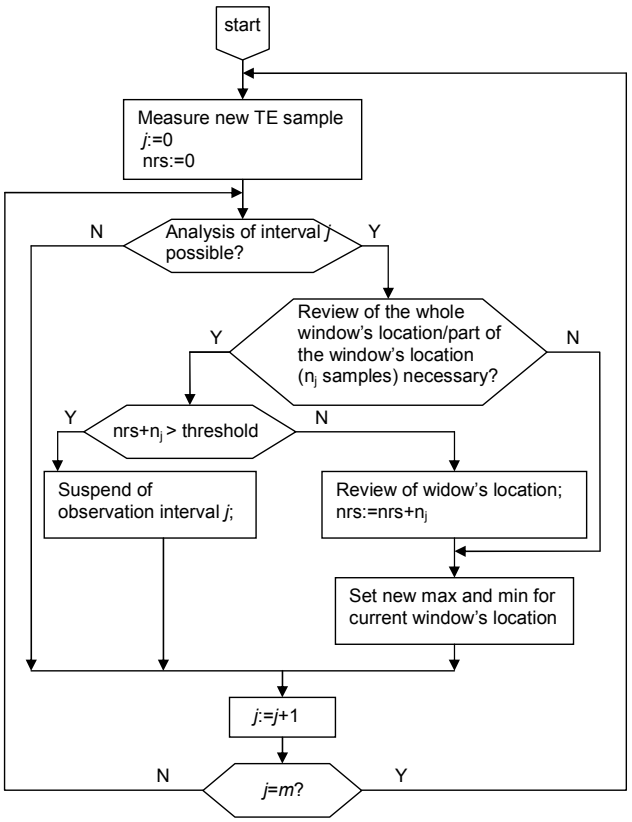


Figure 5. Algorithm of MTIE computation using EF method; j – current number of analyzed observation interval, m – total number of observation intervals, nrs – number of reviewed samples, n_j – number of samples reviewed for j-th interval

IV. Implementation of the Real-Time Computation Method

To test the real-time computation of MTIE in a real measurement process, a special time error measurement test module using FPGA technology was made [10]. This device, which allowed for single-channel time error measurements, was created using a Virtex-5 evaluation module. Data storage and real-time computation were performed using a personal computer connected by the USB interface. The real-time assessment of MTIE was performed using the EF method with an additional control of the computation process implemented especially for the test module. The values of MTIE were computed for 21 simultaneously considered observation intervals from the range of 0.1 s to 1000 s. Measurement and computation sessions were performed for independent oscillator configuration (comparison of two independent oscillators) and synchronized oscillator configuration (comparison of two GPS disciplined oscillators). The examples of the working process of the real-time MTIE assessment for independent oscillator measurement are presented in a form of screenshots in Fig. 6-8. Currently computed MTIE values for observation intervals of 0.1 s, 1 s, 10 s, 100 s, and 1000 s are displayed on the screen in the text mode.

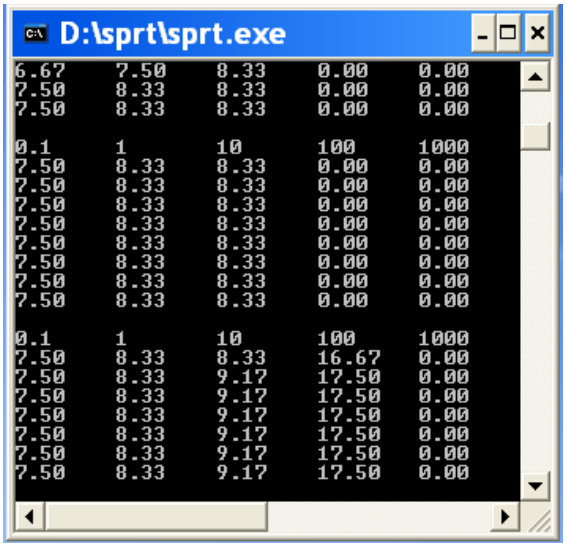


Figure 6. Early stage of the real-time MTIE computation

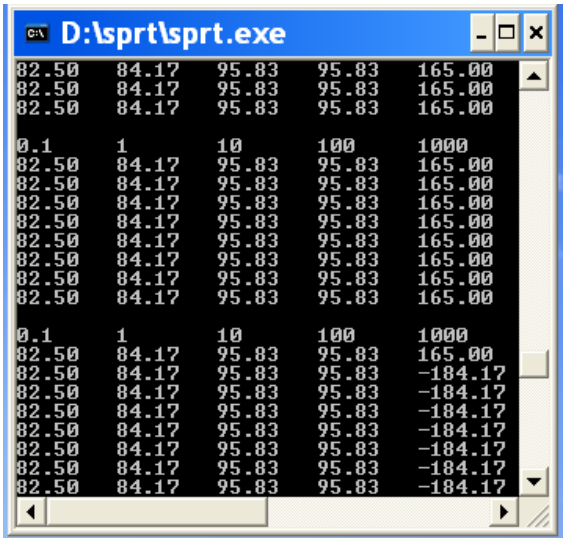


Figure 7. Late stage of the real-time MTIE computation

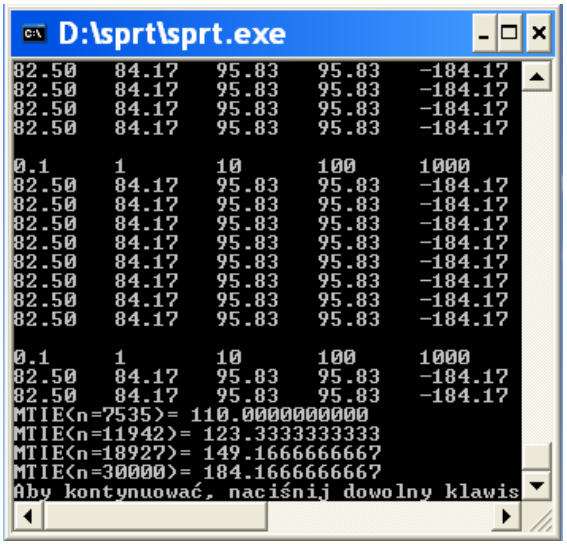


Figure 8. End of real-time MTIE computation

Zero value of MTIE displayed on the screen denotes that the computation for this observation interval does not started yet because of insufficient number of samples measured until now. Negative value of MTIE indicates the suspension of computation for particular observation interval. Threshold value was 12 000 samples. The results of the MTIE assessment successively obtained during the time error measurement process for independent oscillator are presented in the form of three-dimensional plot in Fig. 9. The first segments of the MTIE plot are shorter than the others because the parameters' values for longer observation intervals were not computed for these instants of the time error measurement process. Last segment represents final MTIE curve obtained at the end of the measurement. The results of time error measurement for independent oscillator are presented in Fig. 10. The computations of four observation intervals were suspended during this measurement session (time of measurement T=4000 s).

The examples of the working process of the real-time MTIE assessment for synchronized oscillator measurement are presented in a form of screenshots in Fig. 11-13. The results of the MTIE assessment successively obtained during the time error measurement process for synchronized oscillator are presented in the form of three-dimensional plot in Fig. 14. The results of this time error measurement are presented in Fig. 15. The computations of two observation intervals were suspended during this measurement session (time of measurement T=4000 s).

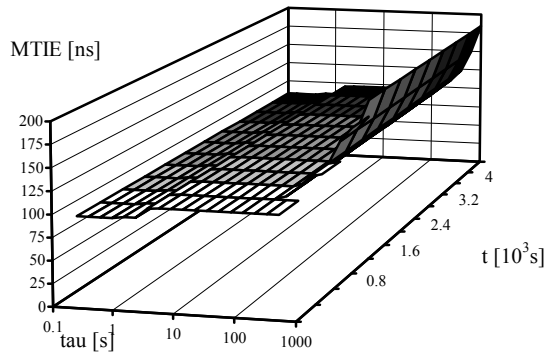


Figure 9. MTIE values successively obtained for independent oscillator measurement

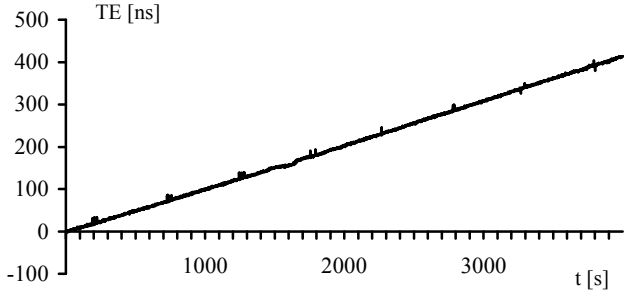


Figure 10. Time error obtained for independent oscillator measurement

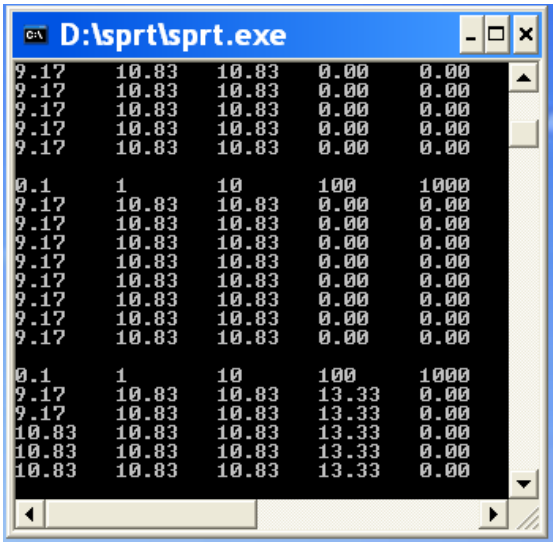


Figure 11. Early stage of the real-time MTIE computation

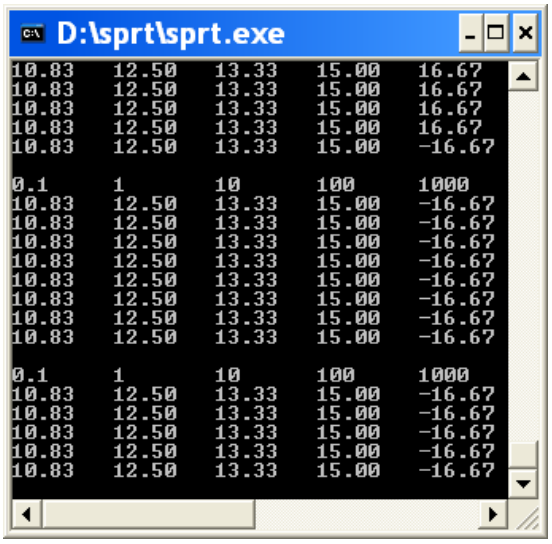


Figure 12. Late stage of the real-time MTIE computation

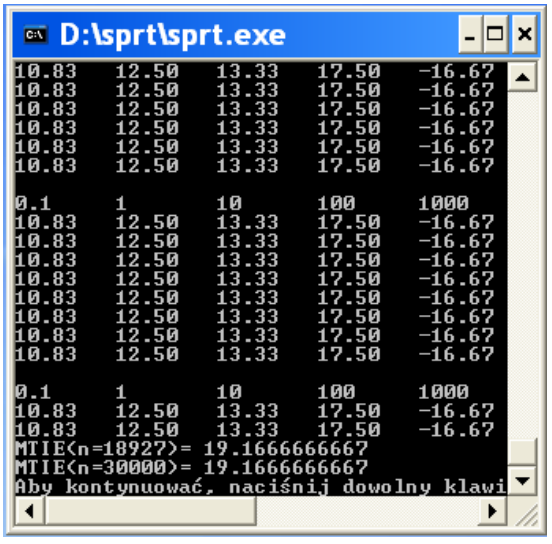


Figure 13. End of real-time MTIE computation

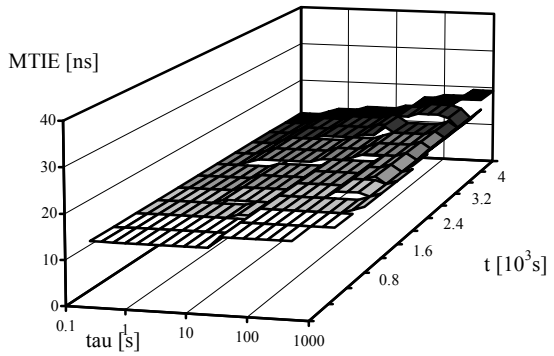


Figure 14. MTIE values successively obtained for synchronized oscillator measurement

Figure 15. Time error obtained for synchronized oscillator measurement

V. CONCLUSIONS

The method applied allows to perform the MTIE computation in the real time, during the time error measurement. Application of the modifications proposed enables effective protection against exceeding the limit of computation time. The results of the tests performed have proved the ability of the real-time MTIE computation for the range of observation intervals from 0.1 s till 1000 s performed during time error measurement with sampling interval τ_0=1/30 s. The method applied shows better performance for the time error measurement in the configuration of synchronized oscillator.

In the future the method presented should be implemented in a multifunctional measuring system enabling evaluation of the synchronization signals' quality in the telecommunication network [10]. The implementation of the EF method using multi-task processing and multi-core microprocessor could enable the real-time MTIE assessment without the suspension of computations for particular observation intervals.

REFERENCES

[1] ETSI EN 300 462, "Generic requirements for synchronization networks" (1998).

[2] ITU-T Rec. G.810, "Considerations on timing and synchronization issues" (08/96).

[3] ANSI T1.101-1999, Synchronization interface standard.

[4] S. Bregni, S. Maccabruni, "Fast computation of Maximum Time Interval Error by binary decomposition," IEEE Trans. Instrum. Meas., vol. 49, No. 6, pp. 1240-1244, Dec. 2000.

[5] A. Dobrogowski, M. Kasznia, "Time effective methods of calculation of Maximum Time Interval Error," IEEE Trans. Instrum. Meas., vol. 50, No. 3, pp. 732-741, June 2001.

[6] A. Dobrogowski, M. Kasznia, "Some problems and their solutions in quasi-parallel MTIE assessment," Proc. of the 2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control Joint 50th Anniversary Conference, pp. 494-499, Montreal, 23-27 August 2004.

[7] A. Dobrogowski, M. Kasznia, "Testing of the methods of real-time MTIE calculation," Proc. of the 2005 Joint IEEE Frequency Control Symposium and Precise Time and Time Interval Systems Application Meeting, pp. 397-403, Vancouver, 29-31 August 2005.

[8] A. Dobrogowski, M. Kasznia, "On-line computation of MTIE using binary decomposition and direct search with sequential data reducing," Proc. of the 2007 IEEE International Frequency Control Symposium Jointly with the 21st European Frequency and Time Forum, pp. 877-882, Geneva, 29 May – 01 June 2007.

[9] A. Dobrogowski, M. Kasznia, "Real-time MTIE Assessment with Flexible Control of Computation Process," Proc. of European Frequency and Time Forum and IEEE Frequency Control Symposium Joint Conference EFTF'09 IEEE-FCS'09, pp. 1102-1107, Besancon, France, 20-24 April 2009.

[10] A. Dobrogowski, M. Jessa, M. Kasznia, K. Lange, M. Jaworski, "Hardware and Software Realization of Time Error Measurements with Real-Time Assessment of ADEV, TDEV, and MTIE," Proc. of 24th European Frequency and Time Forum, Noordwijk, Netherlands, 13-16 April 2010.

[11] M. Kasznia, A. Dobrogowski, M. Jessa, K. Lange, "Advanced methods of synchronization signals' quality evaluation," report of the project no. N N517 1645 33, unpublished.

Generation of 1-pps Timing Signal Controlled by NTP

Andrzej Dobrogowski, Mieczyslaw Jessa, Michal Kasznia, Krzysztof Lange
Chair of Telecommunication Systems and Optoelectronics
Poznan University of Technology
Poznan, Poland
{dobrog, mjessa, mkasznia, lange}@et.put.poznan.pl

***Abstract*—This paper deals with the problem of generation of 1-pps timing signal at the output of computer interface controlled by Network Time Protocol. The methods of synchronization of internal computer clock and transmission of the timing signal to the serial interface are described. The results of the measurement of obtained signals are presented.**

I. INTRODUCTION

In the world facing the technological development, the demand for sources of time signals has been still increasing. Currently, numerous systems which to a high extent meet the needs of the consumers of time signals in every respect are offered. The basic sources of time signal are radio systems, for example DCF-77, information systems, i.e. NTP, and satellite systems like GPS. The offered quality of time signal sources falls into the range of tens of milliseconds for radio sources through hundreds of microseconds for NTP to tens of nanoseconds for satellite systems. It seems that the time signal from the Internet has the broadest application because it is easy to acquire and includes a considerable potential for improving its parameters [5, 6, 8].

The research presented hereby is, therefore, aimed at the application of NTP protocol to generate the timing signal, with a possibility of utilizing the signal outside the Internet network. It was also assumed that the source of signal would be an ordinary computer with an internal quartz clock. This is an important assumption because the utilization of rubidium clock in the NTP receiver would remove a majority of quality problems affecting the acquired time signal, at the expense, however, of a considerable increase in the device price.

Generation of timing signal controlled by NTP could be useful in order to provide some systems with timing signal when we haven't access to good quality timing signal source like atomic clock or GPS controlled oscillator. The authors of this paper have proposed a method of generation of physical timing signal controlled by the NTP [1, 2].

In this paper the method enabling generation of physical timing signal (in the form of 1-pps pulses), controlled by the NTP and brought out using serial port of the computer, is described. The algorithm of control of timing signal using NTP protocol is presented and described. The timing signals obtained were compared with the reference signal controlled by GPS receiver.

II. SOFTWARE AND HARDWARE PLATFORMS OF RESEARCH

The research carried out [1, 2, 3] reveal two basic factors affecting the quality of time acquired from the NTP protocol for which the very NTP system is responsible. These are: the time of servicing a query in the server and the waiting time of the query in the Internet network. Due to the dynamics of phenomena taking place in the network, the compensation of these times by the NTP mechanisms has quality limitations. But when the time signal is led outside computer, the factors resulting from the performance of the computer operating system as well as the factors related to the quality of internal computer clock have an additional influence on the quality of the generated signal. The research carried out proved considerable program difficulties that developed when the MS Windows operating system was utilized to generate the time signal on a computer interface without delays [2]. The computer clock was indeed successfully synchronized by the NTP protocol in time intervals in a range between 2 s and arbitrarily long time intervals. However, difficulties developed in applying the program responsible for transfer the information of time instant from computer internal clock onto 1-pps (NTP second) pulses, which should be transmitted to an external computer interface, like COM or USB. This problem was solved in the Linux operating system. Therefore, in order to generate the timing signal controlled by NTP, this operating system was applied.

The influence of system clock on the parameters of the acquired time signal is significant. Application of a very good clock (with small frequency instability, e.g. rubidium generator) will allow us to refer quite rarely to the synchronizing role of the NTP protocol. Therefore the number of clock adjustments with the potentially occurring noise from the Internet could be smaller. Common oscillators applied in the personal computers feature with frequency inaccuracy and sensitivity to temperature variations. The observations of the performance of computer clocks on different computers

This work was supported by the National Centre for Research and Development in the frame of the project PBZ-MNiSW-02/II/2007 during the years 2007-2011.

indicate that the time error caused by the frequency offset may take the value from 10 μs up to 130 μs for 1 second.

III. GENERATING 1-PPS SIGNAL

A. Setting the computer time from NTP

The "ntpdate" program was applied for setting the NTP time. This is a Linux program operating independently, without any external configuration or system services. The "ntpdate" program is included in a standard NTP packet for time synchronization in the Linux and Windows systems. This application uses a script that makes use of the "ntpdate" command [4] to synchronize the time:

```
#!/bin/bash
interval=$1            # synchronization interval
time_server=$2         # name of server
while :                # infinite loop
do
ntpdate -b $time_server       # synchronize time
with a given server
sleep $interval                # hold script
execution for a given time
done
```

The script call takes the form:

```
./sync.sh vega.cbk.poznan.pl 10
```

The first call parameter is the interval, in seconds, between the instants of time synchronization. The other one is the address of the NTP server. The performance of synchronization script, set for example for a 10-second interval, has the following effect:

```
2 Nov 11:23:18 ntpdate[17247]: step time server
150.254.183.15 offset 0.000030 sec
2 Nov 11:23:28 ntpdate[17253]: step time server
150.254.183.15 offset 0.000024 sec
2 Nov 11:23:38 ntpdate[17264]: step time server
150.254.183.15 offset 0.000034 sec
```

B. Generating 1-pps signal at serial output

Having provided access to a proper, NTP-synchronized time of the system clock, using the "ntpdate" program and assuming the accuracy of synchronization (and its frequency), we applied a Linux program called "OneShot" to transmit the second of computer clock to the serial port output. In this respect, we abandoned the Windows system because it did not ensure as precise control over transmitting the system clock to the COM interface as Linux did it [4]. In this case the DTR signal was applied, "data transfer ready", selected because of easy service of an event from the level of a program written in C language. There is no need to set the port parameters. In this solution there is no transmission of information in the form of ASCII code characters, but the voltage is changed on a selected pin (in this case – pin 4), which represents the 1-pps signal.

The initialized "OneShot" program acquires the system time with resolution 1 μs, and then re-calculates the time till another second. When the time passes, a pulse is sent to the serial port, and then another re-calculation of time is done. The duration of the procedure of sending this pulse is ca 60 μs. Every 1-second re-calculation of time guarantees that the NTP synchronization is taken into account regardless of the assumed synchronization interval. Already in another second, after the system clock is reset, the program can present all variations or inaccuracies resulting from speeding up or slowing down the system clock. Thus the program tries to adjust the pulses, sent every second to the serial port, to the continuously synchronized and variable computer clock. A standard Linux function of C language gettimeofday [4] is applied to acquire the system clock time. This function acquires the current time, given in the form of a packet, from which the time T' that passed from the last full second can be directly acquired. Based on this time, we calculate the time T in which another pulse to the serial port should be sent using the formula:

$$T = 1000000 - T' \quad (1)$$

For this purpose a counter is set up (it can be compared to an alarm), which after given time runs the procedure of sending a pulse onto the serial port. This counter measures the time till another second, based on the Linux setiTimer [4]. This is a function that measures the elapsing real time and generates an alarm after it has passed. The real time measurement means that even when the program is stopped by the system, the counter will still be working and measuring the time till another second. The performance of the "OneShot" program is presented in a diagram in Fig. 1.

IV. MEASUREMENT OF 1-PPS SIGNALS

In order to evaluate the timing signal obtained at the computer interface, two measuring stands were set up. Two computers (with Linux operating systems), connected to the Internet, are supplied with the software described above. 1-pps signal generated at pin no. 4 of PC serial interface is compared with a reference 1-pps signal generated by GPS controlled oscillator. The phase difference between both signals determines the time interval ΔT at the output of a time interval generation module. The interval ΔT is the measure of time error. Next computer (or two computers) is equipped with specialized measuring cards, enabling the measurement and recording of this interval. A block diagram of such measuring system is presented in Fig. 2. The measurement stands applied are presented in Fig. 3.

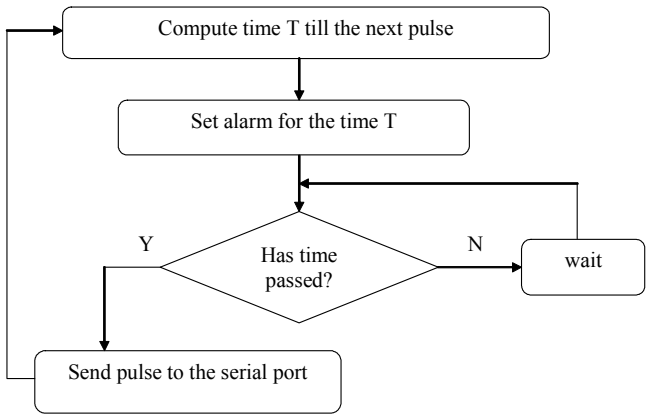


Figure 1. Algorithm of OneShot program performance

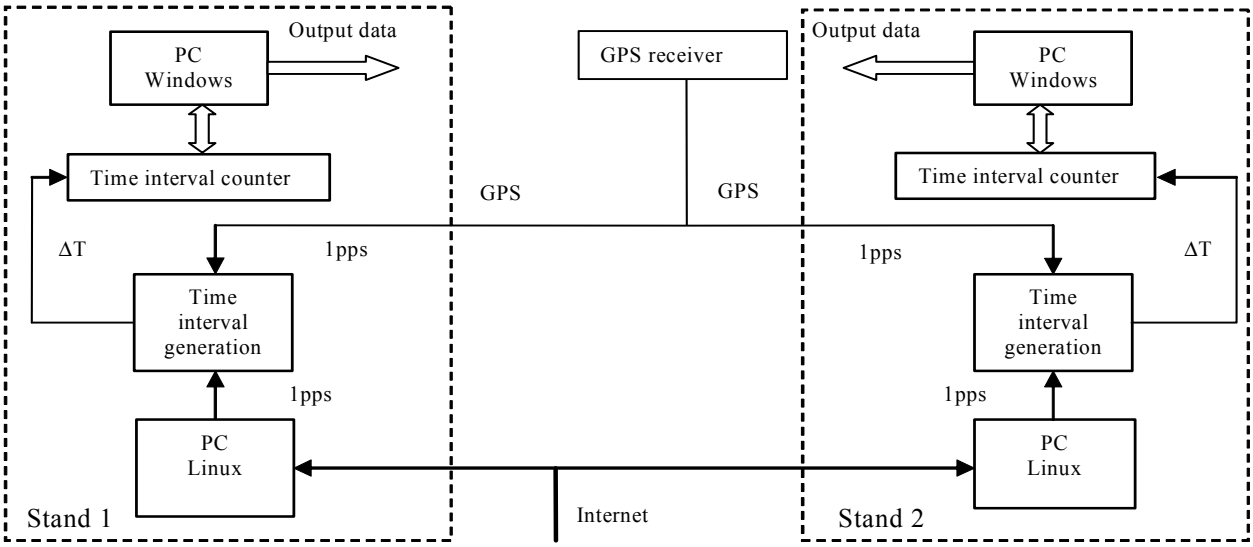


Figure 2. Block diagram of measurement stands

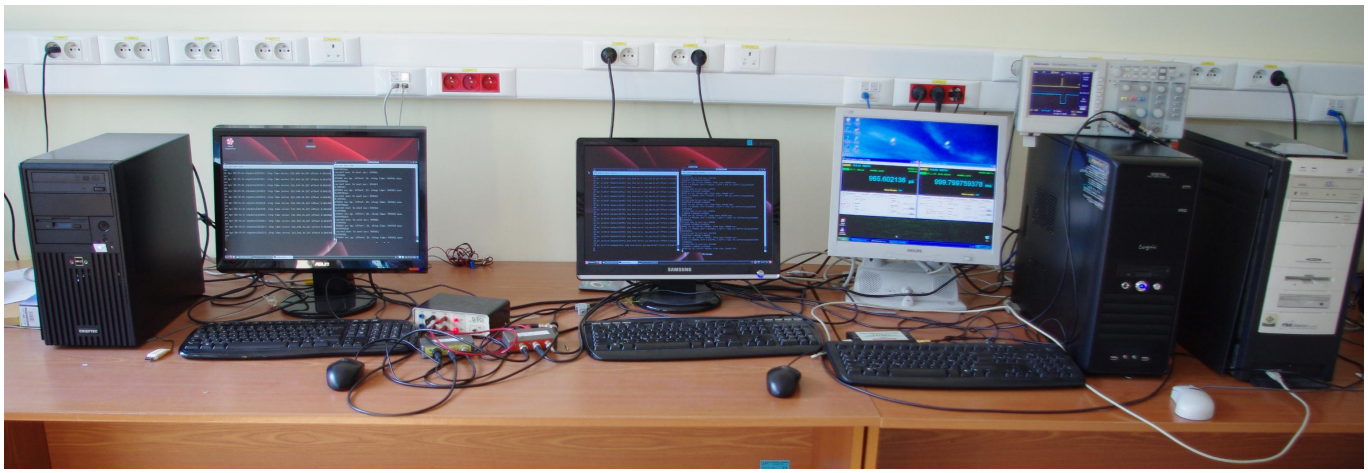


Figure 3. The measurement stands applied

Results of time error measurement of 1-pps signal controlled by NTP server with stratum 1 (tempus1.gum.gov.pl) on both stands simultaneously are presented in Fig. 4. System time of both computers with Linux was synchronized to NTP with 10 s interval.

In the results shown in Fig. 4 a considerable influence of the quality of system clock in computer is visible. The saw-tooth shape of the time error is caused by the frequency offset of the PC internal oscillators. The maximum time error depends on the NTP synchronization interval. Some random peaks of the time error are present in the results. The number and amplitude of these peaks in different results obtained are random. Their amplitudes vary (even up to tens of milliseconds). Similar phenomena were registered in [7].

V. Processing of Controlling Signal

In order to minimize the time difference and the influence of random distortions some additional algorithms controlling the generation of the signal are proposed. The algorithms are working on the values of the time T determining the instants of successively generated pulses.

The methods applied in the algorithms allow averaging or finding a median value of a set of n time intervals T and limiting the changes of the interval T.

In the first case the value T that determines next instant of pulse generation, is average or median value of the set of n last values of interval T. In the second case two thresholds are introduced: threshold t for a maximum value of time T, and threshold c which is an admissible change in the value of time T. The purpose of this threshold algorithm, together with the averaging algorithm, is to limit, and possibly even to eliminate any phase jumps in the time signal, which in the form of peaks can be noticed in Fig. 4.b.

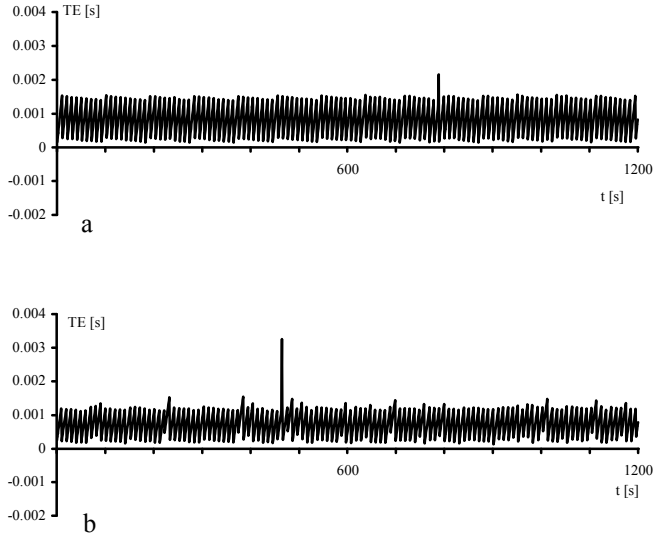


Figure 4. Time error of 1-pps signals controlled by NTP

The algorithm performs in a few steps:

- calculation of the time till another second according to equation (1);
- determining the change in the value of current time in relation to previous time;
- if the absolute value of the change does not exceed the value t, the usual performance of the algorithm is continued;
- otherwise, if a current change does exceed the value t, the difference sign is assumed;
- positive difference means that the time has been elongated, then the previous (last) value increased by c is substituted for a new time T;
- if we put the value c0, then the last value of time T is substituted;
- negative difference means that the time has been shortened, then the previous value, decreased by c, is substituted for a new time T;
- thus calculated new value determines the time till the next second.

The algorithm of determining an admissible threshold in the time signal with value m and the time signal jump to value c for time T is shown in Fig. 5.

Example of execution of the procedure:

```
user@user-desktop ~/time $ timer -m3 -t100 -c100
```

where the parameters are: m3 (or s3) – median (or average) from 3 last values of T; t100 – limit of changes of the value of T set to 100 μs; c100 – fixed change (100 μs) of the value of T, if the limit t is exceeded (for c=0 previous value of the interval T is used).

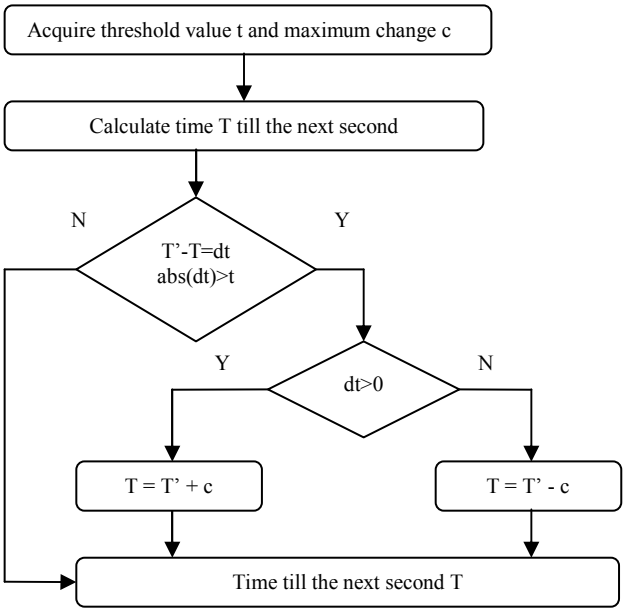


Figure 5. Algorithm for determining threshold t and jump c

VI. MEASUREMENT OF 1-PPS SIGNAL WITH ADDITIONAL CONTROL

In order to evaluate the performance of the proposed algorithms, simultaneous measurements were performed on both measuring stands. Each pair of the measurement results contains one time error series of timing signal obtained without additional procedures and one time error series of timing signal obtained with additional processing. The range of the parameters' values set in the procedures proposed is very wide. Firstly, it is necessary to set properly the number of samples assumed for averaging or calculating the median from successive seconds. Measurements were done for different parameters, and a better stability of the obtained timing signal was confirmed for the median algorithm. The quality of computer clock limits the selection of the number of considered items (values of T) to the value of 10. Still the more difficult is to select the appropriate threshold parameters. Too short intervals of times t and c cause oscillations in the time signal. Setting too long values for those times makes the threshold effect cease. The results of obtained registrations, which illustrate these phenomena, are shown in Fig. 6 and 7.

From the evaluation of results in Fig. 6 and 7 it can be concluded that the peaks in the time error signal were removed by filtering algorithms; whereas the setting of the algorithm in Fig. 6.b is selected inappropriately, which is visible as oscillations in the time signal. The algorithm the performance of which is shown in Fig. 7.b removes peaks and does not introduce oscillations, but it increases the time signal error ca 2 times. The results obtained indicate, however, a perspective for further research; they also prove that it is necessary to optimize the applied algorithms.

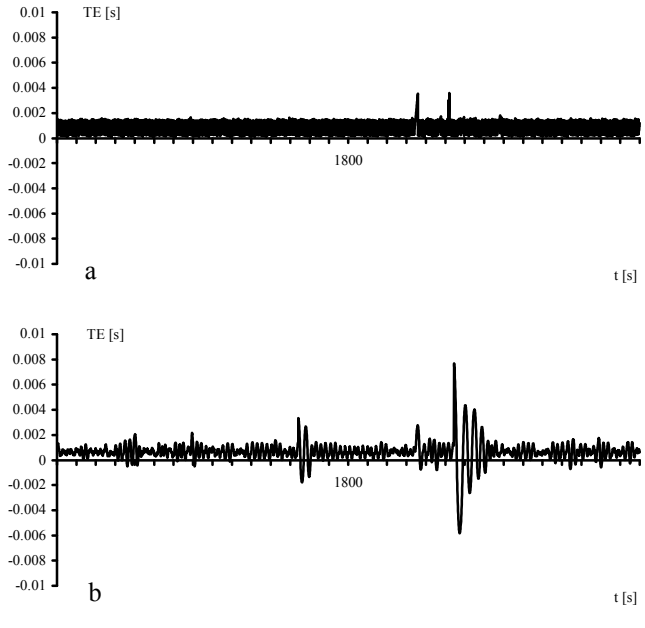


Figure 6. Time error of 1-pps signal without (a) and with additional processing (b) with parameters m3, t100, c100

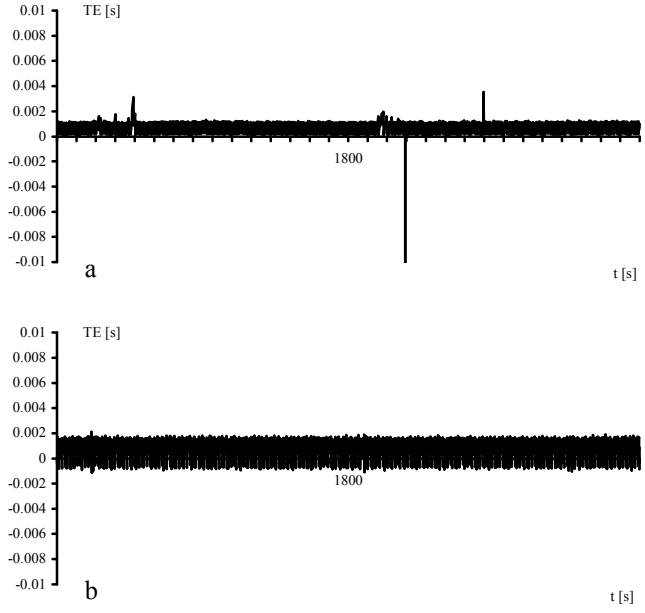


Figure 7. Time error of 1-pps signal without (a) and with additional processing (b) with parameters m3, t7000, c1000

VII. CONCLUSIONS

The technique described in this paper can be utilized for applications different than computer networks at the level of accuracy of the order of several milliseconds. It can be used e.g. for control of the clocks with display in public areas, when the access to the high quality reference is impossible or not necessary. The 1-pps signal controlled by NTP can be applied also in the systems that require the electrical form of the reference signal instead of the data stream.

The quality of the time signal generated on the computer interface depends on the quality of internal computer clock. This quality can be improved by appropriate selection of the parameters of software responsible for cooperation with NTP (intervals of synchronization) and the parameters of software generating the 1-pps signal (control of instants of pulses). Frequent reference to the NTP synchronization procedure makes the time signal more sensitive to the influence of events in the Internet network and may caused other processes simultaneously running on the computer. Rare reference causes increment of the time difference resulting from the frequency inaccuracy of the computer clock.

Some random peaks were observed in the timing signal, which we tried to eliminate using additional algorithms affecting the generated timing signal. The results obtained require, however, the research work to be continued. Further research will make it possible to improve the performance of these algorithms, and to decrease the time error in the obtained signal.

REFERENCES

[1] A. Dobrogowski, M. Jessa, M. Kasznia, K. Lange, "Selected research on time signals cotrolled by NTP, Wybrane badania sygnalow czasu sterowanych protokolem NTP," Przeglad Telekomunikacyjny + Wiadomosci Telekomunikacyjne, no. 8-9, 2009, pp. 1622-1630.

[2] A. Dobrogowski, M. Jessa, M. Kasznia, K. Lange, "Results of evaluation of time signals receiving from NTP servers in Poland," Proc. of 24th European Frequency and Time Forum, Noordwijk, Netherlands, 13-16 April 2010.

[3] A. Dobrogowski, M. Jessa, M. Kasznia, K. Lange, "Research on time signals controlled by NTP, Badania sygnalow czasu sterowanych protokolem NTP," Proc. of the Conference on Teleinformatics Services and Networks Next Generation, Warsaw, Poland, 23-24 November 2010.

[4] Linux Programmer's Manual.

[5] D. L. Mills, "Precision synchronization of computer network clocks," ACM Computer Communication Review 24, April 1994.

[6] D. L. Mills, "Network Time Protocol Version 4. Reference and Implementation Guide," NTP Working Group Technical Report 06-6-1, University of Delaware, June 2006.

[7] V. Smolacha, J. Cermak, J. Palacio, "On calibration of network time services," Metrologia 45 (2008) S51-S58.

[8] P. Rybarczyk, "Expert Network Time Protocol: An Experience in Time with NTP," Apress, 2005.

Active optical pre-compensation in short range frequency transfer in optical single-mode fiber

Per Olof Hedekvist*, Sven-Christian Ebenhag* and Kenneth Jaldehag
Measurement Technology, Communications Lab
SP Technical Research Institute of Sweden
Borås, Sweden
*also with Department of Microelectronics, Photonics Lab
Chalmers University of Technology
Göteborg, Sweden
per.olof.hedekvist@sp.se

***Abstract*— Short distance fiber based optic distribution of time and frequency is often considered sufficiently stable and therefore, uncompensated one way transfer is used. However, when fiber distances increase to a few km, and a substantial amount of the fiber is installed outdoors or in aisles and attics without temperature control, the delay through the fiber may vary substantially. Since the variations are slow, a microwave frequency modulated on an optical carrier may be assumed to remain sufficiently stable. The time delay through the fiber will nevertheless depend on the extremes of the temperature variation which must be taken into account or compensated for. This study presents the design issues of a real time compensation and some results of an optical compensation technique.**

I. INTRODUCTION

The motivation of this work is to transfer a microwave frequency over a short fiber optic link, with a real time compensation of the delay variations of the fiber. The result will be a fiber link in which the transfer time variations, τ, is actively compensated, in comparison to links with actively compensated frequency transfer, operating on the derivative of transfer time variations $d\tau/dt$.

Previous work on multimode fiber A similar setup using light at 850 nm through multimode fiber has been presented previously [1], as well as a setup based on matched electronic delays [2].

II. THEORY

The delay of an optical single-mode fiber will vary as the temperature of the fiber cable changes. A typical value for standard single-mode fiber is 37 ps/km/K [3]. This variation will occur on outdoor installations, especially when sunlit, but also indoor on non- insulate attics, unheated stairwells and sun heated outer walls. The temperature span of an outdoor installation can exceed 20K over a day and 40K over a year, and in indoor and buried fiber the temperature may change as much as 20K[4] over a year.

If a variable delay is adjusted to counteract the fiber delay variations, the net delay remains constant. This requires two way fiber transmission, and a bidirectional optical delay. Two way fiber transmissions can be achieved in parallel fibers (duplex) or return in same fiber (simplex). Furthermore, the Optical Delay Line (ODL) must have a tuning range exceeding the anticipated fiber delay variations.

III. EXPERIMENT AND RESULTS

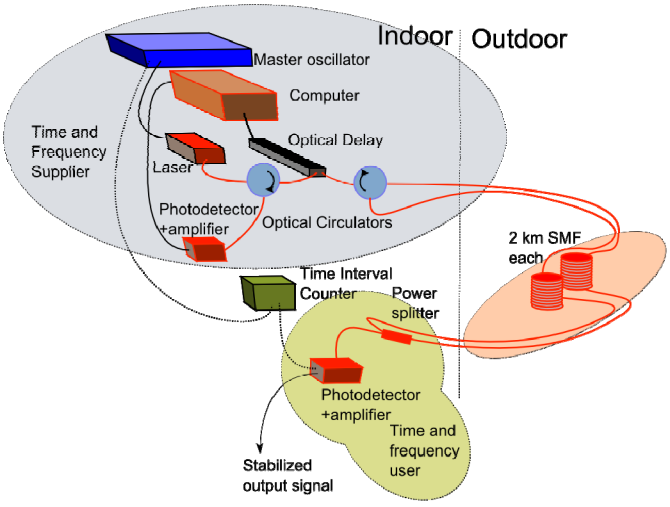


Figure 1. Schematical setup of the experiment.

The experimental setup is schematically described in Fig.1 Two spools of 2 km Standard Single Mode Fiber are installed on the roof. At the user end, 50% of the light is

extracted and 50% returned. Two identical photodetectors are used at both user and supplier end. The returning signal is compared to the transmitted in a double balanced mixer, where the output is detected in a DC-Voltmeter (DCV) set at long integration time. The reading of the DCV is evaluated by a computer that controls the fiber optic delay to keep the voltage constant. The DCV must wait until delay is set at new steady value before starting a new integration of values. This setup is indicated in Fig. 2 below.

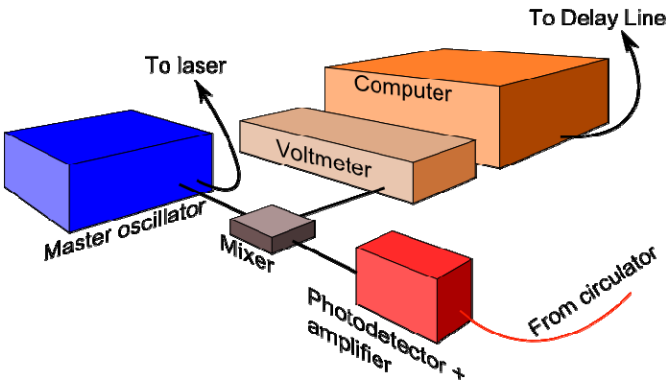


Figure 2. Detailed schematic of return equipment at trensmitter end.

The Time Interval Counter (TIC) is used to evaluate the quality of the fiber system. The Master oscillator is connected to UTC(SP) through an indoor fiber optic link.

A. *Equipment Description*

The laser is a semiconductor laser diode for telecommunication, emitting light at 1550 nm into a single mode fiber. The laser diode is directly modulated by the 10 MHz signal from the Master oscillator.

The photodetectors are semiconductor photodiodes which responds to light at wavelengths of 1300 – 1600 nm, at an electrical bandwidth up to 100 MHz. To increase the electrical signal amplitude, it is amplified before entering the mixer.

The optical circulator will transfer light entering port 1 to exit from port 2. Light entering port 2 will exit from port 3. No light is transferred from port 3 to port 1.

A microwave mixer with identical frequency on both inputs, will emit a DC voltage proportional to the amplitudes of the input signals and *cos(phase),* where *phase* is the phase difference between the two signals. Since there is a fixed relation between *phase* and time, this can be used to estimate the time difference with high resolution.

B. *Experimental Results*

The result of the measurement is shown in the graph in Fig.3. The red dots (left axis) show the time interval measured In the TIC. When outdoor temperature is within 17 – 20°C, the fiber link is stabilized and the transfer time through the fiber link is constant. However, when the temperature decreases below this interval, the ODL reaches its limit, and the transfer time varies with temperature. During test period, outdoor temperature varied between 0 and +20 °C. Furthermore, in active stabilization, a small temperature dependence is still apparent.

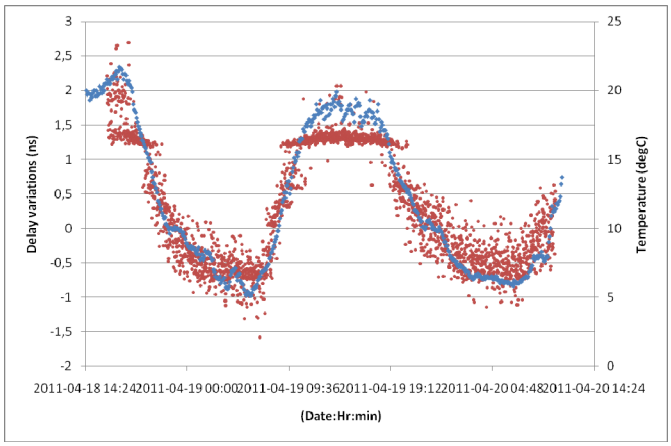


Figure 3. Stabilized output time signal (red, left Y-axis) and outdoor temperature (blur, right y-axis)

IV. CONCLUSIONS

The measurements indicate that the active stabilization of the fiber delay is a feasible method for increased stability on time and frequency transfer over short fiber links. Nevertheless, the ODL must be bidirectional, and its tuning range matched to the fiber link and the installation environment. The commercial ODL used in the experiment have ±100 ps tuning range, which was shown to be insufficient in the present setup. The system regulates the transmitted phase as long as temperature does not change more than a few degrees.

REFERENCES

[1] S.C. Ebenhag, P.O. Hedekvist, C. Rieck, H. Skoogh, P.Jarlemark, and K. Jaldehag, "*A fiber based frequency distribution system with enhanced output phase stability*", 23rd European Frequency and Time Forum joint with IEEE Frequency Control Symposium (EFTF and IEEE-IFCS), Page(s):1061– 1064, 2009.

[2] A. Czubla, L. Sliwczynski, P. Krehlik, L. Buczek, M. Lipinski and J. Nawrocki, *"Stabilization of the propagation delay in the fiber optic frequency distribution using electronic delay lines – first measurement results"*, Presented as Paper 30 at 42nd Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, Va, Nov 15-18, 2010.

[3] K.T.V. Grattan, and B.T. Meggitt "*Optical Fiber Sensor Technology* vol 3" (Dordrecht: Kluwer) 1999

[4] A. Walter, G. Schaefer, *"Chromatic Dispersion Variations in Ultra-Long-haul Transmission Systems Arising from Seasonal Soil Temperature Variations"*, Optical Fiber Communication, OFC02, Paper WU4, March 17-22 Atlanta Ga, 2002

Some remarks on the CCTF CGGTTS format

Jiang Z. and Lewandowski W
Time Department
Bureau International des Poids et Mesures
Pavillon de Breteuil, F-92312, SEVRES CEDEX/France
zjiang@bipm.org

***Abstract*— The CCTF CGGTTS data formats are used worldwide in the applications of various societies from the *UTC*/*TAI* production and scientific studies to all levels of commercial usages. Its design was based on the GPS receiver technique in the 1980s. With the progress of the new GNSS receivers and new T/F transfer techniques, its short points are becoming evident, e.g. the convention that for a tracking arc of 16 minutes only 13 minute data are recorded. In this paper, we discuss the outdate points in the CGGTTS format and how to improve it toward a new CGGTTS convention, namely V03. The new and old versions can co-exist for different applications.**

I. INTRODUCTION

The CCTF GGTTS (version 01) and CGGTTS (version 02) data formats (V01 and V02 hereafter) are used worldwide in the timing community for GNSS time and frequency transfers (T/F). Both are co-exist in the applications of various societies from the *UTC*/*TAI* production and scientific studies to all levels of commercial usages. The designs of the V01 and V02 were based on the state-of-the-art of the receiver manufacture in the 1980s. With the development of the new GNSS receivers and new T/F transfer techniques, some short points, due to the GGTTS convention, become obvious.

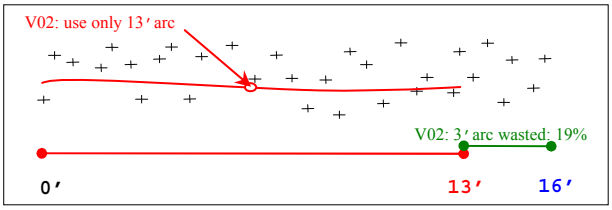


Figure 1. A tracking arc of 16 minutes scheduled by V01 and V02 (among 32 observations of 30 second each, only 26 are recorded)

One of the major defaults in the CGGTTS convention is that for a tracking arc of 16 minutes of measurements only 13 minutes of them are recorded in the CGGTTS (V01 and V02) file. Figure 1 illustrates the case of converting the Rinex observation to CGGTTS. In the earlier days there were only single channel single frequency GPS receivers which followed the BIPM tracking schedule of 16 minutes and for each tracking arc there was a 3 minute's standby for seeking and locking the scheduled satellite. The result is that 3 over 16 (19%) minutes of the raw data rejected this way! This is obviously not reasonable. Theoretically, it is easy to calculate the gain if all the 16 minutes' data will be used: suppose ε is the measurement error of the each 30 seconds' tracking. The error in the average of 13′ is ε/√(26) and that of 16′ ε/√(32). The gain is [ε/√(26)-ε/√(32)]/{ε/√(26)}-1=10%. So nothing but the extra data will have a gain of 10%. As proven below, further improvements will double this gain. Nowadays, dual-frequency multi-channel receivers become popular which allow a continuing satellite locking and data recording. The BIPM tracking schedule and the convention of the 3′ standby become meaningless. This CGGTTS convention is outdating.

The second short-point is that the data are recorded in integer numbers round off at 0.1 ns which accumulates with data processing steps. As we know that in precise and accurate T/F transfers, many corrections, such as the precise orbit, ionosphere and the Earth tides etc. are often less than 0.1ns, i.e. less than the 'round off' errors. In particular, for the future uses of carrier phase, 1 ps recording is necessary.

The third is that the epoch interval of CGGTTS is fixed to 16 minutes. Hence the observation epochs are not the same from one day to next. This is not a convenient practice compared with TW and PPP that are always measured and recorded on integer hours and on the same minute of everyday. The standard MJD of Circular T is on integer hour: 0h. A non-necessary interpolation error is then introduced.

Finally, the recording interval is fixed to 16 minutes. Is it optimal for T/F transfers with modern receivers? Why not varying intervals of 1′, 5′, 10′, 15′, 30′ or 60′ … as options to users according to the specifications of precision and data storing memory required ?

New GNSS systems are coming. To answer the new challenges, we need review the CGGTTS convention. Above we have first investigated the outdating points in CGGTTS and below we will discuss the possibilities to improve the actual version V02 and finally present our considerations toward a new CGGTTS convention, the NEW version, namely V03, is not only for the BIPM *UTC/TAI* computations but also for scientific and commercial purposes. The three, new and old, versions (V01, V02 and V03) can co-exist for different applications in the future.

II. CHARACTERS OF THE CONVENTION: CGGTTS V03

• First the V03 should be as same as possible as the present GGTTS V01 and CGGTTS V02 and compatible to the existing computation programs, e.g. in a Fortran code, it needs only to change the input format from FIX FORMAT to Free FORMAT (*);

• Use and record all the raw data in a tracking arc, i.e. if the

arc is of 16 minutes, all the measurements should be used and converted to the CGGTTS data;

• Optional data collecting intervals for different applications: e.g. 1, 5, 10, 15, 20, 30 and 60 minutes or other user requires. Variable data collecting and recording interval may be interesting for low level or commercial practices. For users interesting only in the long-term stability, a longer recording interval with a less storing memory is a better option. A user can select the best interval according to the specifications of his technical or computer memory etc. Data recording interval is one of the major issues discussed in sections IV, V and VI;

• Optional smoothing methods used in the tracking arcs;

• Allowable, as option, for all the codes, C1, P1, P2 etc.;

• Allowable, as option, for all the GNSS techniques;

• Allowable, as option, for carrier phase information.

III. DATA AND METHOD

A. Data

As numerical examples, we investigate in the paper the P3 data. Hereafter all the analysis is of P3. The standard procedure of Circular T computation is used. In the tables and plots, the solutions are named correspondingly: **V02 13/16** stands for the conventional CGGTTS V02 with tracking arc of 16 minutes of which only 13 minutes are recorded. Similarly, the new solutions are **V03 10/10, 16/16, 20/20, 30/30 and 60/60.** Here the numbers stand for the length of the tracking arcs in minutes, e.g. **20/20** stands for the 20′ arc with all the 20′ data measurements converted to CGGTTS V03. The data sets used are of *UTC* 1001 and *UTC* 1002 (January and February 2010). In order to obtain a reliable statistic result, we analyzed more than 10 short and long baselines coving the continents of Europe, America and Asia. Because the TW links were not very stable since last year, we depend mainly on the comparison with GPS PPP links. We use the Tsoft sub-menu U5 to carry out all the converting computation.

B. Method

The goal is to investigate the relation of the quality of the T/F transfers and the recorded interval length. We use the TW and PPP links as the references which are believed to have a super quality in u_A (uncertainty type A) than the P3 codes. We compare first the standard deviation σ of the Vondrak smoothing residuals in the link computation. Because the computation methods for V02 and V03 are identical and the only difference is the recording intervals, this comparison makes sense. The smaller the σ is, the better the duration of the corresponding interval. Secondly, we compare between the results of V02 and V03 with the TW and PPP. We compute the σ of the differences. It is expected σ2 (V03) < σ1 (V02). We then compute the gain: (σ2-σ1)/σ2. If the gain is bigger than 5%, we consider the gain is significant and the V03 as well as related parameters make sense. We investigate also difference between Comment View and All the View. We present the main statistic results in the plots and tables.

IV. RESULTS AND ANALYSIS

A. Comparison between the arcs 13′/16′ of V02 and V03

We first prove that the Tsoft V03 procedure produces the identical results as V02. Figure 2 illustrates the differences of them. The mean value is 0 ns. The standard deviation σ (of the two raw CGGTTS V02 and V03 without filtering) is ±0.2 ns. Some outliers are up to 2 ns. The differences come from mainly the noises in the raw measurement and how to process them. V03 uses different smoothing and round off at 0.001 ns.

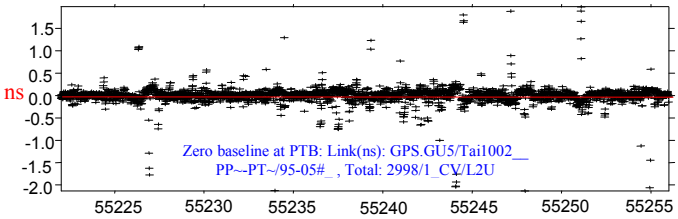


Figure 2. Differences between the V02 13/16 tracking arc link and the V03/U5 13/16 link with σ=±0.2 ns

Figure 3 illustrates the Tdev of an inner-Europe link CH-PTB (640 km). By the Tdev in the figure, we see a significant improve in the short terms until ½ day. Because all the conditions are the same, these gains obtained are thanks to the 19% extra data, i.e. 3 minutes over every 16 minutes.

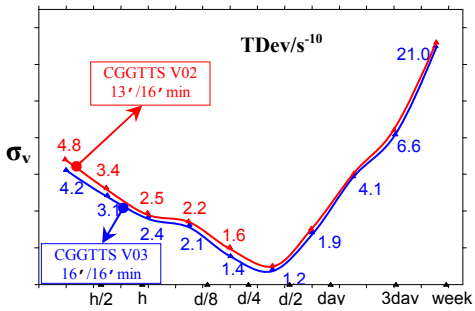


Figure 3. Comparison of Tdev for the link CH-PTB between the V02 13'/16' link (red/up) and the V03 16'/16' link (blue/down)

TABLE I. COMPARISONS OF THE STANDARD DEVIATION OF THE VONDRAK SMOOTHING RESIDUALS BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 16′/16′ LINKS (UNIT IN NS)

Link	Distances km	σ1 V02 13′/16′ /AV epochs	σ2 V03 16/′16′ /AV epochs	Gain (σ1-σ2)/σ1
CH-PTB	640	0.520/2983	0.453/2338	9%
KRIS-PTB	8300	0.629/2840	0.610/2906	3%
NICT-PTB	9100	0.670/3062	0.657/2929	2%
OP-PTB	700	0.764/2124	0.687/2730	10%
OR-PTB	460	0.546/3050	0.516/2933	6%
SP-PTB	1200	0.536/3070	0.501/2943	7%
TL-PTB	9100	0.650/3059	0.607/2939	7%
USNO-PTB	6552	0.611/3080	0.599/2919	2%
KRIS-NICT	1158	0.598/2853	0.548/2916	8%
TL-NICT	2117	0.637/3081	0.575/2995	10%
Average	**3933**	**0.616/2920**	**0.575/2855**	**6.7%**

In the standard *UTC* procedure, we use the Vondrak smoothing to filter the measurement noises. The standard deviation (σ) demonstrates the noises level of a tracking arcs: for V02 the 13′/16′ and for V03 16′/16′. Table I gives the comparisons of the σ for 10 baselines. Most of them are the *UTC* links equipped with HM clocks. For the link CH-PTB, the σ of the residuals is 0.520 and 0.453 ns. We have the gain: (0.520-0.453)/0.520=9%; Similar for TL-PTB (9100 km), we have the σ 0.650 and 0.607 ns. The gain is 7%. Table I shows

that, in all the cases there is a gain, on average 6.7%.

Table II takes the GPS PPP link solution as the reference and gives the results of the comparisons between the V02 and V03 links. Again, in all the cases there is a gain and on average the gain is 6.4%, almost as same as in Table I.

We conclude a gain of 6.6% on average in the V03 format due to the 3 minutes' redundant measurement data.

TABLE II. COMPARISONS OF THE STANDARD DEVIATION OF THE DIFFERENCE BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 16′/16′ LINKS VS. THE PPP LINKS (UNIT IN NS)

Link	Distances km	σ3 V02 13′/16′	σ4 V03 16′/16′	Gain (σ3-σ4)/σ3
CH-PTB	640	0.520	0.474	9%
KRIS-PTB	8300	0.713	0.687	4%
NICT-PTB	9100	0.761	0.754	0.1%
OP-PTB	700	0.779	0.668	15%
OR-PTB	460	0.669	0.637	5%
SP-PTB	1200	0.615	0.576	6%
TL-PTB	9100	0.712	0.671	6%
USNO-PTB	6552	0.684	0.671	2%
KRIS-NICT	1158	0.715	0.656	8%
TL-NICT	2117	0.769	0.711	8%
Average	**3933**	**0.694**	**0.650**	**6.4%**

B. *Comparison between the arcs of V02 13′/16′ and V03 20′/20′*

The length of the tracking arc of 13′/16′ is outdated. The major receivers now used in practices are of the dual-frequency-multi-channel. By testing the 20′, 30′ and 60′, we try to find an optimal tracking length for precise time transfer. We use the same method as above.

TABLE III. COMPARISONS OF THE STANDARD DEVIATION OF THE VONDRAK SMOOTHING RESIDUALS BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 20′/20′ LINKS (UNIT IN NS)

Link	Distances km	σ1 V02 13′/16′ /AV epochs	σ2 V03 20′/20′ /AV epochs	Gain (σ1-σ2)/σ1
CH-PTB	640	0.497/3070	0.427/2338	12%
KRIS-PTB	8300	0.629/2840	0.578/2313	8%
NICT-PTB	9100	0.670/3062	0.626/2330	7%
OP-PTB	700	0.764/2124	0.655/2169	15%
OR-PTB	460	0.546/3050	0.479/2329	12%
SP-PTB	1200	0.536/3070	0.464/2338	13%
TL-PTB	9100	0.650/3059	0.573/2344	12%
USNO-PTB	6552	0.611/3080	0.560/2314	8%
KRIS-NICT	1158	0.598/2853	0.510/2318	15%
TL-NICT	2117	0.637/3081	0.542/2359	15%
Average	**3933**	**0.616/2920**	**0.541/2315**	**12.2%**

TABLE IV. COMPARISONS OF THE STANDARD DEVIATION OF THE DIFFERENCE BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 20′/20′ LINKS VS. THE PPP LINKS (UNIT IN NS)

Link	Distances km	σ3 V02 13′/16′	σ4 V03 20′/20′	Gain (σ3-σ4)/σ3
CH-PTB	640	0.520	0.453	13%
KRIS-PTB	8300	0.713	0.660	7%
NICT-PTB	9100	0.761	0.728	4%
OP-PTB	700	0.779	0.634	19%
OR-PTB	460	0.669	0.611	9%
SP-PTB	1200	0.615	0.542	12%
TL-PTB	9100	0.712	0.639	10%
USNO-PTB	6552	0.684	0.635	7%
KRIS-NICT	1158	0.715	0.627	12%
TL-NICT	2117	0.769	0.684	11%
Average	**3933**	**0.694**	**0.621**	**10.5%**

Table III is the comparisons of the σ between the V02 and V03 links. All the gains are positive, 12.2% on average. Table IV makes the comparisons using GPS PPP as reference. Again, there is a gain in all the cases, 10.5% on average. Increasing the tracking length from 13′/16′ to 20′/20′, a gain of 11% is obtained in u_A. Meanwhile the size of the data file decreases more than 20%.

C. *Comparison between the arcs of V02 13′/16′ and V03 30′/30′*

We continue to increase the length of the tracking arc up to 30 minutes using the same method. Table V presents a gain of 18.7% on average. Table VI makes the comparisons to GPS PPP. Again, there is a gain of 15.9% on average. Increasing the length of tracking arc from 13′/16′ to 30′/30′ makes a gain of 17% in u_A on average. Meanwhile the size of the V03 data file is only half of that of V02.

TABLE V. COMPARISONS OF THE STANDARD DEVIATION OF THE VONDRAK SMOOTHING RESIDUALS BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 30′/30′ LINKS (UNIT IN NS)

Link	Distances km	σ1 V02 13′/16′ /AV epochs	σ2 V03 30′/30′ /AV epochs	Gain (σ1-σ2)/σ1
CH-PTB	640	0.497/3070	0.376/1530	25%
KRIS-PTB	8300	0.629/2840	0.547/1513	12%
NICT-PTB	9100	0.670/3062	0.590/1528	12%
OP-PTB	700	0.764/2124	0.608/1424	20%
OR-PTB	460	0.546/3050	0.445/1534	18%
SP-PTB	1200	0.536/3070	0.430/1537	20%
TL-PTB	9100	0.650/3059	0.526/1533	19%
USNO-PTB	6552	0.611/3080	0.524/1510	14%
KRIS-NICT	1158	0.598/2853	0.469/1529	15%
TL-NICT	2117	0.637/3081	0.499/1549	22%
Average	**3933**	**0.616/2920**	**0.501/1519**	**18.7%**

TABLE VI. COMPARISONS OF THE STANDARD DEVIATION OF THE DIFFERENCE BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 30′/30′ LINKS VS. THE PPP LINKS (UNIT IN NS)

Link	Distances km	σ3 V02 13′/16′	σ4 V03 30′/30′	Gain (σ3-σ4)/σ3
CH-PTB	640	0.520	0.408	22%
KRIS-PTB	8300	0.713	0.631	12%
NICT-PTB	9100	0.761	0.703	8%
OP-PTB	700	0.779	0.577	26%
OR-PTB	460	0.669	0.571	15%
SP-PTB	1200	0.615	0.507	18%
TL-PTB	9100	0.712	0.592	17%
USNO-PTB	6552	0.684	0.610	11%
KRIS-NICT	1158	0.715	0.593	18%
TL-NICT	2117	0.769	0.653	15%
Average	**3933**	**0.694**	**0.584**	**15.9%**

D. *Comparison between the arcs of V02 13′/16′ and V03 60′/60′*

TABLE VII. COMPARISONS OF THE STANDARD DEVIATION OF THE VONDRAK SMOOTHING RESIDUALS BETWEEN THE V02 13′/16′ TRACKING ARC LINKS AND THE V03/U5 60′/60′ LINKS (UNIT IN NS)

Link	Distances km	σ1 V02 13′/16′ /AV epochs	σ2 V03 60′/60′ /AV epochs	Gain (σ1-σ2)/σ1
CH-PTB	640	0.497/3070	0.321/731	34%
KRIS-PTB	8300	0.629/2840	0.537/579	15%
NICT-PTB	9100	0.670/3062	0.585/736	13%
OP-PTB	700	0.764/2124	0.514/679	33%
OR-PTB	460	0.546/3050	0.419/733	23%
SP-PTB	1200	0.536/3070	0.380/735	29%
TL-PTB	9100	0.650/3059	0.498/738	23%
USNO-PTB	6552	0.611/3080	0.465/718	34%
KRIS-NICT	1158	0.598/2853	0.450/730	25%
TL-NICT	2117	0.637/3081	0.440/739	31%
Average	**3933**	**0.616/2920**	**0.461/712**	**26.6%**

We increase the length of the tracking arc up to 60 minutes. Table VII and VIII show the gains of 26.6% and 18.0% correspondingly. The difference of the gains becomes important. The **σ** in Table VII is a self-smoothing. When the tracking arc increases longer than half hour, this **σ** would become too optimism. On the other side, the PPP is not errorless. The u_A of PPP is about 0.3 ns. So the **σ** in Table VIII would be too conservative. Therefore the real gain should be between the two estimations, i.e. 18.0% and 26.6%. We take the average 22% as the gain of increasing the tracking length from 13′/16′ to 60′/60′. Meanwhile the size of the V03 data file is only a fourth of that of V02. Longer length tracking arcs implies smaller size data collection.

TABLE VIII. COMPARISONS OF THE STANDARD DEVIATION OF THE DIFFERENCE BETWEEN THE V02 13′/16′ TRACKING ARC INKS AND THE V03/U5 60′/60′ LINKS VS. THE PPP LINKS (UNIT IN NS)

Link	Distances km	σ3 V02 13′/16′	σ4 V03 60′/60′	Gain (σ3-σ4)/σ3
CH-PTB	640	0.520	0.364	30%
KRIS-PTB	8300	0.713	0.614	14%
NICT-PTB	9100	0.761	0.711	7%
OP-PTB	700	0.779	0.498	19%
OR-PTB	460	0.669	0.565	16%
SP-PTB	1200	0.615	0.470	24%
TL-PTB	9100	0.712	0.598	17%
USNO-PTB	6552	0.684	0.566	17%
KRIS-NICT	1158	0.715	0.614	14%
TL-NICT	2117	0.769	0.598	22%
Average	**3933**	**0.694**	**0.569**	**18.0%**

E. Gains as a function of length of the tracking arcs

The gain is a function of the tracking length or the recording interval. As the length increases, the gain increases. We make in this section a résumé using the averaged gains in the above tables.

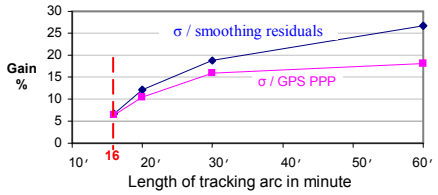


Figure 4. Gains in percentage as a function of the tracking arc length

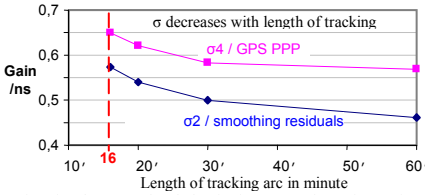


Figure 5. Gains in the standard deviation as a function of tracking length

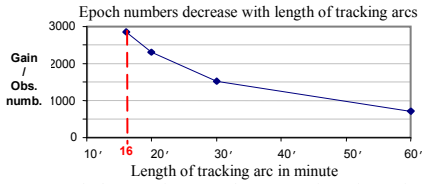


Figure 6. Decrease of observation numbers as a function of tracking length

Figure 4 illustrates the total gains in percentage obtained by increasing the tracking length. From 16′ to 20′, the gain increases more rapid than that between 20′ and 30′. The latter, in turn, increases more rapidly than between 30′ and 60′. Similar phenomena are observed in the Figure 5 where the gains presented with the decrease of the stability of short terms. Meanwhile, as the gain increases, the size of the data decreases by a factor of 4 (Figure 6). However, the tendencies displayed in the figures demonstrate that longer than 1 hour of the tracking arc, the speed of quality improving slows down.

V. TWO REMARKS

A. Comparison with TW

We used also the TW time links as the reference to estimate the gains of V03 against V02. Unfortunately the quality of the European-American TW links degrades since last year. We use then an Asia TW link: KRIS-NICT 1002. The distance is 1158 km. The sigma is only 0.131 ns by the *UTC* standard procedure. The standard deviation of the differences between TW and GPS PPP links is 0.333 ns.

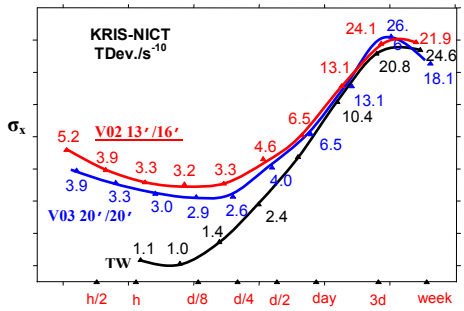


Figure 7. Tdev of the KRIS-NICT baseline: the V02 13′/16′ tracking arc inks and that of the V03 20′/20′ links vs. the TW link

Figure 7 illustrates the time stabilities of the three links: the link of V02 13′/16′, the link of V03 20′/20′ and the TW link. Obviously the TW link is the most stable. The Tdev. of V03 is well below that of the V02 until the averaging time of half day. The three converge after.

We use the same method to compute the gains as did with GPS PPP. The result is listed in Table IX. The total number of the TW measurements is 705. The recording interval is one hour. Therefore for the comparison of V03 60′/60′, it is a point to point comparison. The gains for the lengths of tracking arcs of 16′, 20′, 30′ and 60′ are 7%, 11%, 17% and 21%. The average is 13.9%.

Its comparisons to GPS PPP have been listed in Tables II, IV, VI and VIII. The gains are respectively 8%, 12%, 18% and 14%. The average is 13%. The gains given by the two references of TW and GPS PPP agree with each other.

TABLE IX. COMPARISONS OF STANDARD DEVIATION OF THE DIFFERENCE BETWEEN THE V02 13′/16′ TRACKING ARC KRIS-NICT LINK AND THE V03 16′/16′, 20′/20′, 30′/30′ AND 60′/60′ LINKS VS. THE TW LINK (UNIT IN NS)

σ/V02 13′/16′	σ/V03 16′/16′	σ/V03 20′/20′	σ/V03 30′/30′	σ/V03 60′/60′	Mean.
0.788	0.737	0.701	0.652	0.623	
-	6.5%	11.0%	17.3%	20.9%	13.9%

By conclusion, there is a significant gain in V03 vs. V02, by increase the length of the tracking arc from 20′ to 60′, we have a gain of 14% in u_A on average. However, increasing the recording intervals will decrease the short term details of the

clock behaviors. A user should make the best choices according to the specification of his applications.

B. *Gain with GPS P3 Comment View (CV)*

So far, we discussed only the standard *UTC* AV (All in View) time links. In fact, the CV links used widely in lower level practices have been also analyzed. General speaking, there were gains too: the CV gain is slightly better than that of the AV in short baselines but worse in long baselines.

C. *The reachable uncertainty u_A*

Taking always the P3 as example, the other codes should have the same result. The GPS P3 All in View technique is used for *UTC* time transfer since 2006. The CGGTTS V02 data are used for the computation of the GPS P3 links. The uncertainty type A (u_A) is 0.7 ns as listed in the section 6 of the Circular T.

Above we have theoretically discussed the gain of full use the data 16 minutes is 10%. This implies the u_A(P3) may reduce to 0.63 ns. Taking into account of round off and increasing the tracking length, we can predict the u_A of 0.6 ns for P3 is reachable.

Now we estimate if the V03 is applied, what we can gain. If disregard the lengths of the tracking arcs, the average values of all the σ and gains in the Tables I ~ VIII are listed in Table X. Averagely, the σ(P3/V02)=0.655 ns, knowing that the nominal u_A of GPS P3 is u_A(P3/V02)=0.7 ns; the σ(P3/V03)=0.563 ns. Hence, 0.6 ns is safely reachable in CGGTTS V03 convention. This can be estimated in another way. Because the averaged gain is 14.4% (Table X) and u_A(P3/V02)=0.7 ns, we have u_A(P3/V03)=0.7ns·(1-14.4%)=0.598 ns.

0.6 ns is still a conservative estimation. Because: 1) all the estimations above were based on that the GPS PPP which is errorless; 2) The average gain of 14.4% is average of all the tested lengths of tracking arcs (data recording interval). As suggested to the V03 for *UTC* time transfer, the lengths are 30′ or 60′. The average gain will be 20% and the related stability is 0.529 ns. Same as above, we have therefore the true u_A(P3/V03) equals $\sqrt{[0.529^2 - u_A^2(\text{PPP})]}$. If we use the values 0.3 ns, 0.2 ns and 0.1 ns to replace the term u_A^2(PPP), we will obtain the u_A(P3/V03) becomes respectively 0.432 ns, 0.490 ns and 0.519 ns.

TABLE X. AVERAGE OF THE SIGMA AND GAIN: (UNIT IN NS)

Tables	Distances km	σ /ns V02 13′/16′	σ /ns V03 16′/20′/30′/60′	Gain
I,III,V,IIV	3933	0.616	0.520	16.0%
II,IV,VI,IIIV	3933	0.694	0.606	12.7%
Average		**0.655**	**0.563**	**14.4%**

VI. TOWARD A CCTF CGGTTS VERSION 03 ?

Section II gives the new characters of the suggested CCTF CGGTTS GNSS data recording format convention version 03. The Annex is an example of the version 03. It is the measurement recording of one hour. The options are given in the **COMMENTS** line: the data recording interval is 1200 seconds (20′), the output data are of P3 without carrier phase smoothing. The terms **REFSV, SRSV, SRGPS** and **DSG** are in 0.01 ns but the main term **REFGPS** is in 0.001 ns.

We tested also to change the terms ELV AZTH in 0.01 degree but this made only several ps differences in the REFGPS in the file of LabYYMM.G1I. It seems to makes no sense to record the ELV AZTH in 0.01 degree.

VII. SUMMARY

- We estimated the u_A and the gains obtained using σ of the residuals of Vondrak smoothing (Tables I, III, V, VII) and that of the PPP as reference (Tables II, IV, VI, VIII). Both give almost the same result for intervals shorter than a hour;
- It seems that shorter baselines have bigger gains than longer baselines, e.g. in Table VIII, the average gain of inter-continental baselines (> 6500 km) is 13% while that of the inner-continental baselines (< 2200 km) is 21%; further investigation needed;
- Longer the tracking length, bigger the gains and less the size of storing memory: the gains for 16′, 20′, 30′ and 60′ arcs are respectively 6.6%, 11.4%, 17.3% and 22.3%. For *UTC* time transfer, the data recording interval should not be longer than 1 hour, knowing that the intervals of TW are 1 or 2 hours; It is interesting to notice that a bigger size of data recording would not guaranty a better quality of T/F transfers. The wave length of the filter makes strong influence on the short term uncertainty (< 1 day). However, longer tracking arc implies more tracking data cut off in a passage of satellite. Although the cut off data are of low elevation satellites and it seems this dose not damage the data quality. In fact, this paper proves an improvement of the quality by a reasonable tracking length. Further study should be made on this point;
- The average of total gains is 14.4%. This may be considered as the gain of V03 with respect to V02;
- If V03 (20′/20′) is used in *UTC* computation, the u_A of GPS P3 may be reduced from 0.7 ns to 0.6 ns;
- If V03 is used, the size of the data storing memory will be considerably reduced depending on the length of tracking arc;
- As V02 dose not replace the V01 (we have still V01 data for *UTC* and other applications in T/F transfers). V01 and V02 will co-exist with V03 for as long as needed.

This study concludes:

- The V03 data recording interval may be variable. For *UTC* time transfer, 1800 seconds or 30 minutes is suggested. Increasing properly the data recording interval suggests improving T/F quality and decreasing the size of storing memory. GNSS P code T/F transfers will benefit of a gain of 14.4% on average thanks to this updating;
- It makes sense to convert the tracking data round off 0.001 ns for the related terms;

- As an option, the output C and/or P codes may be smoothed using the carrier phase information;
- Some flexible tolerances or flags may be applicable to allow using more users;
- It is suggested to the CCTF CGGTTS working group to investigate this issue and to pass to the new convention V03.

Finally, it is important to point out that all these suggested modifications are of once for ever with limited software work. Not any investment is required. The gain in uncertainty is based on fully use the potential in the raw measurement data.

REFERENCES

[1] Allan D. **W.** and Thomas C. 1994, Technical directives for standardization **of** GPS, time receiver software, **Metrologia, 31,** 69-79

[2] Jiang Z. 2010, Rinex to CGGTTS - toward a CCTF CGGTTS format version 03, BIPM TM 177

Annex: Example of the CCTF CGGTTS version 03 produced by U5

```
CGGTTS GPS/GLONASS DATA FORMAT VERSION = 03                                   F30195/U5 at 16:40:24/08/19/10
REV DATE =
RCVR =
CH = 12 (GPS)
IMS = F30195
LAB = USN3
X = +1112162.140 m (GPS)
Y = -4842853.628 m (GPS)
Z = +3985496.084 m (GPS)
FRAME = ITRF
COMMENTS =CMD:V03 M13=20 M16=20 -CombLkCp Eps1=1.D1 Eps2=1.D6 Wgt1=1 Wgt2=1 -SMTH -ELV15 -MES -MESN -MESO -MES1 -MES2 -RdRxXYZ
INT DLY =    0.1 ns (GPS P1),    0.1 ns (GPS P2)
CAB DLY =    0.0 ns (GPS)
REF DLY =    0.1 ns
REF = UTC(USNO)
CKSUM = 8C

PRN CL   MJD   STTIME TRKL ELV AZTH   REFSV      SRSV     REFGPS    SRGPS  DSG IOE MDTR SMDT MDIO SMDI MSIO SMSI ISG FR HC FRC CK
             hhmmss  s   .1dg .1dg    .1ns     .1ps/s     .1ns    .1ps/s .1ns     .1ns.1ps/s.1ns.1ps/s.1ns.1ps/s.1ns
  9 FF 55222  020 0 1200 175  446 -1022031.0x  -19.2       11.01    3.6  2.2  93  271  104  -10  -42  -10  -42  22  0  0 L3P 37
 22 FF 55222  020 0 1200 815 1365 -1807158.2x   -7.8       22.71  -15.8  3.6  81   81    2   -9    1   -9    1   7  0  0 L3P B6
 14 FF 55222  020 0 1200 626 3199   -43406.7x  -41.2       11.58    3.2  0.9  68   90   -2   -7    3   -7    3  10  0  0 L3P DD
 18 FF 55222  020 0 1200 425 1322  -444500.0x  -41.7        9.18  -14.4  1.0 112  119   17   20   -1   20   -1  21  0  0 L3P FB
 30 FF 55222  020 0 1200 330 1249 -2138412.6x   -6.7       76.76   26.3  2.5   8  146  -21   29   12   29   12  18  0  0 L3P 47
 26 FF 55222  020 0 1200 609  326   126848.9x   30.2       15.76   -2.8  0.6  40   92    4  -15    0  -15    0  15  0  0 L3P CA
 12 FF 55222  020 0 1200 312  849  1583474.5x  -30.4       -7.74   18.5  0.9 107  154   -7   29  -12   29  -12  25  0  0 L3P 51
 31 FF 55222  020 0 1200 281 2189   539861.3x   26.3       24.82   41.0  1.2  65  169  -39   49  -10   49  -10  21  0  0 L3P 6B
 27 FF 55222  020 0 1200 122  448 -1204040.4x  -38.0       14.88   -3.9  3.0  18  388  188  -18  -28  -18  -28  34  0  0 L3P 6F
  9 FF 55222  040 0 1200  94  460 -1022037.3x  -38.5       31.97  -15.8  2.3  93  510  347  -29   63  -29   63  39  0  0 L3P 71
 22 FF 55222  040 0 1200 718 1473 -1807150.0x   -0.9       21.40   -8.9  1.8  81   84    4   -6    4   -6    4  10  0  0 L3P D5
 14 FF 55222  040 0 1200 656 3402   -43461.7x  -45.1        9.78   -0.8  1.7  68   88   -1  -12  -13  -12  -13  11  0  0 L3P 2C
 18 FF 55222  040 0 1200 338 1378  -444516.8x  -37.8       25.14  -10.6  0.8 112  145   28   22   20   22   20  23  0  0 L3P 36
 30 FF 55222  040 0 1200 391 1155 -2138450.2x  -19.8       78.70   13.2  2.1   8  126  -12   21    8   21    8  19  0  0 L3P 3A
 26 FF 55222  040 0 1200 555  480   126892.0x   32.7       19.36   -0.3  0.9  40   97    6  -16  -16  -16  -16  18  0  0 L3P 2F
 12 FF 55222  040 0 1200 322  741  1583442.9x  -34.9       19.39   14.0  1.2 107  150    1   -3    5   -3    5  25  0  0 L3P F2
 31 FF 55222  040 0 1200 367 2241   539840.9x   20.0       22.21   34.7  1.0  65  133  -22   29    3   29    3  19  0  0 L3P 29
 22 FF 55222  1 0 0 1200 619 1525 -1807154.6x   -7.7        7.28  -15.7  2.4  81   91    7    2   12    2   12  12  0  0 L3P C1
 14 FF 55222  1 0 0 1200 669   33   -43512.8x  -47.9       11.88   -3.6  0.5  68   87    0  -17   23  -17   23  11  0  0 L3P F8
 18 FF 55222  1 0 0 1200 251 1421  -444547.6x  -43.1       27.05  -15.9  1.0 112  191   51   35  -14   35  -14  28  0  0 L3P 32
 30 FF 55222  1 0 0 1200 437 1039 -2138493.5x  -14.5       74.92   18.4  1.5   8  116   -6    8  -25    8  -25  10  0  0 L3P  6
 26 FF 55222  1 0 0 1200 497  606   126931.6x   31.2       19.42   -1.8  0.7  40  105    8  -15  -13  -15  -13  20  0  0 L3P 18
 12 FF 55222  1 0 0 1200 310  637  1583395.3x  -39.1       30.43    9.8  1.1 107  155    9  -14  -20  -14  -20  29  0  0 L3P 3E
 31 FF 55222  1 0 0 1200 453 2308   539835.0x    9.9       34.02   24.7  1.0  65  112  -14    4  -13    4  -13  17  0  0 L3P  D
… …
```

Remark: As mentioned in the introduction, carrier phase information in the Rinex data file may be used to smooth the C and P codes. This may be an option of V03. The study is undertaking at BIPM.

Comparison of the GLONASS Orbit Products for *UTC* Time Transfer

Jiang Z., Petit G., Harmegnies A., Lewandowski W. and Tisserand L.

Time Department, Bureau International des Poids et Mesures
Pavillon de Breteuil, F-92312, SEVRES CEDEX, France
zjiang@bipm.org

***Abstract*— Accurate GNSS (Global Navigation Satellite System) time transfer needs accurate satellite orbit information. Since Nov. 2009, the GLONASS time transfer is introduced in *UTC* generation. The precise and timely GLONASS orbit information is indispensable for Circular T computation. In this paper, we compare the GLONASS orbits produced by different analysis centers. We study the differences of the products, their influences on the common view time transfers and the delays of the availability. The products to investigate are the rapid and final sp3 ephemeredes produced by ESA (Europe Space Agency) and IAC (Information-Analytical Center). IAC belongs to Russia Federal Space Agency. Its primary goal is to analyze the information on the satellite navigation based on the global systems, mainly the GLONASS. The delay of the IAC sp3 rapid products is 24-28 hours. This satisfies the rhythm of the BIPM monthly Circular T computation, of which the deadline of data collection is the 5th of every month. Since Nov. 2010, the BIPM started to use the IAC ephemeredes to compute the GLONASS satellite orbits for *UTC* time transfer. At present the IAC product is the most suitable for the *UTC* GLONASS time link computations.**

I. INTRODUCTION

GLONASS (GLN for short) time transfer was introduced for the first time in *UTC* generation in Nov. 2009 in circular T 267 [1,2,3]. Precise and timely issued GLONASS orbit information is indispensable for the BIPM monthly Circular T computation. There are different GLONASS orbits products supplied by the analysis centers: ESA and IAC (Information-Analytical Center of the Russia Federal Space Agency) [4]. We should make a choice in favor of the *UTC* time transfers.

There are two major indicators for making the choice:

(1) The accuracy and precision of the ephemerid products, including the biases if there are. The quality of the corrections of the on board satellite clocks etc. Their uncertainties in view of the common view (CV) time transfer;

(2) The delay of the issue of these products. Because the BIPM Circular T is a monthly publication, the deadline of data collection is the 5th of a month. The ephemerid products should be available by the deadline. Usually there are the so called rapid and final products. The later is too late for the Circular T computation. However it can be used in this study to estimate the quality of the rapid products.

We study the differences between different SP3 ephemerid products supplied by ESA and IAC.

Most of GNSS users are familiar with the ESA. We give here only a quick introduction on the IAC. More information can be found in its official English web site [4].

IAC belongs to Russia Federal Space Agency. Its primary goal is to analyze the information concerning the satellite navigation based on the global systems GLONASS (Russian Federation), GPS (USA) and GALILEO (European Community). A basis of a web site is the results of its work. It contains the information concerning the GLN and GPS current orbital constellations, Earth maps of the current and daily navigation availabilities, results of the GNSS navigation, fields monitoring in Moscow area in a real-time mode and a number of the other data, cf. Figure 1.

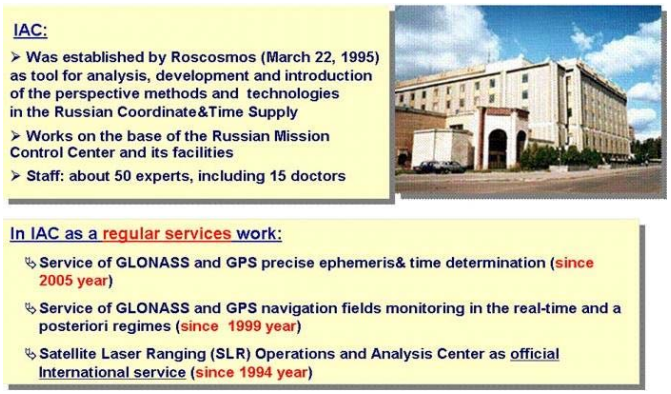


Figure 1. The IAC (Information-Analytical Center) and its missions after http://www.glonass-ianc.rsa.ru/pls/htmldb/f?p=202:1:2253948845654375

IAC supports data archiving and distribution for the Russian and world communities. The main objectives of the archive and FTP-site are to store the GNSS, space geodesy and geodynamics related data and products in a data archive and to disseminate these data and information in a timely manner to Russian and world users.

The IAC archives and distributes the results of the GNSS and SLR data analyzed in IAC and other analysis centers, mainly GLN, GPS and SLR measurements as well as the related products.

The IAC ftp archive consists of 10 head catalogues containing the information about: **MCC**, the results of the IAC work and analysis, including the precise ephemeris and time products for GLN and GPS, daily and monthly bulletins concerning onboard ephemeris and time accuracy, navigation fields quality based on the global network data and others; the **IGS** (International GNSS Service); **IERS** (International Earth Rotation Service); **ILRS** (International Laser Ranging Service); **BIPM** (Bureau International des Poids et Mesures); **NAVCEN**, the official information concerning GPS constellation status from US Coast Guard Navigation Center; **FAF**, the information on indexes of solar and geomagnetic activity; The published official documents, reports and standards of the Russian and international services.

The IAC broadcast navigation information is also indispensable for the precise orbit computations. It can be found in ftp://ftp.glonass-iac.ru/MCC/BRDC/, which however, according to its header, is not an independent product but edited based on all the available IGS broadcast data.

The products supplied by the IAC that we investigated are the rapid and final sp3 ephemerides, which are used to analyze the information on the GLN satellite navigation based on the global systems. The delay of the IAC sp3 rapid products is 24-28 hours vs. a delay of a week by the ESA products.

The timely rapid products of the IAC satisfies the rhythm of the BIPM monthly Circular T production, of which the deadline of the raw data collection of the time links, the clocks and the orbit and ionosphere data is the 5th of every month. As proven in the following sections, the quality of the IAC products is comparable with that of the ESA. Since Nov. 2010, the BIPM started to use the IAC ephemerides to compute the GLN satellite orbits for *UTC* time transfer [2,5,6].

As of this writing, the IAC products are the most suitable for the *UTC* GLN time link computations.

Hereafter are some detailed discussions. In section II, we give the method used in this study and in the sections III-VII we present the results of different numerical tests. Sections VIII and IX are discussion and conclusion.

II. DATA AND METHOD

A. The test data

The main purpose of this paper is to make a comparison in view of *UTC* computation between the products of ESA and IAC:

- different products
- the quality of the products
- availability and rapidity of the products

ESA has not the rapid product but only the final sp3 product. IAC have the both. We have therefore three different products to compare: ESA final sp3, IAC rapid and final sp3. The satellite clock correction is also a topic to study. We can then form four differences:

- between IAC Rapid – IAC Final
- between IAC Rapid – ESA Final
- between IAC Final – ESA Final
- between with and without clock corrections

The time transfer data used is the GLN C/A UTC1009 collected from AOS of Poland, NIS of Egypt, OP of France, PTB of Germany, SU of Russia and UME of Turkey.

Main attention of the numerical analysis is paid at the GLN *UTC* time link SU-PTB. The preliminary conclusion obtained by the baseline SU-PTB was then to be verified with other *UTC* baselines.

B. Method

We use two methods to make the comparisons:

In Method 1, we compare time links computed with the three ephemeredes and we use standard deviation σ of the Vondrak smoothing residuals and Time deviation σ_x of the links as indicators of the quality: smaller the σ and σ_x, the better the ephemeredes. Analysis results are given in sections III-V. In Section VI, we demonstrate the effects of the satellite clock corrections. More results are given in the Annex 1-3 of the reference [5].

In Method 2, we compare the REFSYS values for a single laboratory using different ephemeredes. We estimate how well the quality of the related ephemeredes by computing the σ and σ_x of each difference. This is shown in section VII. Two laboratories in two distant geographic locations are used to provide a variety in geometric configurations: OP in Europe and SG in Asia.

III. TIME LINK COMPARISON USING ESA FINAL AND IAC SP3 FINAL EPHEMEREDES

We first compare the *UTC* time links between SU and PTB and their Tdev obtained using the ESA final (Figure 2) and IAC final (Figure 3) sp3 ephemeredes. Figure 4 illustrates the differences of the two links. The total compared points are 2574 and the σ of the differences is 0.081 ns, which is well below the measurement noise, about 1 ns.

Table I lists the statistics of other 5 baselines. The σ of the differences varies from 0.081 to 0.379 ns. The biggest is the link SG-PTB. SG locates in Singapore. SG-PTB is the longest baseline (6300 km) in this test. The related measurement noise is up to 2 ns. The 4 days' data missing may also explain the big σ. In general, the differences between the GLN time links using the ESA and IAC final sp3 is acceptable vs. their measurement uncertainties.

TABLE I. *UTC* TIME LINK DIFFERENCES USING ESA FINAL AND IAC FINAL SP3 EPHEMEREDES

Baseline	N	σ/ns
AOSPTB	2510	0.200
NISPTB	1264	0.174
OPPTB	2525	0.109
SGPTB	1446	0.379
SUPTB	2574	0.081
UMEPTB	2315	0.148

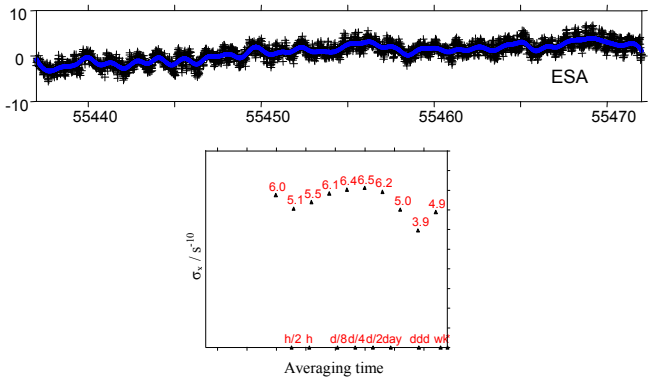


Figure 2. SU-PTB Time link and Tdev using ESA final ephemerid / ns (UTC1009, total CV 3098, smoothing σ 1.068 ns)

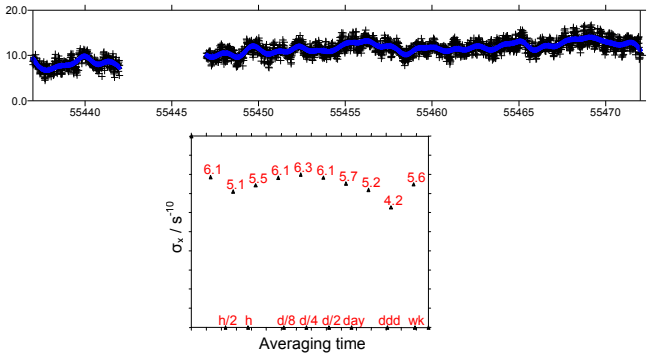


Figure 3. SU-PTB Time link and Tdev using IAC final ephemerid with 4 days' data missing /ns (UTC1009, total CV 2650, smoothing σ 1.055 ns)

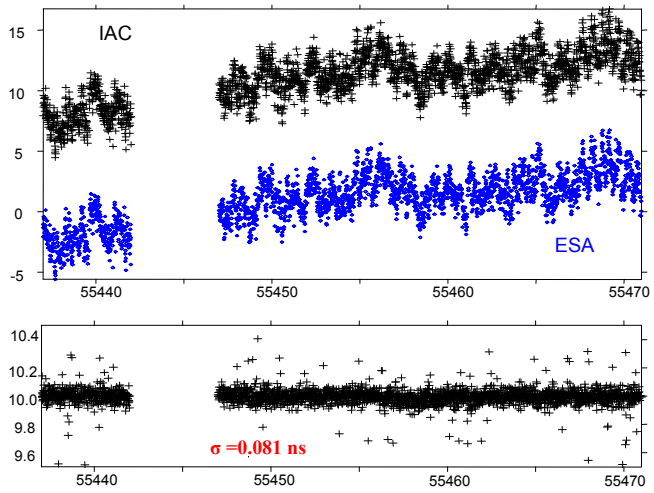


Figure 4. Differences between SU-PTB Time link using ESA and IAC final ephemerid /ns (UTC1009, total points 2574, σ of difference 0.081 ns)

IV. TIME LINK COMPARISON USING ESA FINAL AND IAC RAPID SP3 PRODUCTS

On the 5th every month, we start to compute the Circular T. The final products of either ESA or IAC are too late for the *UTC* computation. Our main concern is then the 2-day delay rapid product of IAC. In this section, comparisons are made between the ESA final and the IAC rapid sp3 ephemeredes. We first compare the *UTC* time links between SU and PTB and their TDev obtained using the ESA final and IAC rapid sp3 (Figure 5) ephemeredes. Figure 6 illustrates the differences of the two results. The total compared points are 3011 and the σ of the differences is 0.089 ns, which is well below the measurement noise and agrees with the statistic result between the final products (Figure 4).

Table II lists the statistics of other 5 baselines. The σ of the differences varies from 0.089 to 0.186 ns. The biggest is still the link SG-PTB due to the long distance. In general, the differences between the GLN time links using the ESA final and IAC rapid sp3 is acceptable vs. their measurement uncertainties. The difference between them is negligible for the purpose of the GLN CV time transfer.

TABLE II. UTC TIME LINK COMPARISON USING ESA FINAL AND IAC RAPID PRODUCTS

Baseline	N	σ/ns
AOSPTB	2942	0.182
NISPTB	1667	0.134
OPPTB	2961	0.125
SGPTB	1676	0.186
SUPTB	3011	0.089
UMEPTB	2734	0.135

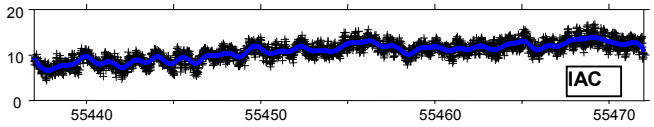


Figure 5. SU-PTB Time link using IAC rapid ephemerid /ns (UTC1009, total CV 3098, smoothing σ 1.066 ns)

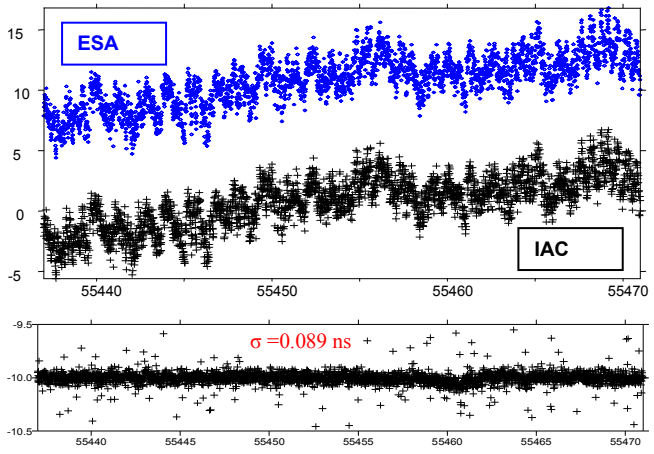


Figure 6. Differences between SU-PTB Time link using ESA final and IAC rapid ephemerid /ns (UTC1009, total points 3011, σ of difference 0.089 ns)

V. TIME LINK COMPARISON USING IAC FINAL AND RAPID

In last section, we conclude the in view of the CV links, the IAC rapid sp3 gives no considerable difference vs. the ESA final one. We now make a comparison within the IAC products between the final and rapid sp3 ephemeredes.

We first compare the *UTC* time links between SU and PTB. Figure 7 illustrates the differences of the two results. The IAC final, as pointed out in Figure 3 and 4, has 4-days hole. The total compared points are 2574 and the σ of the differences is 0.095 ns, which is well below the measurement noise and agrees with the above statistic results.

Table III lists the statistics of other 5 baselines. The σ of the differences varies from 0.095 to 0.377 ns. The biggest is still the link SG-PTB due to the long distance and data missing. In general, the difference is acceptable vs. the measurement uncertainties.

TABLE III. TIME LINK COMPARISON BETWEEN IAC FINAL AND RAPID

Baseline	N	σ/ns
AOSPTB	2510	0.084
NISPTB	1264	0.126
OPPTB	2525	0.100
SGPTB	1446	0.377
SUPTB	2574	0.095
UMEPTB	2315	0.132

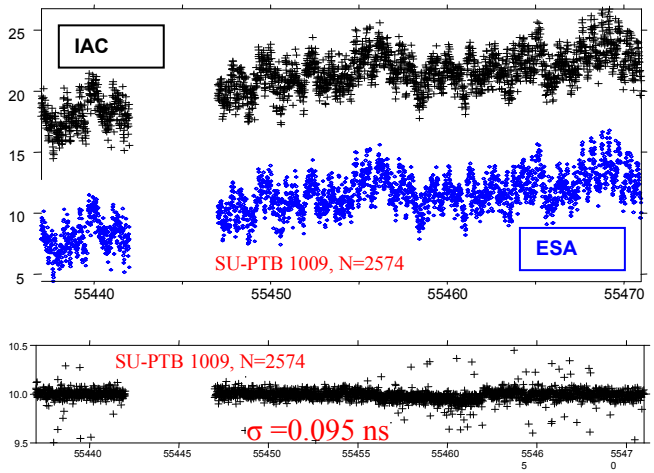


Figure 7. Differences between SU-PTB Time link using IAC rapid and final ephemerid /ns (UTC1009, total points 2574, σ of difference 0.095 ns)

VI. DIFFERENCES BETWEEN ESA AND IAC WITH OR WITHOUT SATELLITE CLOCK CORRECTIONS

In above sections, we did not use the satellite clock corrections given in the sp3 data files. The conclusions drawn from these analysis hold also for the results where the clock corrections are used. This is because, in the CV links, especially for the short or middle baselines, the clock effects are almost cancelled.

Figure 8 illustrates the differences of which the σ is 0.089 ns, as same as the σ obtained among different comparisons in the above sections. The residual influences of the satellite clock corrections are covered by the measurement noises.

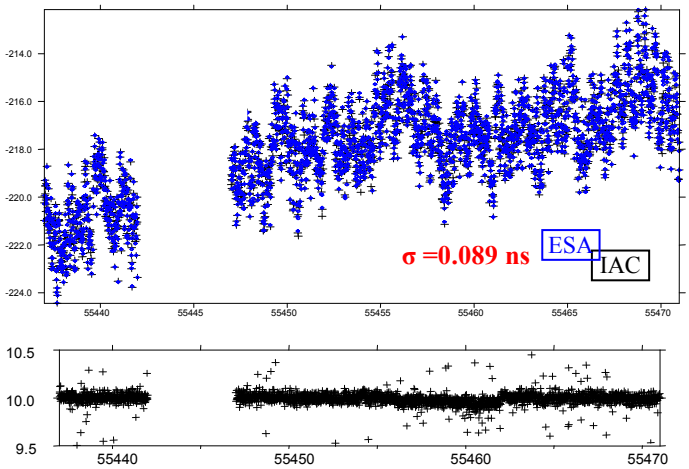


Figure 8. Differences between ESA and IAC final sp3 ephemeredes with or without clock corrections /ns (UTC1009, total points compared 2574, σ of difference 0.089 ns)

VII. COMPARISON OF REFSYS VALUES USING FINAL AND RAPID SP3 EPHEMEREDES FROM ESA AND IAC

In sections III-VI, we analyzed the time links. In this section, we study the influence of the different ephemeredes on the REFSYS values which are the differences between the ground clock and the satellite clock, the satellite time reference system. This is a zero baseline comparison which gives directly the differences between the ephemeredes without the influences of the distance etc.

The comparison results are illustrated in Figures 9, 10 and 11. There are 3 plots. In each figure: the up plot is the difference in ns and down-left plot is the Modified Allan deviation and the down-right the Time deviation.

The comparisons are made between:

- REFSYS differences between ESA and IAC final: Figure 9
- REFSYS differences between ESA and IAC rapid: Figure 10
- REFSYS differences between IAC final and rapid: Figure 11

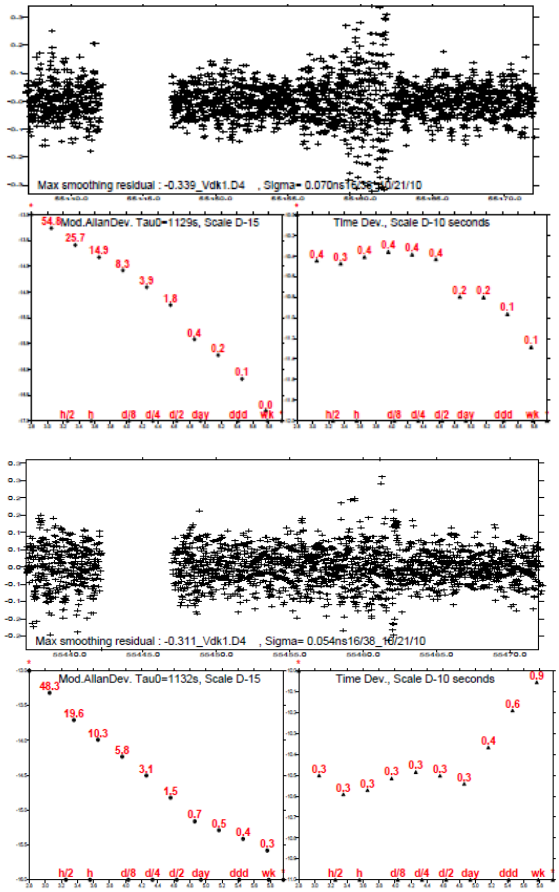
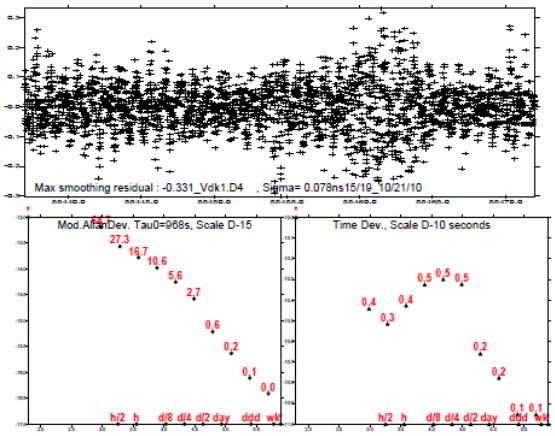


Figure 9. REFSYS differences between ESA and IAC final for SG (up) and OP (down) /ns



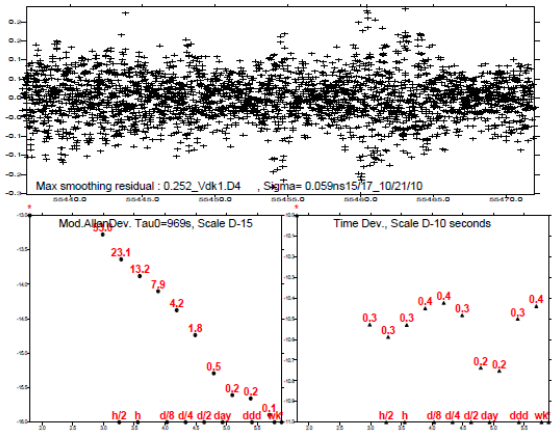


Figure 10. REFSYS differences between ESA and IAC rapid for SG (up) and OP (down) /ns

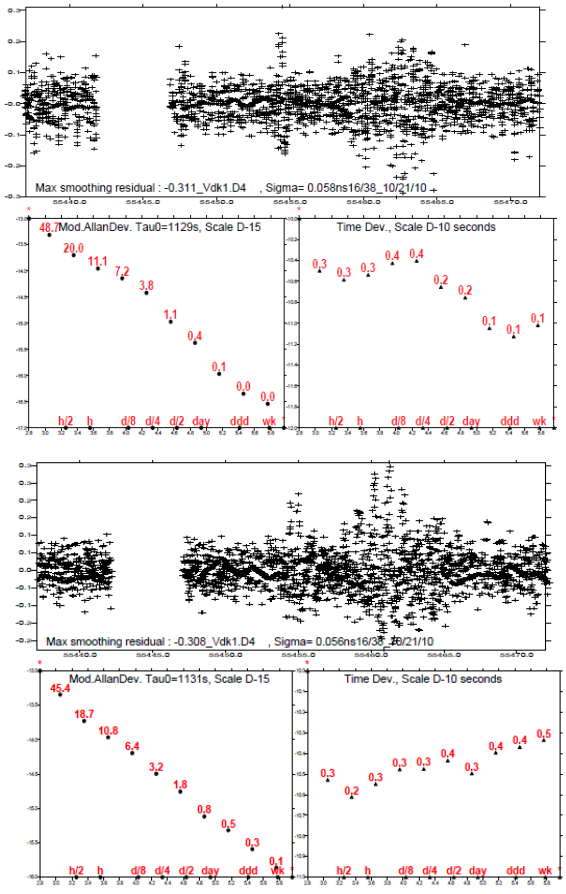


Figure 11. REFSYS differences between IAC final and rapid for SG (up) and OP (down) /ns

TABLE IV. STANDARD DEVIATION OF THE DIFFERENCES OF REFSYS USING ESA/IAC FINAL AND RAPID SP3 EPHEMEREDES

Difference between	SG/ns	OP/ns
ESA final - IAC final	0.070	0.054
EAS final - IAC rapid	0.078	0.059
IAC final – IAC rapid	0.058	0.056

Table IV gathers the statistic results shown in the Figures 9, 10 and 11. The σ varies between 0.05 ns to 0.08 ns. The differences are therefore negligible with respect to the measurement noises.

VIII. DISCUSSION

Method 1 shows that the σ and the first point of Tdev changes very little (1-2%) when using the three different ephemeredes. No significant discrepancies are found between these three products.

Method 2 shows that the three ephemeredes differ from each other less than 0.1 ns in the REFSYS values. In Table IV, the values for the three differences for both laboratories, SG and OP, are similar hence the three ephemeredes are expected to be of similar quality.

The noise level of the GLN C/A code CV time links is about 1 ns [2,6]. The u_A value for CitrT is 1.5 ns. As the difference of the comparisons is very small compared to u_A, we can consider the three products as nearly equivalent. However, we consider that IAC rapid is the best taking into account that it is available more timely than the other ones.

In addition, in Section VI we compared also the CV links with and without clock corrections. The differences due to the clock corrections are less than 0.1 ns. Hardly to say it is a gain by applying the clock corrections. However, we can always use the clock corrections expecting the quality being increased afterwards. Precise clock is the base of the GNSS All in View time transfer. Further study is undertaking at BIPM to use the AV for GLN time transfer.

As the quality of IAC rapid and final are as same as that of ESA, we will use IAC for UTC computation and keep ESA as back up.

IX. CONCLUSION

Since Nov. 2009, the GLONASS time transfer is introduced in *UTC* generation.

Accurate GLN time transfer for *UTC* computation needs accurate satellite orbit information. At present there are available the EAS final and the IAC rapid and final sp3 ephemerid products. Both of the final products are not timely issued for the monthly *UTC* computation. With a 2 day delay, the IAC rapid ephemerid is the only choice.

In this paper, we carefully compared the differences of the IAC rapid products with the final products which are a priori more accurate and precise than the rapid one.

We conclude that, in view of the *UTC* CV time links, the differences between these products and their influences on the CV time transfers are of the order of 0.1 ns which is negligible compared with the measurement noise of 1 ns about.

We should use the IAC rapid sp3 products for *UTC* computation and use the ESA ones as backup.

REFERENCES

[1] BIPM Circular T 263, 2009, ftp://ftp2.bipm.org/pub/tai/publication/cirt.263

[2] Lewandowski W, Jiang Z, 2009, Use of Glonass at the BIPM, *Proc. PTTI2009*, 5-13

[3] BIPM Circular T 274, 2009, ftp://ftp2.bipm.org/pub/tai/publication/cirt.274

[4] IAC, Information-Analytical Center, Russia Federal Space Agency http://www.glonass-ianc.rsa.ru/pls/htmldb/f?p=202:1:2253948845654375

[5] Jiang Z, Tisserand L, Harmegnies A, Lewandowski W and Petit G, 2010, BIPM TM183: Use of the IAC GLN products in UTC time link computation

[6] Jiang Z, Lewandowski W, 2011, Use of GLONASS for UTC Time Transfer, submitted to metrologia, unpublished

Generation of an Ultrastable 578 nm Laser for Yb Lattice Clock

M. Pizzocaro , G. A. Costanzo, and M.Zoppi
Electronic Department, Politecnico di Torino,
Torino, Italy

D. Calonico, F. Levi, A. Mura and A. Godone
Optics Division, INRiM
Torino, Italy
d.calonico@inrim.it

***Abstract*—In this paper we described the development and the characterization of a 578 nm laser source to be the clock laser for an Ytterbium Lattice Optical clock. Two indipendent laser sources have been realized and the characterization of the stability with a beat note technique is presented.**

I. INTRODUCTION

A yellow laser (λ = 578 nm) to be used for the interrogation of the $^1S_0 \leftarrow {}^3P_0$ clock transition in cold ytterbium atoms [1] was realized and characterized. In Figure 1 we report for convenience the relevant transitions for Ytterbium: the two cooling transitions at 399 nm and 556 nm and the clock transition at 578 nm. The yellow light is generated by frequency sum in a waveguide non-linear crystal of a 1319 nm Nd:YAG laser and a 1030 nm Yb doped fiber laser [2,3]. Two independent, ultrastable lasers are obtained from the same light source in a setup with two Fabry-Pérot cavities.

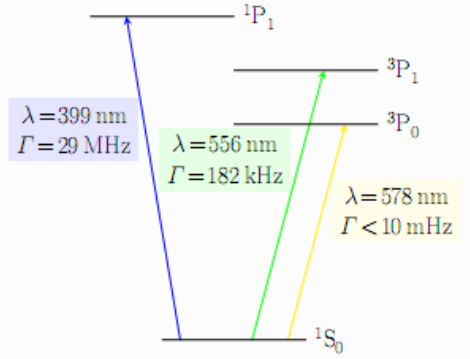


Figure 1. Ytterbium laser cooling and laser clock transitions

II. SUM FREQUENCY GENERATION

The yellow laser is obtained by sum frequency generation of two infrared laser sources (1030 nm and 1319 nm).
The non-linear crystal is a waveguide, magnesium-doped, periodically poled lithium niobate (PPLN) device and works in single pass. Infrared light of both lasers is delivered to the crystal through a polarization-maintaining fiber coupler (y-shaped). The output of the coupler is a bare fiber, aligned very close to the input face of the crystal to couple the light into one of the waveguide channels (the distance is of the order of 1 μm). The crystal is put on a translator to easily align the fiber with the working waveguide channel. In Figure 2 the optical bench scheme for the Sum Frequency Generation is shown.

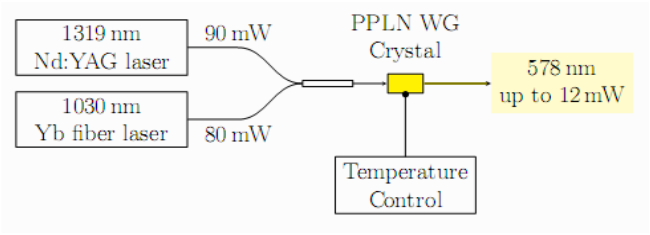


Figure 2. Scheme of the 578 nm laser source Sum Frequency Generation

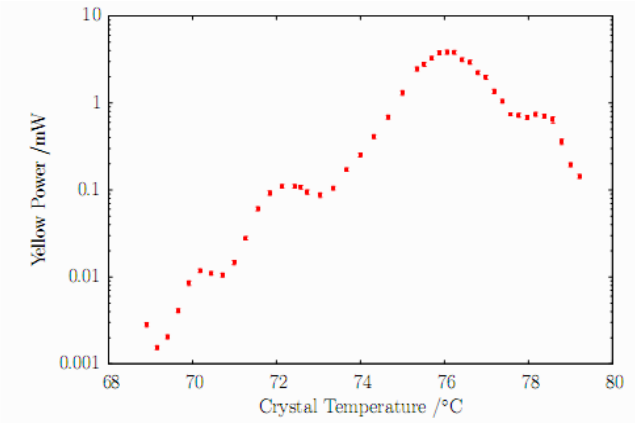


Figure 3. Yellow laser output power vs nonlinear crystal temperature.

The output power is maximized tuning the temperature of the crystal. The power shows a diffraction-like behavior around the phase-matching temperature found at ~76 ◦C, as shown in figure 3. We have generated up to 12 mW of light at 578 nm. Unfortunately the crystal seems to suffer some

This work is supported by Regione Piemonte (YTRO project)

degradation and the maximum power is decreasing slowly. The time instability of the laser intensity is reported in Figure 4. For time intervals up to 100 s, the laser intensity instability is less that 1×10^{-3}.

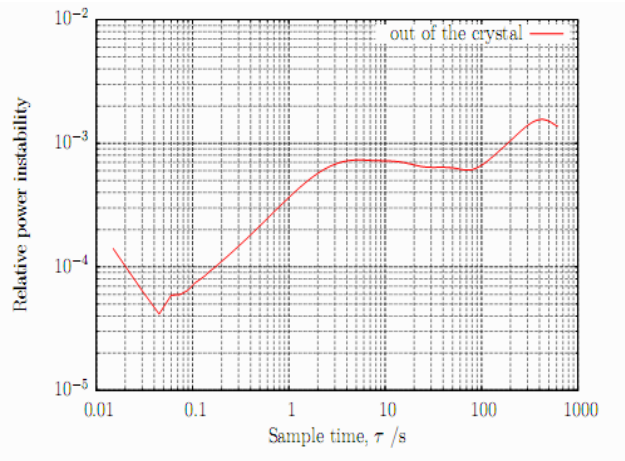


Figure 4. Yellow laser output power temporal stability.

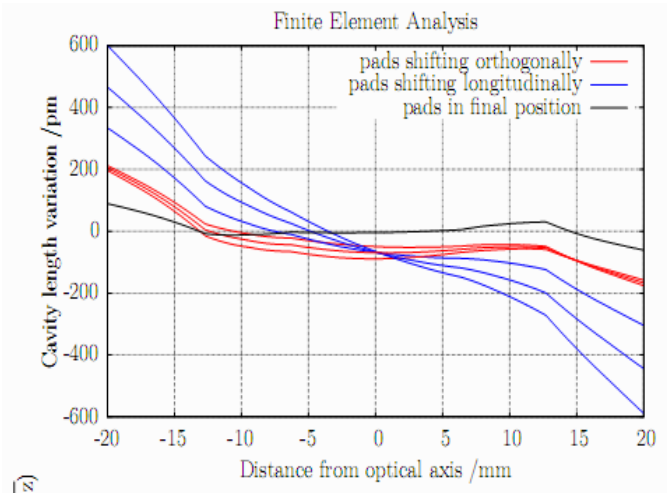


Figure 5. Finite Element Analysis shows a cavity length sensitivity to vibrations of less than 2 pm/g , being g the gravity acceleration.

III. TWO ULTRASTABLE CAVITIES

A couple of ultrastable Fabry-Pérot cavities are used as references to independently stabilize the frequency of the clock lasers. The two cavities are completely made of Corning Ultra Low Expansion Glass (ULE) with a free spectral range of $\Delta\nu_{FSR}$ = 1.5 GHz and a finesse of F1 = 151 000 ± 1000 and F2 = 224 000 ± 4000 respectively, measured with a cavity ring-down experiment.

The ULE spacers are notched, held horizontally and supported by four Viton pads on an aluminum structure. We have performed a finite element analysis to investigate the elastic deformation of the Fabry-Pérot cavity. The final position of the four pads was chosen to minimize the variation of the distance between the two mirrors due to the gravitational acceleration (the cavity is said to be supported at the Airy points).

Each cavity is held in a vacuum chamber (pressure $<2\times10^{-4}$ Pa), placed on top of a passive vibration isolation platform inside an acoustic shielding enclosure.

Finite element analysis shows that the vibration sensitivity of the cavity is less than 2 pm/g , being g the gravity acceleration, that for our cavities (L=10 cm), in terms of relative units means less than 2×10^{-11}s^2/m.

The sensitivity is shown in Figure 5, where we report the cavity length variation when a load of g is applied with respect to distance from the optical axis. Different curves are for different pad positions.

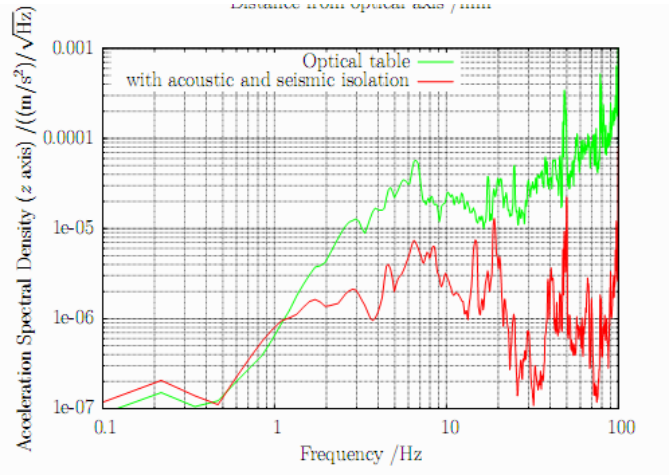


Figure 6. Seismic noise on the acoustic enclosure floor (upper green curve) and on the seismic dumping board (lower red curve).

The noise on the isolation platform was measured with a seismometer. In Figure 6 we report the seismic noise on the acoustic enclosure floor and on the vibration dumping board. The residual seismic noise is compatible with an instability in the laser frequency of 1×10^{-15} @ 1 s when the cavity is held at the Airy points. Thermal isolation is achieved with a copper shield surrounding the cavity. A digital control stabilizes the temperature of the shield, acting on two heaters external to the vacuum chamber. The ULE has a zero in the coefficient of thermal expansion around room temperature. To quickly estimate its value, we give a linear ramp to the temperature of the shield of one of the two cavities, while monitoring the beat note with the other. Aside from an

exponential transient and a time lag, the temperature of the cavity itself should vary linearly. The results are presented in Figure 7. This method allows us to estimate the zero at ~21 ◦C but has some trouble measuring the delay of the cavity temperature (this should translate in a shift of ~0.5 ◦C).

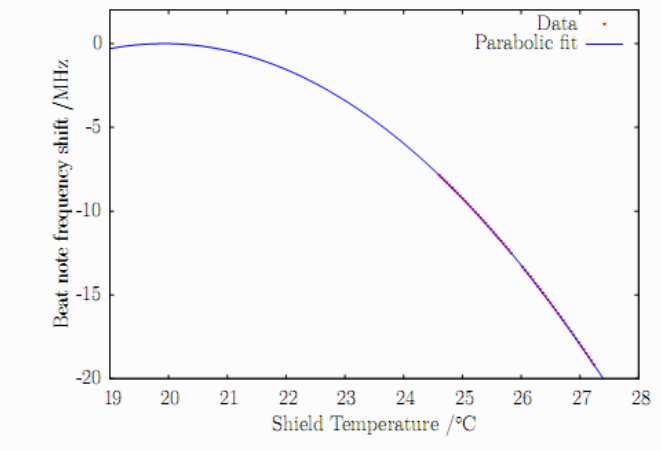


Figure 7. Frequency of the beat note changing one cavity temperature to measure the zero CTE point of the ULE cavity.

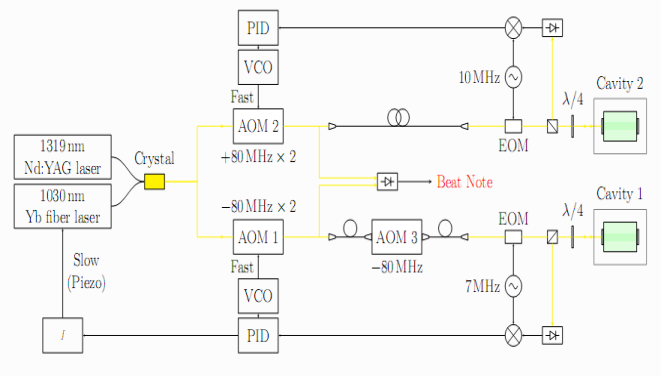


Figure 8. Scheme of the optical bench to lock the lasers on the two ultrastable cavities.

IV. STABILIZATION WITH POUND DREVER HALL TECHNIQUE

The Pound-Drever-Hall technique is used to stabilize the frequency of the yellow laser against the resonances of the cavities. The 578 nm laser at the output of the crystal is split and delivered to the cavities through polarization-maintaining fibers (about 60 µW of laser are impinging on each cavity). Two different beams from the same laser are independently frequency stabilized. In both case an electro-optic modulator (EOM) is used to add sidebands to the carrier of the yellow laser (clearly visible in the transmission). Thus, an optical isolator allows to extract the reflected beam and to obtain the Pound-Drever-Hall error signal.

Two double pass acousto-optic modulators (AOMs) are used for the fast locking of the laser frequency on each cavity. A proportional-integral-derivative (PID) control is used, while a voltage-controlled oscillator (VCO) drives each AOM. The bandwidth of the control is ~ 150 kHz, limited by the delay in the AOMs.

In Figure 8 the optical bench scheme is reported, while Figure 9 shows the transmission signal of the cavity as well as the Pound-Drever-Hall signal.

The Pound-Drever-Hall signal from the first cavity is also used for a slow lock on the piezoelectric actuator of one of the two infrared lasers, in order to correct for the drift. A third AOM is used to bridge the frequency of the two cavities.

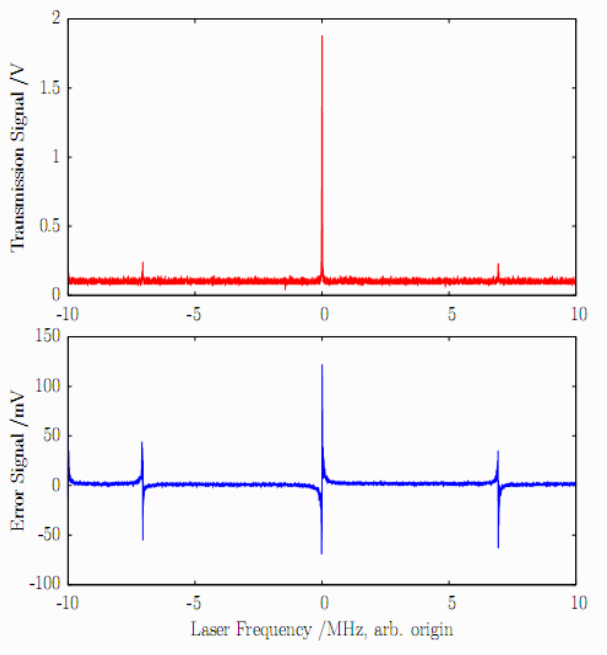


Figure 9. Trasmission and Pound-Drever-Hall signals from the cavities.

V. COMPARISON BETWEEN TWO CAVITIES

The laser beams, stabilized against the two different cavities, are exctracted before the optical fibers delivering them to the anti-acoustic chambers and are superimposed on a fast photodiode. The resulting beat note is then filtered, amplified, down-converted and counted using an hydrogen maser as a

reference. The noise of the two lasers is independent within the bandwidth of the two controls. The noise of the two lasers is independent within the bandwidth of the two controls.

Figure 10 shows the beat note at the spectrum analyzer allowing the measurement of the bandwidth, that is about 150 kHz. Figure 11 presents just the close up scan of the beat note, that allows to estimate a lorentzian linewidth of (3.8 ± 0.1) Hz.

carrier, that has a linewidth of 3.8 Hz (linewidth of the real beat note of the two cavities).

The beat note is then down converted to about 10 MHz and its instability is measured with a phasometer: results are presented in Figure 12. Here, we show the instability of the beat note when the dumping board is not operated (upper green line), showing a large degradation in the short term behavior as expected ($1/\tau$ instability). Then, the red line presents the instability of the lasers when the dumping board is under operation: the short term is much lower than in the previous case, and a linear drift is observed as large as 0.8 Hz/s. The linear drift removal confirms that the frequency noise between 0.5 s and 20 s is mainly flicker.

The frequency flicker floor is at 4×10^{-15}, probably limited by AM to FM noise conversion.

We expect the limit from the thermal noise, dominated by the thermal fluctuation in the ULE mirrors substrate in the cavities to be between 1.5×10^{-15} and 3×10^{-15}.

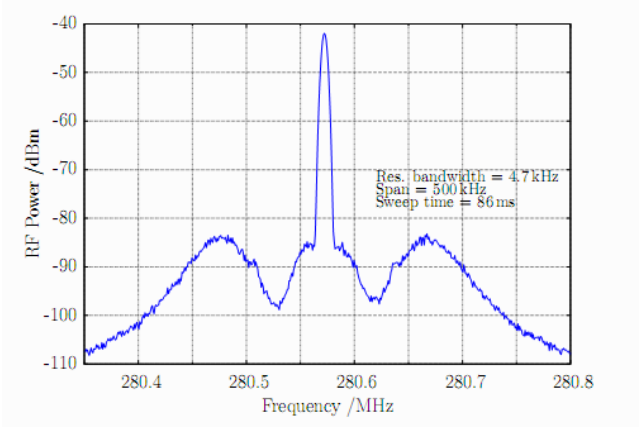


Figure 10. Beat note between the two independent locked lasers, showing the control bandwidth (150 kHz).

VI. CONCLUSIONS

We have realized a laser source at 578 nm for an ytterbium lattice clock capable of 12 mW and a frequency instability $<4 \times 10^{-15}$ corresponding to a lorentzian linewidth <4 Hz. This result is close to the expected thermal limit of the ULE mirror substrate, and we think that the excess of noise is due to AM to PM conversion. Further measurements and the implementation of a power stabilization servo will be useful to reach the thermal noise limit.

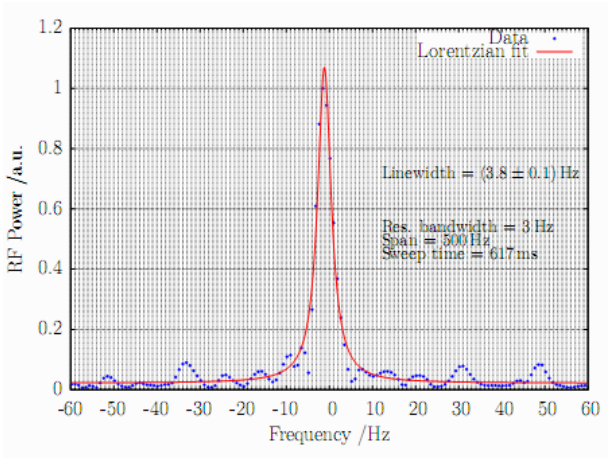
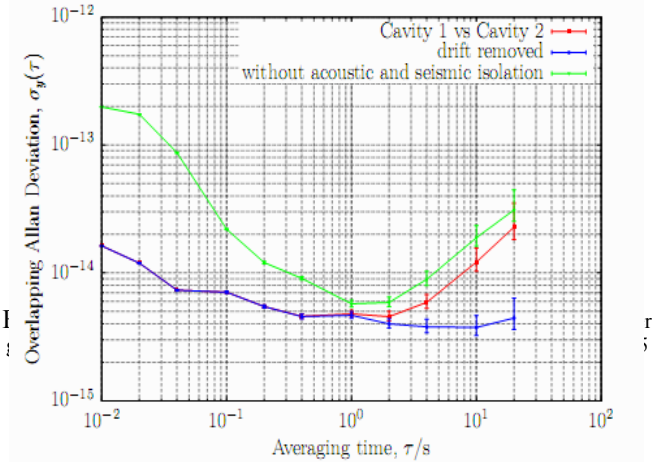


Figure 11. Beat note between the two independent locked lasers, showing a linewidth of 3,8 Hz.



REFERENCES

[1] N. D. Lemke, A. D. Ludlow, Z. W. Barber, T. M. Fortier, S. A. Diddams, Y. Jiang, S. R. Jefferts, T. P. Heavner, T. E. Parker, and C. W. Oates, "Spin ½ Optical Lattice Clock", Phys. Rev. Lett. 103, 063001, 2009.

[2] C. W. Oates, "Stable Laser System for Probing the Clock Transition at 578 nm in Neutral Ytterbium", Frequency Control Symposium, 2007 Joint with the 21st European Frequency and Time Forum. IEEE International, Geneve, 29 May -1 June 2007, pp 1274

[3] K. Hosaka et al., "Evaluation of the clock laser for an Yb lattice clock using an optic fiber comb", IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol 57, p 606, 2010

[4] A. Yu. Nevsky et al., "A narrow-line-width external cavity quantum dot laser for high-resolution spectroscopy in the near-infrared and yellow spectral ranges", Appl Phys B, 92: 501, 2008 .

Measurement of Radiative Decay and Cold Collision Trap Losses in Laser-Cooled Ytterbium

M. Zoppi, G.A. Costanzo and M. Pizzocaro
Politecnico di Torino
Torino, Italy.

D. Calonico, F. Levi, L. Lorini,
E.K. Bertacco and A. Godone
Optics Division
I.N.Ri.M.
Torino, Italy

***Abstract*—Ytterbium lattice frequency standard is an outstanding optical clock with a demonstrated relative frequency accuracy of $3.4 \cdot 10^{-16}$, presently limited by the uncertainty on the blackbody radiation (BBR), lattice polarizabilities and cold collisions shifts. In particular the BBR shift is looming as a major component in the uncertainty budget of many optical frequency standards. In fact the uncertainty on BBR systematic shift in Ytterbium atoms is due both to experimental and theoretical limits in the knowledge of the atomic system: presently Ytterbium has the largest theoretical uncertainty on the BBR shift. Then, both for atomic physics and metrology, new measurements of the Ytterbium atomic properties are strongly demanded, such as the determination of transition matrix elements, lifetimes, branching ratios, polarizabilities and hyperfine constants. In this work we present a set of new measurements for the radiative decay and for the cold collisions trap losses in laser-cooled Ytterbium atoms. These data are relevant to the blackbody radiation shift evaluation and to the cold collisions shift theory.**

I. INTRODUCTION

In recent years microwave frequency standard performances based on Cs atoms have been overcome by optical clocks which permit to obtain an "almost" factor 10 of improvements of the metrological properties of the standards. In fact several scientific groups [1-7] have shown fundamental results on the accuracy and the stability of optical clocks based on a set of alkaline-earth and alkaline-earth like atoms.

Among them Yb frequency standards are one of the best candidates for the possible re-definition of the Second, the Time unit in the International System: the astonishing results of the group of Lemke [7], showing a relative accuracy of few parts in 10^{16}, are pushing the scientific community to a more active research on this fields tackling the Yb atoms properties for a better knowledge. A lack in Yb atomic property data limits both evaluation of relative frequency accuracy in Yb optical frequency standards and quantum information experiments. For this reasons more experimental data are needed in order to improve the level of knowledge of this atom.

Thanks to the atomic structure of the Yb atom, the possibility of a deeper comprehension of the Yb atomic level structure can be accomplished by comparison of the different isotope dynamics taking advantages of the laser cooling possibility. In fact Yb can be laser-cooled down to the mK regime with the $^1S_0 \rightarrow {}^1P_1$ (399nm) transition, with the natural linewidth of 29MHz and a Doppler limit temperature of about 690μK.

When the atomic sample reaches the cold regime, a set of interesting parameters, useful for frequency standard applications, can be measured. Among them the measurement of the cold collision coefficient β can clarify the collisional shift theory for Yb optical clocks, improving their accuracy. Moreover the measurements of the so-called D-branching ratio $a_{2,0}$ (which is defined as the de-excitation rate of atoms from the $^1P_1$ state to the ^3P triplet, via the ^3D states) allow to improve the theoretical evaluation of the blackbody radiation shift in Yb optical clocks.

II. EXPERIMENTAL SETUP

In our experimental set up, the Yb source is an effusion oven operating between 360 and 390°C with a nozzle realized with a microchannel bundle, maintained at a temperature of 400°C. The latter is made by 200 microtubes in Monel400 with 100μm of internal diameter, 10mm long, that implies a thermal beam divergence of 17mrad.

The physical structure is maintained under Ultra High Vacuum by two ion pumps which ensure a differential vacuum: the trap region is maintained at a pressure of 10^{-6}Pa while the oven region pressure is 10 times higher. The MOT region is placed about 50cm from the nozzle, where the thermal beam has a diameter of about 20mm. The Zeeman

This work is supported by Regione Piemonte (YTRO project).

slower is not used in order to simplify the structure and avoid any stray magnetic field in the clock implementation. Yb atoms interact with three retro-reflected laser beams in a classical XYZ (0,0,1) configuration.

Each beam has 4mm (1/*e*) diameter and is red detuned of about 30MHz by means of two double-pass acousto-optic modulator systems. The MOT magnetic static gradient is generated by two anti-Helmholtz coils that provide a magnetic field gradient of 0.12T/(m·A). The trapped atom cloud is approximately spherical and has a diameter of 1mm (1/*e*). Atomic fluorescence light is collected by an EMCCD camera placed outside the vacuum chamber, 25cm far from the interaction region.

In order to increase the number of atoms in the trap an additional laser cooling beam has been added, counter-propagating to the Yb atomic beam: this beam works at a fixed frequency, which is 140MHz red detuned from the atomic resonance of the 399nm transition for each isotope.

The laser radiation is provided by the 798nm radiation obtained from a Ti-Sa laser frequency doubled using a 20mm long PPKTP crystal placed in an enhancement ring cavity. Up to 70mW of 399nm radiation have been obtained from a 600mW infrared radiation. The resonant ring cavity is then locked to the Ti-Sa laser using the Hänsch-Couillaud technique. The typical power of each laser beam is about 3mW: about $3\cdot10^4$ atoms of ^{174}Yb are trapped without the pre-cooler.

III. Evaluation of the Two-Body Collision Coefficient

In order to measure the β coefficient, the usual time evolution of the trapped atoms number *N(t)*, described by

$$\frac{dN(t)}{dt} = R - \alpha \cdot N(t) - \beta \iiint n^2(\vec{r},t)d^3\vec{r} \qquad (1)$$

must be analyzed.

In (1), *R* is the atom capture rate, α the linear loss rate of the trap, β the two-body collision coefficient and *n* the spatial density of the trap. If the constant volume regime is assumed and the spatial extension of the MOT is approximated as a sphere with a radius *a*, the atom spatial density is described by $n(r,t)=n_0(t)\cdot\exp(-r^2/a^2)$ and the non-linear second order differential equation is evaluated. The latter is a particular case of the more general class of Riccati differential equations:

$$\begin{cases} \dfrac{dN(t)}{dt} = R - \alpha \cdot N(t) - \dfrac{\beta}{(\sqrt{2\pi}a)^3} \cdot N^2(t) \\ N(0) = N_{in} \end{cases} \qquad (2)$$

Once fixed the boundary condition this equation has only one analytical solution. From (2) we introduce a different and new method for β evaluations, strongly different from methods reported in [8-10]. As in [11], we use a loading/decay technique where both processes are measured in the same experimental setup but, compared to the method developed in [11], without the pre-cooler beam during the acquisition.

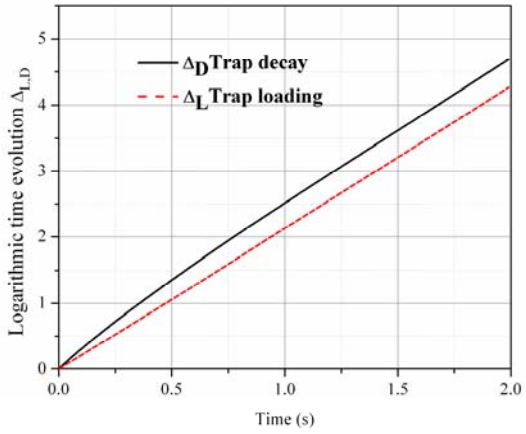


Figure 1. Simulated logarithmic time evolution of the trapped atoms number during trap loading (Δ_L, dashed line) and decay (Δ_D, straight line) assuming $\beta = 4\cdot10^{-15}\mathrm{m^3s^{-1}}$, $a = 0.7$mm, $\alpha = 1\mathrm{s^{-1}}$, $N_{in} = 10^6$ and $R = 10^6\mathrm{s^{-1}}$.

This experimental choice enables us to develop an exact theoretical model reducing the uncertainty on the final result. As a matter of fact we introduce the functions Δ_L and Δ_D defined as:

$$\Delta_{L,D}(t) \equiv -\ln\frac{N_{L,D}(\infty) - N_{L,D}(t)}{N_{L,D}(\infty) - N_{L,D}(0)} \qquad (3)$$

where Δ_L and Δ_D concern the loading and the decay process.

Finally in (3) $N_{L,D}(\infty)$ is the asymptotic value of trapped atoms in the loading/decay processes. As shown in Fig. 1, we generally observe a linear behavior, but close to the beginning of the processes a non linearity enables to distinguish the two situations: this arises because of the different number of atoms, with consequently a different loss rate due to the collisions between the cold atoms.

If we consider the difference $\Delta_{log} = \Delta_L - \Delta_D$, whose behavior is reported in Fig. 2, it is possible to evaluate some important parameters of the atomic sample. The evaluation of the derivative in the arbitrary time origin and the asymptotic value of Δ_{log} leads to a measurement of α and β coefficients using (4). In fact from the solution of (2), it can be easily obtained:

$$\left.\frac{d\Delta_{\log}(t)}{dt}\right|_{t=0} = kN_{Max} \qquad (4)$$

$$\lim_{t\to\infty}\Delta_{\log}(t) = \ln\left(1 + \frac{2kN_{Max}}{a+\varsigma}\right)$$

where $k \equiv \beta/\left(\sqrt{2\pi}a\right)^3$ and $\varsigma \equiv \sqrt{4kR+\alpha^2}$; N_{max} and a are directly measurable using the EMCCD camera imaging, while R is evaluated as the slope in the origin of the solution (4) for N_{in}=0.

IV. Experimental results

The only measurement of the two-body cold collision coefficient β actually available in literature for Yb is accomplished in [12] using ^{174}Yb atoms in the level $^3P_2$ with estimated temperature of 41μK.

In the present work we use laser cooled Yb atoms of different isotopes with a temperature of few mK. While the loading curve is obtained by chopping the six laser beams with a mechanical shutter within closing time lower than 100μs, the decay curve is obtained by switching on-off the RF of the acousto-optic modulator which red detunes the pre-cooler beam of 140 MHz from the atomic resonance and only taking the values of the fluorescence when the RF is off. Each measure is the average of about 30 acquisitions reducing the laser noise.

The image analysis by EMCCD camera signal shows that the spatial density distribution of the atoms in the trap is well described by a gaussian function with a radius of about 0.4mm, depending on the isotope.

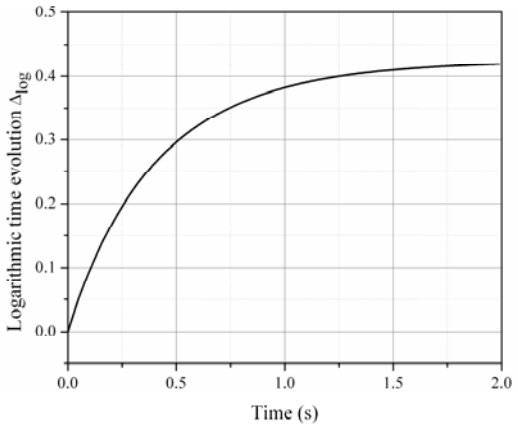


Figure 2. Simulated time evolution of Δ_{log} using the same parameters in the previous figure.

In Figs. 3 and 4 the measurements for ^{174}Yb are shown with a laser trap intensity of 75% of the saturation value. The results for a single run acquisition are reported in Tab. I, where β is obtained solving the equations system (4).
The values of τ_L and τ_D reported in Table I are the loading and decay times of the experimental curves for different isotopes, obtained from an exponential fit of the data. The coefficients reported in the Table I are the first measurement of β for Yb in the mK temperature regime.

V. Branching ratio measurements

In order to estimate the branching ratio $a_{2,0}$, we developed the following simple model: if we neglect the non-linear term in (1), the solution becomes:

$$N(t) = \frac{R}{\alpha}(1-e^{-\alpha t}) \tag{5}$$

This approximation is acceptable if we consider only the loading curve, where the number of the trapped atoms is lower and the non-linear contribution is negligible. In [13-14] the evaluation of $a_{2,0}$ is performed fixing the isotope and changing the laser trap intensity. The latter changes the fraction f of the excited atoms.

The linear loss rate α can factorized in two main contributions as $\alpha = \alpha_b + a_{2,0}\, f$. The first term is the loss due to the collisions between the trapped atoms and the atoms in the thermal beam. The second is the loss due to the atoms decay from the excited state to the metastable ^{3}P triplet.

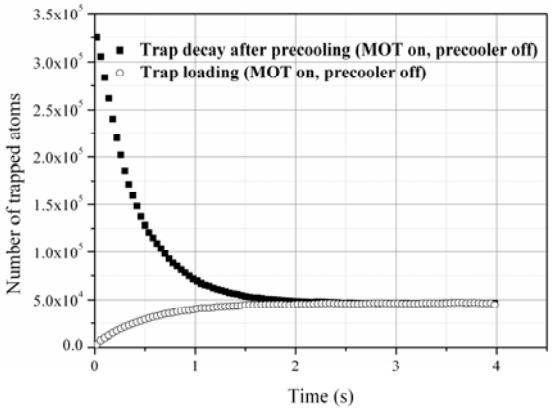


Figure 3. Experimental number of trapped atoms during MOT loading without pre-cooler (empty circles) and during the decay from the pre-cooler switch off to the MOT regime (full squares).

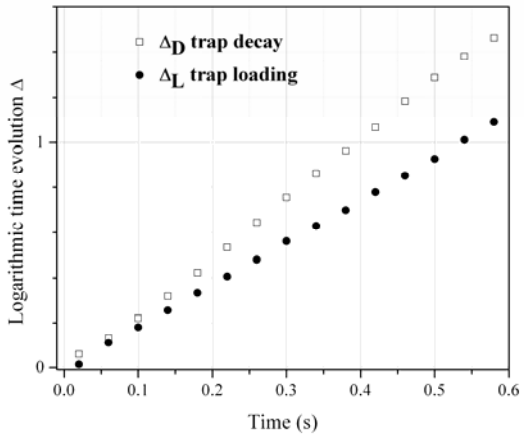


Figure 4. Experimental time evolution of Δ_D (empty squares) and Δ_L (full circles) during trap loading and decay for the ^{174}Yb isotope.

With this assumption $a_{2,0}$ is the D-branching ratio. In this work $a_{2,0}$ is obtained fixing f and comparing couples of loading times of different isotopes to obtain a linear systems with two unknown quantities, L and $a_{2,0}$. This method implies to assume both of the last two parameters constant over all the isotopes and writing α as in the following equation:

$$\alpha_{isot} = \eta_{isot} L + a_{2,0} f \quad (6)$$

where L is a loss term constant over all isotopes, due to collisions between trapped atoms and thermal beam and η is the relative abundance for each isotope. The subscript *isot* denotes quantities that vary for each isotope, while $a_{2,0}$ is assumed to be constant for the whole set of isotopes. The experimental results, in good agreement with [13], are reported in Table II.

TABLE I.

β experimental results					
Isotope	***τ_L (ms)***	***τ_D (ms)***	***a (mm)***	***N_{max} (10^5)***	***β ($10^{-15} m^3 s^{-1}$)***
171	594(5)	502(5)	0.33(5)	1.4(0.8)	1.3(0.8)
174	521(4)	396(1)	0.46(5)	3.4(1.9)	3.6(1.9)

TABLE II.

***$a_{2,0}$* experimental results**			
This work	***Exp. Ref. [13]***	***Exp. Ref. [14]***	***Exp. Ref. [15]***
25(3) s^{-1}	23(11) s^{-1}	13(3) s^{-1}	7(5) s^{-1}

VI. CONCLUSIONS

We have developed a new method for the measure of the two-body collision coefficient β and the branching ratio $a_{2,0}$ of Yb atoms in the mK cold collisions regime. These theoretical methods have been then implemented for laboratory experiments allowing us to improve the knowledge of the fundamentals data of the Yb atoms.

REFERENCES

[1] C. W. Hoyt, Z. W. Barber, C. W. Oates, T. M. Fortier, S. A. Diddams, and L. Hollberg, Phys. Rev. Lett. 95, 083003, 2005.

[2] M. M. Boyd, A. D. Ludlow, S. Blatt, S. M. Foreman, T. Ido, T. Zelevinsky, and J. Ye, "^{87}Sr Lattice Clock with Inaccuracy below 10^{-15}", Phys. Rev. Lett. 98, 083002, (2007).

[3] A. D. Ludlow, T. Zelevinsky, G. K. Campbell, S. Blatt, M. M. Boyd, M. H. G. de Miranda, M. J. Martin, J. W. Thomsen, S. M. Foreman, Jun Ye, T. M. Fortier, J. E. Stalnaker, S. A. Diddams, Y. Le Coq, Z. W. Barber, N. Poli, N. D. Lemke, K. M. Beck, and C. W. Oates, "Sr Lattice Clock at $1 \cdot 10^{-16}$ Fractional Uncertainty by Remote Optical Evaluation with a Ca Clock", Science 319, 1805 (2008).

[4] X. Baillard, M. Fouché, R. Le Targat, P. G. Westergaard, A. Lecallier, F. Chapelet, M. Abgrall, G. D. Rovera, P. Laurent, P. Rosenbusch, S. Bize, G. Santarelli, A. Clairon, P. Lemonde, G. Grosche, B. Lipphardt, and H. Schnatz, "An optical lattice clock with spin-polarized ^{87}Sr atoms", Eur. Phys. J. D 48, 11, (2008).

[5] Ch. Lisdat, J. S. R. Vellore Winfred, T. Middlemann, F. Riehle, and U. Sterr, "Collisional Losses, Decoherence, and Frequency Shifts in Optical Lattice Clocks with Boson", Phys. Rev. Lett. 103, 090801, (2009).

[6] M. Petersen, R. Chicireanu, S.T. Dawkins, D.V. Magalhaes, C. Mandache, Y. Le Coq, A. Clairon, and S. Bize, "Doppler-Free Spectroscopy of the $^1S_0$-$^3P_0$ Optical Clock Transition in Laser-Cooled Fermionic Isotopes of Neutral Mercury", Phys. Rev. Lett. 101, 183004 (2008).

[7] N. D. Lemke, , A. D. Ludlow, Z. W. Barber, T. M. Fortier, S. A. Diddams, Y. Jiang, S. R. Jefferts, T. P. Heavner, T. E. Parker, and C. W. Oates, "Spin-1/2 Optical Lattice Clock", Phys. Rev. Lett. 103, 063001, (2009).

[8] J. Weiner, V. S. Bagnato, S. Zilio, and P. S. Julienne, "Experiments and theory in cold and ultracold collisions", Rev. Mod. Phys. vol. 71, no. 1, (1999) and references therein.

[9] A. Browaeys, J. Poupard, A. Robert, S. Nowak, W. Rooijakkers, E. Arimondo, L. Marcassa, D. Boiron, C. I. Westbrook, and A. Aspect, "Two body loss rate in a magneto-optical trap of metastable He", Eur. Phys. J. D vol. 8, no. 2, pp. 199-203, (2000).

[10] T. P. Dinneen, K. R. Vogel, E. Arimondo, J. L. Hall, and A. Gallagher, "Cold collisions of Sr*-Sr in a magneto-optical trap", Phys. Rev. A vol. 59, pp. 1216-1222, (1999).

[11] R. L. Cavasso-Filho, A. Scalabrin, D. Pereira, and F. C. Cruz, "Observing negligible collision trap losses: the case of alkaline-earth metals", Phys. Rev. A vol. 67, 021402R, (2003).

[12] A. Yamaguchi, S. Uetake, D. Hashimoto, J. M. Doyle, and Y. Takahashi, "Inelastic collisions in optically trapped ultracold metastable ytterbium", Phys. Rev. Lett. vol. 101, 233002, (2008).

[13] T. Loftus, J. R. Bochinski, R. Shivitz, and T. W. Mossberg, "Power-dependent loss from an ytterbium magneto-optic trap", Phys. Rev. A vol. 61, 051401R, (2000).

[14] U. D. Rapol, A. Krishna, A. Wasan, and V. Natarajan, "Laser cooling and trapping of Yb from a thermal source", Eur. Phys. J. D vol. 29, 409, (2004).

[15] S. G. Porsev, Yu. G. Rakhlina, and M. G. Kozlov, "Electric-dipole amplitudes, lifetimes, and polarizabilities of the low-lying levels of atomic ytterbium", Phys. Rev. A vol. 60, 2781, (1999).

Quantum Cascade Laser: a promising candidate toward an ultra-stable laser in the mid-IR

L. Tombez, J. Di Francesco, S. Schilt, G. Di Domenico, D. Hofstetter and P. Thomann
LTF Laboratoire Temps-Fréquence
Institut de Physique, Université de Neuchâtel
Neuchâtel, Switzerland
lionel.tombez@unine.ch

***Abstract*—The frequency noise properties of free-running single-mode mid-IR Quantum Cascade Lasers (QCLs) operated near room-temperature in CW mode and emitting at 4.6 µm are presented. The frequency noise level measured in this paper is globally a factor of 100 lower than recently reported for a similar QCL but operated at cryogenic temperature. The results are very promising for the future realization of a QCL-based ultra-stable laser in the mid-IR.**

I. INTRODUCTION

Narrow-linewidth lasers exhibiting a high spectral purity have important applications in various fields, such as high-resolution spectroscopy, coherent optical communications or time and frequency metrology. In the latter field, such lasers can be used as ultra-stable reference for low phase noise microwave generation when frequency-stabilized to a high-finesse optical cavity for linewidth narrowing. Indeed, the relative stability of the ultra-stable optical reference can be transferred to the microwave domain using an optical frequency comb (OFC). Such optical references were demonstrated based on external cavity and fiber lasers in the near-infrared (NIR) [1], but also in the mid-infrared (MIR) with a 3.39-µm HeNe laser stabilized to a CH_4 transition [2]. In this latter case, the relative stability of the HeNe/CH_4 laser was transferred using a MIR comb obtained by difference frequency generation (DFG) between two spectral parts of a NIR comb. As a result of the DFG process, the generated MIR comb has a null carrier-envelope offset (CEO) frequency, independently of the value of the NIR comb CEO. In other words, the translational degree of freedom of the NIR comb can stay free without affecting the generated MIR comb. Therefore, the CEO does not need to be stabilized nor to be measured, which relaxes the need for an octave spanning comb and the associated f-to-2f interferometer.

With the objective to implement a similar optical-to-microwave transfer based on a MIR quantum cascade laser (QCL), we studied the frequency-noise characteristics of QCLs, aiming at assessing their suitability as a potential ultra-stable frequency reference. Indeed, efficient laser linewidth narrowing by frequency-stabilization to a high-finesse cavity with a reasonable feedback-loop bandwidth can be achieved only if the frequency noise of the free-running laser is sufficiently low. An estimate of feedback loop bandwidth required for the realization of an ultra-narrow linewidth QCL will be assessed from our experimental results.

II. EXPERIMENTAL SETUP

We investigated the frequency-noise properties of free-running commercial single-mode DFB-QCLs (Alpes Lasers SA, Switzerland) operated in CW mode near room temperature and emitting at 4.6 µm. The frequency noise measurement was performed by converting the laser frequency fluctuations into intensity fluctuations using the side of a molecular absorption in standard single-pass direct absorption spectroscopy. The emission wavelength of the laser was tuned to the centre of the linear range in the flank of the 250-MHz wide R(14) transition in the fundamental vibration band of carbon-monoxide (CO). The laser was driven (350 mA at 10V) by a home-made ultra-low noise current source (<350pA/Hz$^{1/2}$) and temperature-stabilized with a one-stage Peltier cooler at 277 K. At this operating temperature, the output power was on the order of 5 mW. The beam was collimated with an aspherical ZnSe lens and traversed a 1-cm long absorption gas cell filled with pure CO at a pressure of 20 mbar. The resulting absorption line is shown in Fig 1 and compared to the absorption profile simulated from Hitran

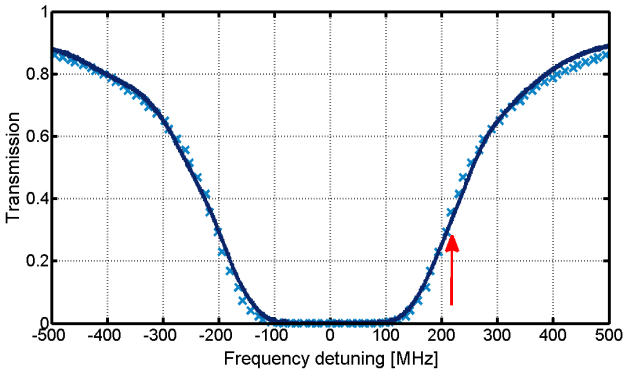


Fig. 1 Carbon-monoxide absorption line used as frequency discriminator (20 mbar, 1-cm pathlength, pure CO). The simulation from HITRAN database is also shown (crosses). The arrow represents the laser frequency operating point.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

database [3]. Another ZnSe lens was used to focus the transmitted light onto a low-noise liquid-nitrogen-cooled HgCdTe detector with 20-MHz bandwidth. The detected signal was amplified and measured with a fast Fourier transform analyzer for low Fourier frequencies (up to 100 kHz) and with an electrical spectrum analyzer at higher frequencies.

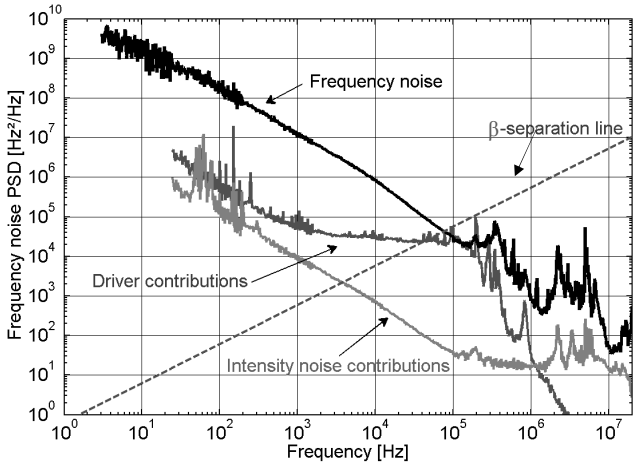


Fig. 2 Frequency noise PSD of the free-running DFB-QCL operated in CW mode near room temperature (T=277 K, I=350 mA, I_{th}=290 mA, U=10 V, P_{op}=6 mW).

III. RESULTS

The measured frequency noise power spectral density (PSD) revealed a 1/f-like behavior, with a noise level of $2x10^8$ Hz^2/Hz at 100 Hz dropping down to a very low level below 100 Hz^2/Hz at 10 MHz (Fig. 2). The contributions of the driver current noise and laser intensity noise are also shown in Fig. 2 and are both well below the measured frequency noise (up to 100 kHz Fourier frequency), which can therefore be attributed to the laser itself. Above 100 kHz, one can observe additional technical noise that deteriorates the measurement, which probably originates from the current driver since similar peaks are observed in the intensity noise spectrum. Despite this technical noise, the frequency noise decreases significantly up to 10 MHz. Moreover, above 10 kHz, it seems to deviate from the 1/f trend observed at low frequency into a steeper slope of $1/f^{3/2}$. However, one cannot exclude that an even steeper behavior is not hidden in that noisy part of the spectrum as previously reported for a cryogenic QCL [5]. As our goal was to assess the required feedback loop bandwidth required for linewidth reduction, the part of the spectrum above 100 kHz is not of prime concern since it lies below the β-separation line also presented in Fig. 2, which is relevant for the determination of the linewidth and thus for the required feedback bandwidth [4]. It is however interesting to note that the white frequency noise level is not reached below 10 MHz and only an upper limit of ~100 Hz^2/Hz can be inferred from our measurement, corresponding to an intrinsic linewidth of ~ 300 Hz. While the intrinsic linewidth is a fundamental limit in a laser, the real observed linewidth is highly broadened by the 1/f noise. The laser FWHM linewidth can be calculated from the frequency noise PSD [4] and depends on the observation time in presence of 1/f noise. Fig. 3a shows the calculated FWHM linewidth of the laser versus the observation time; for instance, a 550-kHz linewidth is obtained for 5 ms observation time. This value agrees very well with the heterodyne beat spectrum linewidth of ~1 MHz measured between two identical QCLs in a complementary experiment (Fig. 3b), since the beat lineshape corresponds to the convoluted lineshapes of the two identical QCLs. Finally, the measured PSD also shows that a feedback loop bandwidth on the order of 100 kHz would be sufficient for a high reduction of the laser linewidth. Because of the low frequency noise measured on this laser, this can be achieved with standard PID servo-controllers.

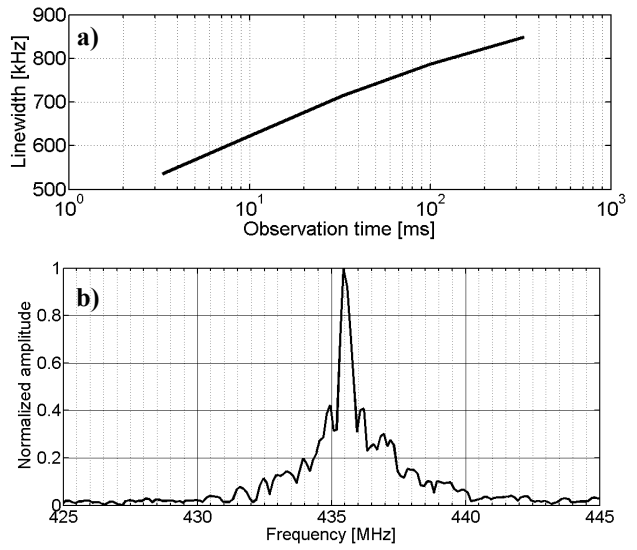


Fig. 3 (a) Laser FWHM linewidth calculated from the frequency noise PSD vs. observation time. (b) Heterodyne beat signal measured between two identical 4.6-µm room temperature QCLs (4 ms sweep time).

Frequency noise properties of QCLs are not widely known and complete PSD characterizations were reported only by a few research groups and only for cryogenic devices so far [5-6]. Interestingly, our results are a factor of 100 lower than data recently reported by Bartalini *et al*. [5] for a similar laser (4.3µm DFB-QCL from Alpes Lasers) but operated at cryogenic temperature. This significantly lower frequency noise is believed to be a consequence of the difference in the laser temperature operation conditions [7], but further investigations are required to fully understand this different behaviour. On the other hand, our results are on the same order of magnitude than what was measured by Myers *et al*. [6], but with a QCL emitting at a considerably longer wavelength of 8.5 µm.

IV. CONCLUSION

We have investigated the frequency noise properties of free-running commercial single-mode DFB QCLs operated at 277 K. The measured frequency noise indicated that ultra-narrow linewidth operation can be achieved by frequency-stabilization with a moderate feedback loop bandwidth on the order of 100 kHz, which is very promising for the future realization of an ultra-stable reference in the mid-IR range, based on a QCL.

V. ACKNOWLEDGMENT

This research is supported by the Swiss National Science Foundation, the Swiss Confederation Program Nano-Tera.ch scientifically evaluated by the SNSF and was initiated by the Gebert-Ruef Foundation in Basel, Switzerland. We would like to thank Prof. Sigrist and M. Gianella from ETHZ for providing the gas cell and Stéphane Blaser (AlpesLasers) for fruitful discussions and information about the lasers.

REFERENCES

[1] J. Millo, M. Abgrall, M. Lours, E.M.L., English, H. Jiang, J. Guéna, A. Clairon, M.E. Tobar, S. Bize, Y. Le Coq and G. Santarelli, "Ultralow noise microwave generation with fiber-based optical frequency comb and application to atomic foutnain clock", Appl. Phys. Lett. 94 (14), p.141105 (2009)

[2] S. M. Foreman, A. Marian, J. Ye, E. A. Petrukhin, M. A. Gubin, O. D. Mücke, F. N. C. Wong, E. P. Ippen, and F. X. Kärtner, "Demonstration of a HeNe/CH_4-based optical molecular clock", Opt. Lett. **30**, 570 (2005)350.

[3] L.S. Rothman et al, "The *HITRAN* 2008 molecular spectroscopic database", J. Quant. Spectrosc. Radiat. Transfer **110**, 533 (2009)

[4] G. Di Domenico, S. Schilt and P. Thomann., "Simple approach to the relation between laser frequency noise and laser line shape", App. Opt. 49, No. 25, pp. 4801-4807 (2010)

[5] S. Bartalini, S. Borri, P. Cancio, A. Castrillo, I. Galli, G. Giusfredi, D. Mazzotti, L. Gianfrani, and P. De Natale, "Observing the Intrinsic Linewidth of a Quantum-Cascade Laser: Beyond the Schawlow-Townes Limit", Phys. Rev. Lett. **104**, pp. 083904 (2010)

[6] T. L. Myers, R. M. Williams, M. S. Taubman, C. Gmachl, F. Capasso, D. L. Sivco, J. N. Baillargeon, and A. Y. Cho, "Free-running frequency stability of mid-infrared quantum cascade lasers", Opt. Lett. 27(3), pp. 170–172 (2002)

[7] L. Tombez, J. Di Francesco, S. Schilt, G. Di Domenico, J. Faist, P. Thomann, and D. Hofstetter, "Frequency noise of free-running room temperature DFB quantum cascade lasers", in preparation

Stability limits of an optical frequency standard based on free Ca atoms

J. A. Sherman and C. W. Oates
NIST
325 Broadway
Boulder, CO 80027
Email: jeff.sherman@nist.gov

***Abstract*—We have quantified a short term instability budget for an optical frequency standard based on cold, freely expanding calcium atoms. Such systems are the subject of renewed interest due to their high frequency stability and relative technical simplicity compared to trapped atom optical clocks. By filtering the clock laser light at 657 nm through a high finesse cavity, we observe a slight reduction in the optical Dick effect caused by aliased local oscillator noise. The ultimately limiting technical noise is measured using a technique that does not rely on a second clock or fs-comb.**

I. Introduction

Optical frequency standards based on trapped atoms and ions offer unmatched absolute uncertainties [1], [2]. However, a clock based on cold but untrapped atoms may still offer competitive levels of short term stability while remaining technically simpler and more readily commercialized. Here we discuss limits to the stability of an optical standard featuring cold, freely expanding, Ca atoms [3].

Briefly, atoms are collected in a magneto-optical trap (MOT) by use of the strong ${}^1S_0 \leftrightarrow {}^1P_1$ transition at 423 nm. Though this MOT period lasts only 2.8 ms, many atoms are retrapped between clock cycles, yielding a steady-state atom number of order 10^6. An external cavity diode laser at 657 nm, stabilized to a ULE glass Fabry-Perot cavity (linewidth $\sim$ 9 kHz), is made resonant with the allowed inter-combination transition ${}^1S_0 \leftrightarrow {}^3P_1$. The MOT is shuttered for 0.44 ms, and a sequence of 657 nm pulses forms a Ramsey-Bordé atom interferometer [4] in the freely falling and expanding atoms. Counter-propagating, horizontally aligned pairs of Ramsey pulses suppress residual first-order Doppler shifts, enabling high resolution, and high S/N, spectroscopy [5], [6]. Resonant 423 nm light is used to measure the atoms' ground state population before and after the clock pulse sequence. Dividing these collected fluorescence signals yields a spectroscopic feature normalized against atom number fluctuations with a linewidth (FWHM) of $\sim$ 1 kHz, and signal-to-noise ratio (S/N) of about 50 in 3.3 ms. The clock instability limit implied by these results is about $2 \times 10^{-15}(\tau/1\mathrm{s})^{-1/2}$; which is roughly consistent with the quantum projection noise limit given a coherent fringe contrast of $\sim 10\%$.

We measure the Ca clock instability to be consistently $3 - 4 \times 10^{-15}(\tau/1\mathrm{s})^{-1/2}$ when compared against a superior local optical standard [7] via a fs-comb. We calculate that the optical Dick effect [8], [9] contributes $3.6 \times 10^{-15}(\tau/1\mathrm{s})^{-1/2}$ to this instability. About half of the Dick effect is eliminated by filtering the clock laser through a high-finesse optical cavity to remove high frequency local oscillator noise. We measure a comparable amount of instability, $1.8 \times 10^{-15}(\tau/1\mathrm{s})^{-1/2}$, due to technical sources. In this paper we discuss these contributions and highlight measurement techniques used to quantify them with minimal use of a second clock or fs-comb.

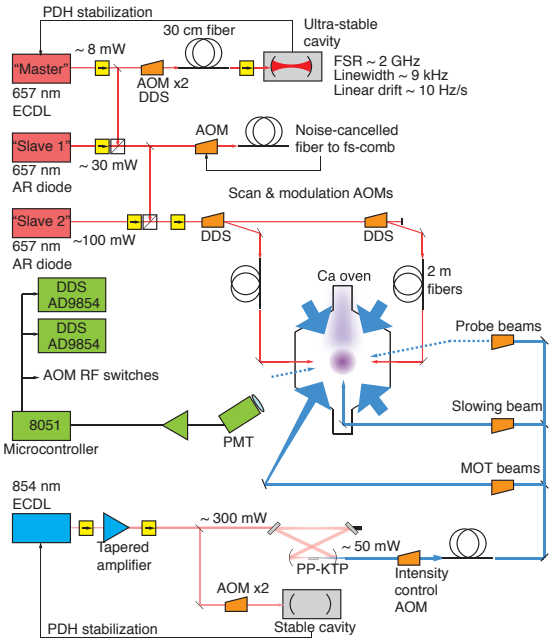


Fig. 1. The free Ca clock is among the simplest optical frequency standards. Slowing, cooling, trapping, and probing radiation at 423 nm is derived from a frequency-doubled external cavity diode laser. A diode laser at 657 nm, AR-coated and placed in a Littman external cavity, is stabilized to an ultra-high finesse optical cavity. Injection-locked 'slave' laser systems boost the 657 nm power to enable very short Ramsey pulses and homogenous broadening of the clock transition up to the residual Doppler linewidth of $\sim$ 2 MHz. A microcontroller triggers the clock sequence every 3.3 ms, acquires fluorescence data, and servos the clock laser on the atomic resonance.

A. A simplified apparatus

Figure 1 shows the current version of the apparatus, which has been optimized for stability and fast cycle rate, rather than accuracy. For instance, the MOT gradient magnetic field

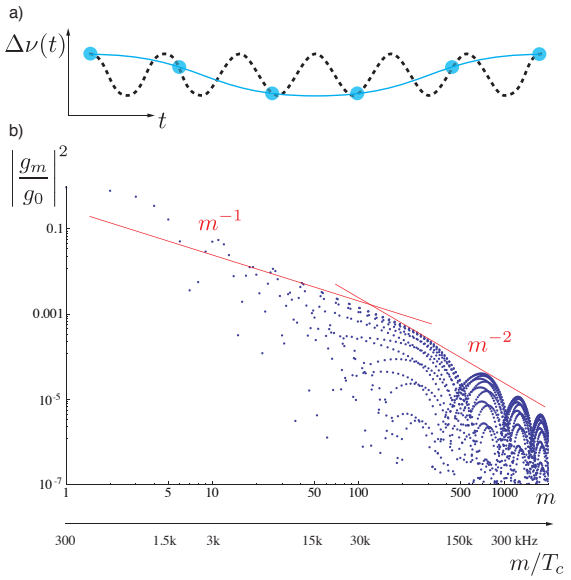


Fig. 2. (a) A cartoon illustrates the optical Dick effect as an aliasing, or downsampling of noise. The dashed line represents a particular Fourier component of laser frequency noise close to the clock measurement repetition rate; the width of each blue circle represents the atomic interrogation period of each clock cycle. Since these interrogation windows are short, it takes many clock cycles to properly average over the noise component. The aliased noise (blue line) is not genuine; if written by a servo onto the local oscillator it can actually increase the short term instability of a clock. (b) Given the Ramsey-Bordé pulse sequence parameters, we calculate sensitivity coefficients g_m/g_0 in Eq. 1 which describe how a local oscillator frequency noise spectrum is aliased into white noise.

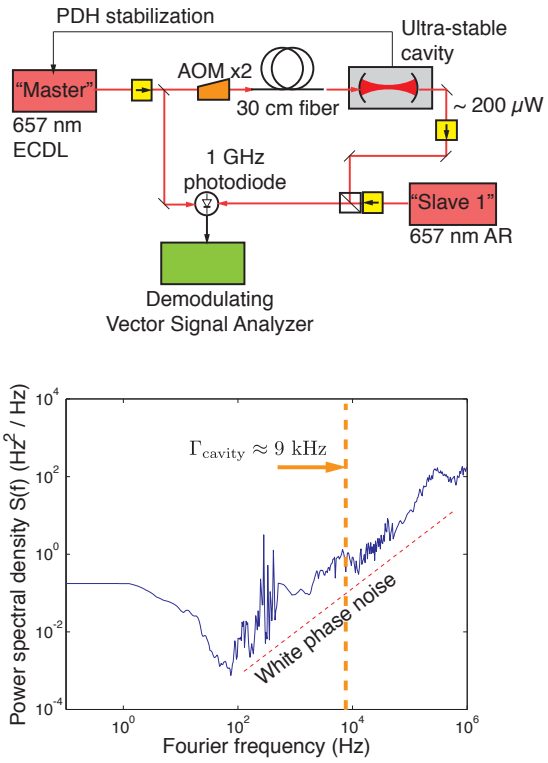


Fig. 3. Fast laser frequency noise is suppressed by filtering through a high-finesse optical cavity; the cavity acts to average out fluctuations faster than the photon ring-down timescale of $\tau = 1/\Gamma_{\text{cavity}} \sim 0.1$ ms. To determine the transfer function of the cavity, we heterodyne laser beams before and after transmission through a high finesse optical cavity used for laser stabilization. The noise added by the injection locking to a second laser is independently evaluated.

remains on permanently, eliminating vibrations that result from rapid toggling [10]. We forgo a quenched second-stage of atom cooling [10] in order to minimize the sequence duration and maximize the atomic interrogation duty cycle. The effusive atomic beam source is located about 15 cm away from the interaction region; no Zeeman slower or chirped slowing beam is employed. Key acousto-optic modulators between the clock laser and the atoms, and between the laser and the stabilization cavity, are driven with maser referenced DDS synthesizers (labeled DDS in Figure 1). The only fiber optics path that is actively noise-cancelled is the long run to the fs-comb.

II. Local Oscillator Noise

A. The optical Dick effect

Since the clock's repetition rate is 300 Hz, a digital servo locking the laser to the atomic resonance should strongly suppress local oscillator noise at 1 s averaging times. However, since the fraction of the clock sequence spent interrogating the atomic resonance—the duty cycle—is small, high frequency fluctuations are not averaged to zero but instead are down-sampled. This process, known as the optical Dick effect, results in extra white noise with a fractional instability:

$$\sigma_d^2(\tau) = \frac{1}{\tau}\sum_{m=1}^{\infty} S_y(m/T_c)\left|\frac{g_m}{g_0}\right|^2. \qquad (1)$$

Local oscillator noise S_y at harmonics m of the cycle rate $1/T_c$ are aliased by the imperfect interrogation. Sensitivity coefficients,

$$g_m \equiv \frac{1}{T_c}\int_0^{T_c} g(t)e^{-2\pi i m t/T_c}\,dt, \quad (m>0), \qquad (2)$$

tend to shrink as the interrogation duty cycle approaches 100% because the sensitivity of the atomic population to oscillator phase $g(t) = 2\,\delta P/\delta\phi|_{\delta_\phi\to 0}$ approaches a constant value. The sensitivity function $g(t)$ for the Ramsey-Bordé pulse sequence is derived elsewhere [8].

Figure 2 illustrates the aliasing effect and shows calculated Dick effect coefficients for our nominal clock cycle having an interrogation duty cycle of 13%. The Ramsey $\pi/2$ pulses last $T_R = 0.75\mu$s. The duration between Ramsey pulses is $T = 215\mu$s. The dead-time between pairs of opposing Ramsey pulses is $T_\nu = 5.8\mu$s. Increasing the two Ramsey dark periods lasting T would increase resolution and the interrogation duty cycle, but would also inflict a loss of signal; the lifetime of the excited clock state is 0.4 ms. Numerical analysis of Eq. 1 with our estimated laser noise spectrum S_y shows that more than 99% of the Dick effect instability is due to harmonics $m < 100$, or laser noise frequency components $f < 30$ kHz.

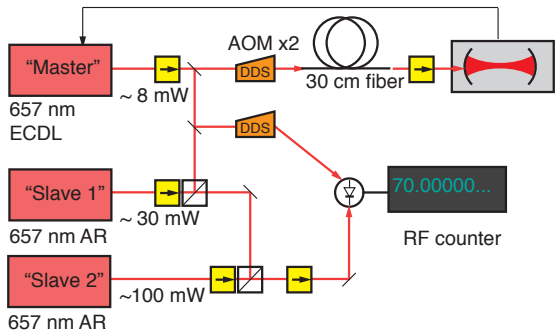


Fig. 4. Illustrated here is a heterodyne between the master external cavity diode laser and an injection-locked 'slave' laser, which is itself seeded by another 'slave' laser. An RF beatnote is created by shifting one of the beams with an acousto-optic modulator (AOM) driven by a maser-referenced synthesizer (labeled DDS). The Allan deviation of maser-referenced RF counter is related to the optical frequency fractional Allan devation relevant to Ca clock operation with $\sigma_y(\tau) = \sigma_\nu(\tau)/f_{\text{optical}}$. Direct counting of a DDS frequency yields a baseline of $\sigma_y(1\text{s}) = 6.0 \times 10^{-18}$. Table I shows the results for several such heterodyne measurements; all instabilities are much lower than the observed clock instability. The noise spectra have not been studied in detail.

TABLE I
INSTABILITY CONTRIBUTIONS DETERMINED BY HETERODYNE MEASUREMENTS.

Heterodyne test	$\sigma_y(\tau = 1\text{s})$
DDS counting baseline	6.0×10^{-18}
Before/after cavity	6.3×10^{-16}
Before/after 1.5m PM fiber	1.3×10^{-16}
Master vs. injection locked 'slave 1'	1.6×10^{-16}
Master vs. injection locked 'slave 2'	1.8×10^{-16}

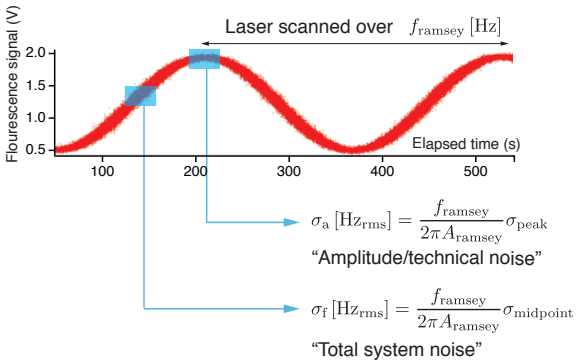


Fig. 5. In the absence of a second local oscillator or atomic reference, one can still derive useful absolute instability of a clock by using the atomic reference as a frequency discriminator. Here, the unlocked clock laser is slowly scanned over a Ramsey-Bordé fringe. Point-to-point rms deviations in the fluorescence signal are larger near the frequency-sensitive nodes. Technical noise, which is presumed to be largely laser frequency independent, is observable at the peaks and troughs of the fringe. The clock's operational parameters can be optimized to minimize each noise level. The technical noise level sets a limit on ultimate system stability even in the absence of local oscillator noise. The noise level at the fringe nodes is quantitatively consistent with the overall clock performance when compared against superior optical standards.

B. Clock laser filtering

We find the instability of the cavity-stabilized clock laser (while unlocked to the atomic resonance) to be 2.5×10^{-15} after 1 second of averaging by comparison to a superior optical reference. This instability is consistent with that derived from the laser's noise spectrum from 1 Hz–1 kHz found by demodulating the laser's beatnote with a referenced fs-comb. Faster laser noise components, perhaps due to residual phase noise not suppressed by the Pound-Drever-Hall stabilization scheme, can be downsampled by the Dick effect. To examine what fast noise might be suppressed by filtering the clock laser, we heterodyne clock laser beams before and after transmission through the high-finesse stabilization cavity. As shown in Figure 3, noise rising as f^2 is observed that contributes directly very little to the local oscillator instability at 1 second; however, through the Dick effect mechanism, the noise above $\Gamma_{\text{cavity}} \approx 9$ kHz is responsible for roughly half of the Dick effect instability. The unfiltered and filtered laser Dick effect instability limits are

$$\sigma_{\text{d,}}(\tau) = 3.6 \times 10^{-15}(\tau/1\text{s})^{-1/2}$$
$$\sigma_{\text{d, filt.}}(\tau) = 2.4 \times 10^{-15}(\tau/1\text{s})^{-1/2}$$

We do not see as large an improvement from clock laser filtering as other groups [11] due to our shorter cycle time and larger interrogation duty cycle.

C. Other noise sources

We performed heterodyne measurements between the master ECDL beam and several derived laser beams—after injection locking, after short fiber optic runs, after acousto-optic modulation (AOM)—to identify other sources of noise that might contribute to the clock instability. An example arrangement is shown in Figure 4. In each heterodyne experiment, AOM frequencies and the RF counter were maser referenced. Table I expresses these noise contributions as fractional Allan deviations of the optical frequency at 1 second.

III. TECHNICAL AND TOTAL NOISE EVALUATION

A. Self-evaluation of instability

Without the benefit of a second, independent clock laser or atomic reference, one can still evaluate some aspects of the absolute instability of a Ca clock apparatus. With a clock laser detuning ν, frequency fluctuations $\delta\nu(t)$, and Ramsey dark period T, the spectroscopy method produces a fluorescence signal [4]:

$$\rho_e(t) \propto A\cos\left(2\pi \times 4T \times (\nu + \delta\nu(t)\right) + \delta\rho_e(t). \quad (3)$$

The laser detuning ν can be slowly scanned over this feature; laser noise $\delta\nu(t)$ is translated into fluctuations in fluorescence proportional to the local slope $\partial\rho_e/\partial\nu$. At the Ramsey fringe nodes, this slope is constant and computable using the Ramsey fringe amplitude A and separation $f_{\text{Ramsey}} = 1/(4T)$. The root-mean-square deviation of successive measurements of ρ_e can be expressed in as a frequency instability by

$$\sigma_\nu(\tau = T_c)[\text{Hz}] = \frac{1}{2\pi A(4T)}\sigma_\rho[V_{\text{rms}}]. \quad (4)$$

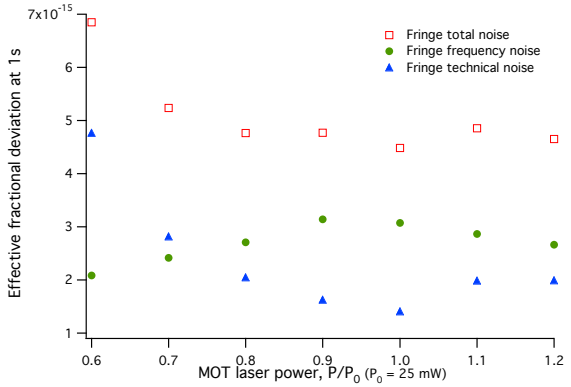


Fig. 6. The instability of the clock system is measured by using the atomic resonance as a frequency discriminator in open loop operation. Noise on the Ramsey fringe peaks and nodes can be related to effective Allan deviations due to technical and frequency-dependent noise, respectively. Here we see how these noises vary as a function of MOT trapping power. An optimum operation condition, the minimum of the total noise, is found without using a second clock or fs-comb.

At the peaks and troughs of the Ramsey fringe we observe technical noise $\delta\rho_e(t)$ which is largely independent of the clock laser frequency but nonetheless contributes noise when the servo is locked to a nodal point. At these lock-points, the technical noise will be equivalent to uncorrelated frequency noise using the same conversion factor $1/(2\pi A4T)$. Several potential sources of technical noise are suppressed by normalizing the signal to atom number, and by dithering the Ramsey-Bordé fringe pattern to remove the incoherent excitation background. The remaining technical noise could include shot-to-shot atom number fluctuations imperfectly accounted for by the normalization pulse, fluctuations in 657 nm pulse power, and frequency noise on the 423 nm probe pulses. Path length, pointing, and wavefront fluctuations between Ramsey pulses in a single clock cycle also appear as excess technical noise, but are dependent on the laser frequency as well. In general, fluctuations in any systematic effect leading to inaccuracy adds instability as well, though dithering of the Ramsey fringe suppresses such effects varying slower than half the cycle rate.

Fourier transforms of recorded $\rho_e(t)$ series at the peaks, toughs, and nodes indicate these noises are largely white from 0.5—150 Hz, and thus have equivalent noise spectra of $S_\nu(f) = \sigma_\nu^2/(f_{\text{rep}})$, where f_{rep} is the clock cycle rate. Effective fractional Allan deviations can be constructed with $\sigma_y^2(\tau) = (S_\nu/2)(1/\tau)(1/f_{\text{optical}})$.

Figure 6 demonstrates the utility of this technique. The MOT laser power is varied around its nominal power of 25 mW, and the effective deviation at 1s implied by the open-loop fringe noise is plotted. Technical and frequency noise (corresponding to peak and node noise levels) are separately plotted, and their sum agrees well with the observed clock instability when the unlocked local oscillator noise is subtracted.

IV. Conclusion

Previous efforts have aggressively pursued an uncertainty budget for the Ca optical frequency standard. In Table II, we present instead an instability budget at 1s of averaging with no account of ultimate accuracy. We conclude that three different methods of evaluating the instability share good agreement: direct measurement against a superior optical standard, a quadrature sum of the calculated optical Dick effect instability and known technical noise sources, and an open loop measurement using the atoms as a frequency discriminator. The latter technique never relies on a second clock or fs-comb, but it is directly sensitive to the open-loop local oscillator noise.

Another important result is that while filtering the clock laser through our high-finesse cavity decreases the Dick effect limit somewhat, it remains comparable to the technical noise that would ultimately limit the system if the local oscillator were significantly improved. Further study must focus on identifying and reducing the principle sources of technical noise. Since this level is thought to be near that of quantum projection noise, an increase in the number of atoms might be necessary for further improvements.

Acknowledgment

The authors would like to thank colleagues T. Fortier and S. Diddams for fs-comb measurements, and A. Ludlow, N. Lemke, R. Fox, and Y. Jiang for comparison measurements with the Yb optical lattice clock and local oscillator. We thank R. Fox, A. Curtis, F. Bondu, G. Wilpers, and Y. LeCoq for their contributions to earlier versions of the Ca apparatus.

References

[1] H. Katori, "Optical lattice clocks and quantum metrology," *Nature Photonics*, vol. 5, pp. 203–210, 2011.

[2] C. Chou, D. Hume, J. Koelemeij, D. Wineland, and T. Rosenband, "Frequency comparison of two high-accuracy Al^+ optical clocks," *Physical review letters*, vol. 104, no. 7, p. 70802, 2010.

[3] C. Oates, F. Bondu, R. Fox, and L. Hollberg, "A diode-laser optical frequency standard based on laser-cooled Ca atoms: Sub-kilohertz spectroscopy by optical shelving detection," *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics*, vol. 7, no. 3, pp. 449–460, 1999.

[4] C. Bordé, C. Salomon, S. Avrillier, A. Van Lerberghe, C. Bréant, D. Bassi, and G. Scoles, "Optical ramsey fringes with traveling waves," *Physical Review A*, vol. 30, no. 4, p. 1836, 1984.

[5] G. Wilpers, C. Oates, and L. Hollberg, "Improved uncertainty budget for optical frequency measurements with microkelvin neutral atoms: Results for a high-stability ^{40}Ca optical frequency standard," *Applied Physics B: Lasers and Optics*, vol. 85, no. 1, pp. 31–44, 2006.

[6] C. Degenhardt, H. Stoehr, C. Lisdat, G. Wilpers, H. Schnatz, B. Lipphardt, T. Nazarova, P. Pottie, U. Sterr, J. Helmcke, *et al.*, "Calcium optical frequency standard with ultracold atoms: Approaching 10^{-15} relative uncertainty," *Physical Review A*, vol. 72, no. 6, p. 062111, 2005.

[7] Y. Jiang, A. Ludlow, N. Lemke, R. Fox, J. Sherman, L. Ma, and C. Oates, "Making optical atomic clocks more stable with 10^{-16}-level laser stabilization," *Nature Photonics*, vol. 5, pp. 158–161, 2011.

[8] A. Quessada, R. Kovacich, I. Courtillot, A. Clairon, G. Santarelli, and P. Lemonde, "The dick effect for an optical frequency standard," *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 5, p. S150, 2003.

[9] G. Dick, "Local oscillator induced instabilities in trapped ion frequency standards," California Institute of Technology, Jet Propulsion Lab, Pasadena, CA, Tech. Rep., 1987.

TABLE II
A SHORT-TERM INSTABILITY BUDGET FOR A FREE CALCIUM OPTICAL CLOCK.

Noise source	Symbol	Measurement technique	Fractional Allan deviation at 1 s
Local oscillator, unlocked	σ_L	Noise spectrum measured from beat with Yb-clock on fs-comb	2.5×10^{-15}
Dick effect, unfiltered	σ_{d}	Calculated	3.6×10^{-15}
Dick effect, filtered through cavity	$\sigma_{\mathrm{d,filt.}}$	Filtered noise obtained by filtered/unfiltered heterodyne	2.4×10^{-15}
Technical noise, and atom/photon shot noise	$\sigma_{\mathrm{tech.}}$	'Self-evaluated' open-loop Ramsey fringe noise at peaks/troughs	1.8×10^{-15}
Additional noise from injection locking, fibers, etc.	$\sigma_{\mathrm{add.}}$	Heterodyne measurements	0.6×10^{-15}
Quadrature sum of $\sigma_{\mathrm{d,\ filt.}}, \sigma_{\mathrm{tech.}}, \sigma_{\mathrm{add.}}$ (filtered laser case)			3.0×10^{-15}
Quadrature sum of $\sigma_{\mathrm{d}}, \sigma_{\mathrm{tech.}}, \sigma_{\mathrm{add.}}$ (unfiltered laser case)			4.0×10^{-15}
Measured instability against Yb lattice clock, filtered laser, typical			$3.5(5) \times 10^{-15}$
'Self-evaluated' open loop Ramsey fringe noise at nodes, σ_L subtracted in quadrature			3.6×10^{-15}

[10] G. Wilpers, C. Oates, S. Diddams, A. Bartels, T. Fortier, W. Oskay, J. Bergquist, S. Jefferts, T. Heavner, T. Parker, *et al.*, "Absolute frequency measurement of the neutral 40-Ca optical frequency standard at 657 nm based on microkelvin atoms," *Metrologia*, vol. 44, p. 146, 2007.

[11] T. Nazarova, C. Lisdat, F. Riehle, and U. Sterr, "Low-frequency-noise diode laser for atom interferometry," *JOSA B*, vol. 25, no. 10, pp. 1632–1638, 2008.

Evaluation of Fabry-Perot cavity length by the stabilized optical frequency comb and acetylene absorption

Radek Šmíd, Martin Čížek, Zdeněk Buchta, Josef Lazar, and Ondřej Číp
Institute of Scientific Instruments of Academy of Sciences of the Czech Republic, Brno, Czech Republic.
email: smid@isibrno.cz

***Abstract*— The ultra low expansion cavities play a crucial role in laser stabilization, and they are essential in ion clocks. We present a method to monitor the temperature distance changes in the Fabry-Perot cavity with a spacer from an ultra low expansion material with sub-nanometer resolution. The FPC was placed into a temperature-stabilized stainless steel chamber.**

I. INTRODUCTION

The ultra-low expansion materials are crucial in Fabry-Perot cavities. The laser stabilization to the cavity by Pound-Drever-Hall technique [1] use stable cavities with low expansion material such as Zerodur ceramics [2] and ULE glass [3]. They play a key role in ion clocks [4]. Thermal dilatation on the cavity of length of 20 cm is such as 4 nm per 1 K. The knowledge of the absolute length and dilatation itself is necessary. In this work we present a method of measurement of thermal dilatation of low-expansion materials by the use of a femtosecond frequency comb [5].

The method uses a Fabry-Perot cavity (FPC) itself as tool for length changes measurement. FPC spacer is made from Zerodur. Temperature of the cavity is stabilized and effects of refractive index of air are suppressed by placing into the evacuated chamber. Any fluctuation of FPC length is demonstrated in the change of free-spectral range (FSR) of the cavity while exhibits changes in the selected mode takes effect on the change of tunable laser frequency. Changes of frequency are down-converted from the optical domain by the use of optical frequency comb.

The optical frequency comb stabilized to GPS disciplined oscillator represents a stable source in optical domain and works as a stable frequency rule in optical domain [5-10]. This work is based on evaluation of changes of FPC length by the use of optical frequency comb and tunable DFB laser.

This work was partially supported by European Commission and Ministry of Education, Youth, and Sports of the Czech Republic (project No. CZ.1.05/2.1.00/01.0017). The authors also ac-knowledge the support of Ministry of Education, Youth and Sports of CR, projects No.: LC06007, 2C06012, the AS CR, projects No.: AV0 Z20650511, Ministry of Industry and Com-merce, projects No: 2A-1TP1/127, FT-TA3/133, 2A-3TP1/113 and Grant Agency of CR projects: GA102/09/1276, GA102/07/1179 and GP102/10/1813.

II. METHODOLOGY

A. Stabilized optical frequency comb

The proposed method employs the optical frequency comb is based on the femtosecond laser and FPC. The optical frequency comb produces periodically a train of femtosecond pulses. Optical frequency spectrum of comb lines around the central optical frequency (wavelength) is produced by fs laser. The frequency f_i of a comb spectral component is represented by expression:

$$f_i = f_{ceo} + i \cdot f_{rep} \tag{1}$$

where i is the number of a comb line (typically in the order of 10^6) and f_{rep}, and f_{ceo} are frequencies typically set in radio frequency (RF) domain, called repetition and offset frequency, respectively. The repetition frequency f_{rep} is indirectly proportional to the period of the train pulses and the offset frequency f_{ceo} is caused by phase-shifts between envelope of the pulses and their carrier wavelength [11]. The offset frequency has to be retrieved from comb spectrum by self-referencing method called as the f-2f interferometric technique [5]. The offset frequency f_{ceo} is as well as the repetition frequency f_{rep} phase locked to the stable oscillator or atomic clocks RF signal by a servo-loop. The result is the optical comb frequency spectrum with relative stability of all of frequency components coupled to the relative stability of the RF oscillator or atomic clocks (thus more than 10^{-14}).

B. Fabry-Perot cavity

The optical resonator FPC consists of two mirrors separated by the geometrical distance L_{cav}. This length is represented by the spacer made from the low-expansion material. It transfers only optical frequency ν fulfilling the relation:

$$f_m = m\nu_{FSR} = \frac{mc}{4nL_{cav}}. \tag{2}$$

where ν_{FSR} is FSR of FPC, *m*-th is a longitudinal mode of the FPC in confocal geometry with mirrors separated by L_{cav}, *c* is speed of the light, *n* is the refractive index of air inside the cavity.

C. Measurement of cavity length

The comb spectrum is transmitted through FPC if some of optical frequency spectral lines f_i (1) meet the resonance frequency from (2). Reasonable number of lines transmit only than FRS of the FPC is an integer multiple of f_{rep} thus: $\nu_{FSR} = x \cdot f_{rep}$ [12]. The optical signal at the output of the FPC depends on the distance between mirrors in FPC. Then the optical output of the FPC contains number of separated resonance lines [13]. Otherwise the optical output from FPC represents the similar signal to the white light continuum and the signal doesn't contain any resonance peaks. These problems can be solved by The use of a single-mode stable laser has one resonance peak per FSR of the cavity. The use of tunable laser source is advantageous for monitor of small changes of length. Optical frequency of the closest optical frequency comb line is used as a precise frequency reference to the tunable laser source. The optical frequency of the tunable laser referenced to optical comb line (1):

$$\nu_{opt} = f_{ceo} + i \cdot f_{rep} \pm f_{beat} \tag{3}$$

where f_{beat} is the beat frequency between the closest optical comb spectral line and the tunable laser locked to the FPC. The beat frequency polarity can be easily evaluated from the nature of temperature length variations with respect to the beat frequency f_{beat}. Let us equal f_m (2) to ν_{opt} (3), then:

$$\nu_{opt} = f_m = f_{ceo} + i \cdot f_{rep} \pm f_{beat} = \frac{mc}{4nL_{cav}} \tag{4}$$

Extracting the exact length of the FPC from (4) then:

$$L_{cav} = \frac{mc}{4n\,(f_{ceo} + i \cdot f_{rep} \pm f_{beat})} \tag{5}$$

Because *i* and *m* are integer constants and *c* is exact speed of the light in vacuum, the exact length L_{cav} of FPC is determined by frequencies f_{ceo}, f_{rep} and f_{beat} . Repetition and offset frequencies are controlled by RF reference by f-to-2f stabilization. The influence of refractive index of air on the result is suppressed by placing the FPC into the vacuum chamber where refractive index equals 1:

$$L_{cav} = \frac{mc}{4(f_{ceo} + i \cdot f_{rep} \pm f_{beat})}. \tag{6}$$

The FPC mirror distance is defined by f_{beat} frequency between optical frequency comb and the tunable laser.

FPC of initial length L_{cav0} at temperature t_0 is enlarged to length L_{cav1} at temperature t_1 characterized by CTE - α:

$$L_{cav1}(t_1) = [1 + \alpha \cdot (t_1 - t_0)] \cdot L_{cav0}(t_0) \tag{7}$$

Relative changes of cavity length caused by temperature (7):

$$\alpha(\Delta t) = \frac{\Delta L_{cav}}{\Delta t} = \frac{\Delta f_{beat}}{\Delta t \cdot (f_{ceo} + i \cdot f_{rep} \pm f_{beat})} \tag{8}$$

CTE of low-expansion materials reaches values down to $0.01 \cdot 10^{-6} K^{-1}$ [2,3]. Precision of the RF stabilized oscillator 10^{-12} is equivalent to temperature change in order of 0.1 mK. So temperature instabilities are easily detected.

Next, the integer number *i* (the number of the comb line components) is last unknown variable that has to be defined for evaluation of α.To determine the *i* integer of the comb line has to be measured by "the ruler" with better resolution than by ν_{opt} / f_{rep} . The number *i* characterizing the closest component to the tunable laser frequency is evaluated by tuning the laser frequency f_{cw} (Fig. 1):

$$\Delta f = f_{cw} - (f_{ceo} + i \cdot f_{rep}). \tag{9}$$

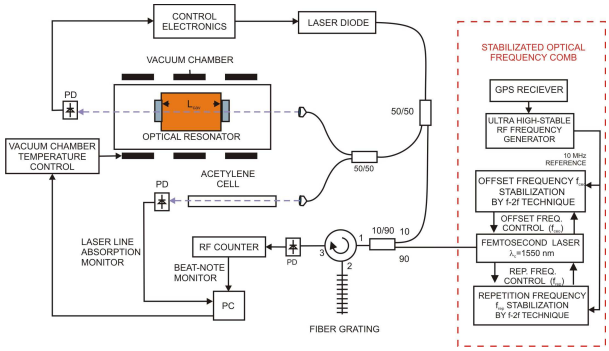


Figure 1. Scheme of the optical setup for measurement of expansion coefficient of Zerodur material represented by optical resonator using stabilized optical frequency comb and additional laser diode and acetylene cell for exact length evaluation.

The number of frequency comb component *i* can be extracted by tuning at center of absorption lines f_{ac1}, f_{ac2}, f_{ac3}, etc. (*j*=1,2,3,…) of absorption gas (for ex. Acetylene):

$$\Delta f_{acj} = f_{acj} - (f_{ceo} + k \cdot f_{rep}) \tag{10}$$

With known repetition and offset frequency of the frequency comb and measuring the beat frequency between the laser diode and the frequency comb we can easily evaluate CTE of the FPC spacer from (8).

III. Experimental Set-up

Method was verified on Zerodur ceramics tube used as spacer in confocal FPC of 400 MHz FSR. Finesse of the cavity was 339 and mirrors were deposited on the 10 mm thick fused silica substrates are optically connected to the Zerodur spacer of 187.5 mm length. The FPC was placed on four points holder in the axis of temperature stabilized stainless steel vacuum chamber evacuated down to 10^{-4} Pa. Temperature stabilization of the vacuum chamber was done by two resistance wires surrounding the chamber and controlled to 0.5 K. The temperature was measured outside the vacuum chamber as well as around the Zerodur spacer by four

temperature sensors. The temperature at the FPC surface was stable down to 0.5 mK for a period of tens of minute.

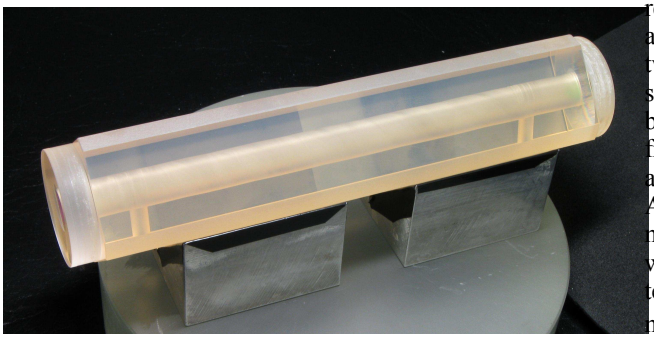


Figure 2. Photo of Fabry-Perot cavity tube with Zerodur spacer and fused silica mirror substrates.

The DFB laser diode works at central wavelength of 1542.14 nm (194.40 THz) and can be tuned by the operating temperature control for more than 2 nm (250 GHz). Another temperature control within the range from -10°C to 80°C was used to increase SNR and the FBG's central filtering wavelength was tuned within the same interval as DFB laser diode. The FBG had 0.186 nm FWHM bandwidth. Optical frequency comb was Er:doped stabilized femtosecond mode-locked laser with 100 MHz repetition frequency and central wavelength at 1550 nm and the fiber spectral output 1500-1600 nm was used. The optical frequencies of the femtosecond mode-locked laser were stabilized by f-2f interferometry [8] to 10 MHz RF reference represented by the temperature stabilized quartz oscillator disciplined by GPS receiver to Cs clock in long term.

The light from the DFB laser diode is mixed with the optical frequency comb fiber output and the beat signal is through the fiber circulator backreflected on the FBG with 0.186 nm (FWHM). Beat frequency between the closest comb components and the DFB laser diode were detected by infrared fast photodetector (PD), counted by a RF counter and recorded by the computer.

Second part of the DFB laser diode output was splitted in 50/50 coupler to the acetylene cell and to the FPC. Infrared PDs monitor absorption at the acetylene cell and the output resonance mode spectrum of the FPC. The absorption cell with the help from optical frequency comb and FPC resonance help to determinate the exact absolute value of the optical frequency of the DFB laser. The DFB tuning range covers the range from P(13) to P(11) $^{13}C_2H_2$ absorption lines [14].

Exact number of a spectral component of the frequency comb closest to the DFB frequency is determined by tuning DFB while monitoring the absorption of acetylene cell and counting repeating beat signal (10).

IV. RESULTS

Our measurement was carried out in temperature stabilized vacuum chamber at 20-30°C. The temperature range was chosen because it is close to the lowest CTE of prefabricated Zerodur material. Signals from PD's are presented in Fig. 3. Intensity of the DFB laser light transmitted through the FPC cavity is represented by the red curve. The data represents detuning of the DFB laser diode frequency from one resonance transmission peak to the closest acetylene absorption line P(11) represented by the green line to another two absorption lines P(12) and P(13) while detune over several resonance mode of FPC (red line). The beat signal between DFB laser diode frequency and the closest optical frequency comb was measured in the same time (blue line) and used for calculation of DFB frequency detuning change. After such evaluation of the k component (10) the long-term measurement at fixed FPC resonance transmission peak FPC was done. Similar measurement was done at various temperatures. Relative length changes of FPC spacer in the measurement temperature range were from $0.8\cdot10^{-7}$ to $5.8\cdot10^{-7}$ and the standard deviation for the data was $2\cdot10^{-9}$. It corresponds to absolute change of 15 to 464 nm in the case of 187.5 mm long FPC with the standard deviation of 0.4 nm. The factory CTE of the Zerodur [2] spacer was $(0.028\pm0.005)\cdot10^{-6}K^{-1}$ at 23°C. Measured CTE within the temperature range was always higher than this value in the range from $0.03\cdot10^{-6}K^{-1}$ even to $0.1\cdot10^{-6}K^{-1}$. According to [15,16] measurement of CTE of Zerodur in this temperature range could have either positive or negative value with zero crossing between 22.5-30°C. Even that the standard deviation of data may seem to be high the resolution of the method is in the order of resolution of the beat frequency counter 1 [Hz] to the optical frequency of the DFB laser - 194.40 [THz] or by the stability of the reference whichever apply first. For FPC with 187.5 mm long spacer it reaches the resolution down to 1 fm.

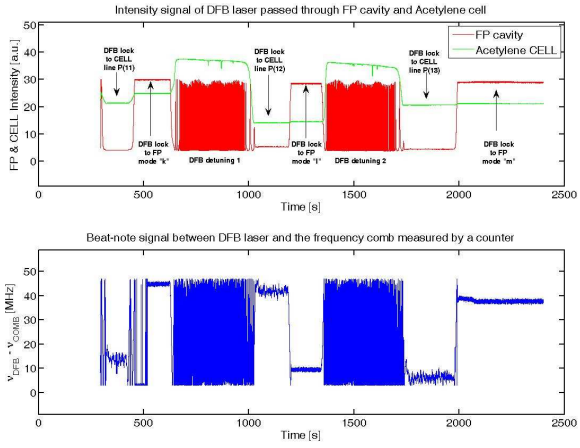


Figure 3. Measurement of DFB laser frequency detuning with from Fabry-Perot cavity mode to P(11), P(12) and P(13) acetylene absorbtion line. Fabry-Perot cavity modes (red line), absorption in acetylene cell (green line). Beat signal between optical frequency comb and DFB frequency.

V. CONCLUSION

We experimentally demonstrated the method to monitor the cavity length changes and expansion coefficient of an

ultra-low expansion material with DFB laser. Optical frequency comb was used as a frequency "ruler" reference. Acetylene cell was used as the coarse reference for finding of the optical comb component.

REFERENCES

[1] R. W. P. Drever,"Laser phase and frequency stabilization using an optical resonator". Appl Phys B, vol. 31 (2), p. 97, 1983.

[2] Technical information TIE37: Thermal expansion of ZERODUR®. http://www.schott.com/advanced_optics/english/download/schott_tie-37_thermal_expansion_of_zerodur_v2_august_2006_en.pdf

[3] ULE® Corning Code 7972, Ultra Low Expansion Glass http://www.corning.com/docs/specialtymaterials/pisheets/UleBro91106.pdf

[4] S. A. Diddams, T. Udem, J. C. Berquist, E. A. Curtis, R. E. Drullinger L. Hollberg, W. M. Itano, W. D. Lee, C. W. Oates, K. R. Vogel, and D. J. Wineland, "An Optical Clock Based on a Single Trapped $^{199}Hg^{+}$ Ion," Science, vol. 293, pp.825–828, 2001.

[5] S. T. Cundiff, and J.Ye, "Colloquium: Femtosecond optical frequency combs," Review of Modern Physics, vol. 75, pp.325-342, 2003.

[6] S. A. Diddams, L. Hollberg, L.-M. Ma and L. Robertsson, "A femtosecond-laser-based optical clockwork with instability $\leq 6.3 \times 10^{-16}$ in 1 s," Optics Letters, vol.27(1), pp.58-60, 2002.

[7] L.-S. Ma, Z. Bi, A. Bartels, L. Robertsson, M. Zucco, R. S. Windeler, G. Wilpers, C. Oates, L. Hollberg, and S. A. Diddams, "Optical Frequency Synthesis and Comparison with Uncertainty at the 10^{-19} Level," Science, vol. 303, pp.1843-1845, 2004.

[8] J. Ye, J. L. Hall, and S. A. Diddams, "Precision phase control of an ultrawide-bandwidth femtosecond laser: a network of ultrastable frequency marks across the visible spectrum," Optics Letters, vol. 25(22), pp.1675-1677, 2000.

[9] S. T. Cundiff, J. Ye, and J. L. Hall, "Optical frequency synthesis based on mode-locked lasers," Review of Scientific Instruments, vol. 72(10), pp.3749-71, 2001.

[10] A. Bauch, and H. R. Telle, "Frequency standards and frequency measurement," Reports on Progress in Physics, vol. 65, pp.789-843, 2002.

[11] H. R. Telle, G. Steinmeyer, A. E. Dunlop, J. Stenger, D. H. Sutter, and U. Keller, "Carrier-envelope offset phase control: A novel concept for absolute optical frequency measurement and ultrashort pulse generation," Applied Physics B, vol. 69, pp.327-332, 1999.

[12] R. Šmíd, O. Číp and J. Lazar, "Precise length etalon controlled by stabilized frequency comb," Measurement Science Review, vol. 8, pp.114–117, 2008.

[13] R. Šmíd, O. Číp, M. Čížek, B. Mikel, and J. Lazar, "Conversion of stability of femtosecond mode-locked laser to optical cavity length," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 57(3), pp. 636-640, 2010.

[14] K. Nakagawa, M. de Labachelerie, Y. Awaji, and M. Kourogi, "Accurate optical frequency atlas of the 1.5-mu m bands of acetylene", J. Opt. Soc. Am. B, Vol. 13 (12), pp. 2708-2714, 1996.

[15] I. Mitra, J. Alkemper, R. Müller, U. Nolte, A. Engel, H. Hack, H. Kohlmann, V. Wittmer, W. Pannhorst, M. J. Davis, L. Aschke, K. Knapp, "Optimized Glass-ceramic Substrate Materials for EUVL Applications," [Proc. SPIE 5374, Emerging Lithographic Technologies VIII, 2004]

[16] J. H. Burge, T. Peper, and S. F. Jacobs, "Thermal Expansion of Borosilicate Glass, Zerodur, Zerodur M, and Unceramized Zerodur at Low Temperatures," Applied Optics, vol. 38, pp.7161-7162,1999.

Precision calculation of blackbody radiation shifts for metrology at the 18th decimal place

M. S. Safronova
University of Delaware
Newark, DE, 19716
Email: msafrono@udel.edu

M. G. Kozlov
PNPI
Gatchina, Russia
Email: mgk@mf1309.spb.edu

Charles W. Clark
Joint Quantum Institute
NIST and University of Maryland
Email: charles.clark@nist.gov

***Abstract*— We developed a theoretical method within the framework of relativistic many-body theory to accurately treat correlation corrections in atoms with few valence electrons, and applied it to calculate polarizabilities and the blackbody radiation (BBR) shifts of atomic frequency standards. We have performed the calculations of the BBR shifts in B^+, Al^+, and In^+ that allowed to reduce the ultimate uncertainties due to this effect at room temperature to 10^{-18} level for B^+ and In^+ and to 4×10^{-19} for Al^+. These uncertainties approach recent estimates of the feasible precision of currently proposed optical atomic clocks.**

I. Introduction

Recent advances in atomic and optical physics led to unprecedented improvements in the accuracy of the optical frequency standards. In 2010, the construction of optical clock with a fractional frequency inaccuracy of 8.6×10^{-18}, based on quantum logic spectroscopy of an Al+ ion, was reported [1]. Development of more precise frequency standards will open ways to more sensitive quantum-based standards for applications such as measurements of the fundamental constants and testing of physics postulates, inertial navigation, magnetometry, and tracking of deep-space probes.

The operation of atomic clocks is generally carried out at room temperature, whereas the definition of the second refers to the clock transition in an atom at absolute zero. This implies that the clock transition frequency should be corrected in practice for the effect of finite temperature of which the leading contributor is the blackbody radiation (BBR) shift. Experimental measurements of the BBR shifts are difficult.

While the BBR shift in Al^+ frequency standards is anomalously small, its contribution becomes significant at the current level of precision. The BBR frequency shift of the clock transition can be related to the difference of the static electric-dipole polarizabilities between the final and initial clock states. Owing to severe (98%) cancellation between the static polarizabilities of the two clock states in this frequency standard, accurate calculation of the BBR shift in Al^+ is very difficult and required development of new all-order methodology. We developed a theoretical method within the framework of relativistic many-body theory to accurately treat correlation corrections in atoms with a few valence electrons [2]. This method combines the all-order approach currently used in precision calculations of properties of monovalent atoms with the configuration-interaction (CI) approach that is applicable for many-electron systems. In this work, we have extended this method to accurate calculation of the ground and excited state polarizabilities of divalent ions. The resulting polarizabilities are used to evaluate the BBR shifts at 300K in the $ns^2 - nsnp\ ^3P_0$ clock transitions in Al^+, B^+, and In^+. Frequency-dependent corrections are also evaluated and found to be negligible in all three cases.

II. Calculation of polarizabilities

The BBR frequency shift of the clock transition can be related to the difference of the static electric-dipole polarizabilities between the clock states, $\Delta\alpha_0$, by [3]

$$\delta\nu = -\frac{1}{2}(831.9\ \mathrm{V/m})^2 \left(\frac{T(K)}{300}\right)^4 \Delta\alpha_0(1+\eta), \quad (1)$$

where η is a small dynamic correction due to the frequency distribution and only the electric-dipole transition part of the contribution is considered. The magnetic-dipole contributions are suppressed by a factor of α^2 [3]. We estimated that the contribution to the Al^+ BBR shift due to the $^3P_0 - ^3P_1$ M1 transition is below 10^{-5} Hz and is negligible at the present level of accuracy. The evaluation of the dynamic correction η requires the knowledge of the electric-dipole matrix for transitions giving the dominant contributions to the polarizabilities.

In the combined CI + all-order approach used in the present work, core excitations are incorporated in the CI method by constructing an effective Hamiltonian using fully converged all-order excitation coefficients. This approach is described in detail in [2]. Its application to the polarizability calculations is described in [4]. The valence part of the polarizability is determined by solving the inhomogeneous equation of perturbation theory in the valence space [5]. The ionic core part of the polarizability is calculated separately in the random-phase approximation (RPA). The small valence-core (VC) term that corrects the ionic core polarizability for the presence of the valence electrons is also calculated in the RPA. DHF calculations are carried out as well for both of these contributions to evaluate the uncertainty associated with this term.

We note that the ionic core contribution is the same for both clock states and so it does not contribute to the BBR shift. On the other hand, the VC contribution is different for the two clock states. It is negligible for B^+. It is the largest for the

978-1-61284-110-6/11/$26.00 ©2011 IEEE

TABLE I

RELATIVE DIFFERENCES OF THE B^+ AND AL^+ CALCULATED ENERGY LEVELS AND NIST DATA (IN %). EXPERIMENTAL ENERGY LEVELS ARE LISTED FOR REFERENCE IN CM^{-1}. TWO-ELECTRON BINDING ENERGY IS LISTED IN THE FIRST ROW FOR EACH ION. ALL OTHER LEVELS ARE COUNTED FROM THE GROUND STATE.

Ion	Level	Expt.	CI	CI+MBPT	CI+All
B^+	$2s^2\ ^1S_0$	508818	0.2	0.007	-0.001
	$2p^2\ ^3P_0$	98911	-1.0	-0.114	-0.005
	$2p^2\ ^3P_1$	98920	-1.0	-0.116	-0.007
	$2p^2\ ^3P_2$	98933	-1.1	-0.124	-0.015
	$2p^2\ ^1D_2$	102363	-0.8	-0.188	-0.113
	$2p^2\ ^1S_0$	127661	-0.5	-0.264	-0.223
	$2s3s\ ^3S_1$	129774	0.2	0.010	0.014
	$2s3s\ ^1S_0$	137622	-0.2	-0.116	-0.093
	$2s3d\ ^3D_1$	150650	0.2	-0.005	-0.004
	$2s3d\ ^3D_2$	150650	0.2	-0.006	-0.004
	$2s3d\ ^3D_3$	150650	0.2	-0.006	-0.005
	$2s2p\ ^3P_0$	37336	-0.7	-0.028	0.043
	$2s2p\ ^3P_1$	37342	-0.7	-0.040	0.037
	$2s2p\ ^3P_2$	37358	-0.8	-0.054	0.020
	$2s2p\ ^1P_1$	73397	-1.3	-0.395	-0.272
	$2s3p\ ^3P_0$	143989	0.1	0.004	0.009
	$2s3p\ ^3P_2$	143990	0.1	0.003	0.008
	$2s3p\ ^3P_1$	143993	0.1	0.002	0.008
	$2s3p\ ^1P_1$	144103	0.04	-0.016	-0.004
Al^+	$3s^2\ ^1S_0$	381308	1.2	0.043	0.006
	$3p^2\ ^1D_2$	85481	2.3	0.071	-0.022
	$3s4s\ ^3S_1$	91275	1.4	0.068	0.015
	$3p^2\ ^3P_0$	94085	1.6	0.036	0.008
	$3p^2\ ^3P_1$	94147	1.6	0.032	0.004
	$3p^2\ ^3P_2$	94269	1.6	0.024	-0.004
	$3s4s\ ^1S_0$	95351	1.4	0.053	0.003
	$3s3d\ ^3D_3$	95549	1.4	-0.002	-0.026
	$3s3d\ ^3D_2$	95551	1.4	-0.002	-0.026
	$3s3d\ ^3D_1$	95551	1.4	-0.001	-0.025
	$3s3p\ ^3P_0$	37393	3.1	0.151	0.007
	$3s3p\ ^3P_1$	37454	3.1	0.140	0.008
	$3s3p\ ^3P_2$	37578	3.1	0.120	-0.017
	$3s3p\ ^1P_1$	59852	0.4	-0.175	-0.141
	$3s4p\ ^3P_0$	105428	1.4	0.068	0.020
	$3s4p\ ^3P_1$	105442	1.4	0.067	0.020
	$3s4p\ ^3P_2$	105471	1.4	0.065	0.018
	$3s4p\ ^1P_1$	106921	1.3	0.046	0.007

TABLE II

RELATIVE DIFFERENCES OF THE IN^+ CALCULATED ENERGY LEVELS AND NIST DATA (IN %). EXPERIMENTAL ENERGY LEVELS ARE LISTED FOR REFERENCE IN CM^{-1}. TWO-ELECTRON BINDING ENERGY IS LISTED IN THE FIRST ROW FOR EACH ION. ALL OTHER LEVELS ARE COUNTED FROM THE GROUND STATE.

Ion	Level	Expt.	CI	CI+MBPT	CI+All
In^+	$5s^2\ ^1S_0$	378299	5.8	-1.1	-0.25
	$5s6s\ ^3S_1$	93923	8.8	-1.5	-0.42
	$5s6s\ ^1S_0$	97030	8.1	-1.5	-0.48
	$5p^2\ ^1D_2$	97628	10.5	-2.3	-0.66
	$5p^2\ ^3P_0$	101608	7.0	-1.8	-0.42
	$5s5d\ ^3D_1$	102088	8.6	-1.4	-0.30
	$5s5d\ ^3D_2$	102174	8.6	-1.4	-0.30
	$5s5d\ ^3D_3$	102308	8.6	-1.4	-0.31
	$5p^2\ ^3P_1$	103249	7.2	-1.9	-0.44
	$5p^2\ ^3P_2$	105565	7.3	-1.9	-0.45
	$5s5p\ ^3P_0$	42276	13.2	-3.7	-1.08
	$5s5p\ ^3P_1$	43351	13.0	-3.6	-0.97
	$5s5p\ ^3P_2$	45830	13.0	-3.6	-1.06
	$5s5p\ ^1P_1$	63038	4.2	-0.4	-0.09
	$5s6p\ ^3P_0$	107662	8.1	-1.4	-0.34
	$5s6p\ ^3P_1$	107842	8.0	-1.4	-0.34
	$5s6p\ ^3P_2$	108430	8.0	-1.4	-0.35
	$5s6p\ ^1P_1$	109780	7.4	-1.3	-0.35

$^3P_0$ polarizability of In^+ to which it contributes only 0.5%. However, its contribution to the BBR shift is much larger, 1.8% and 5% in Al^+ and In^+, respectively, owing to the large degree of cancelation between $^1S_0$ and $^3P_0$ polarizabilities. We estimate the dominant uncertainty in this term as the difference of the DHF and RPA values, and assume that all other uncertainties do not exceed this dominant uncertainty. Adding these two uncertainties in quadrature, we estimate that VC term leads to the 0.6% and 2% uncertainties in the BBR shifts for Al^+ and In^+.

In order to establish the accuracy of our approach to the calculation of valence polarizability, we also perform the CI and CI+MBPT [6] calculations carried out with the same parameters (configuration space, basis set, number of partial waves, etc.). No core excitations are added in the pure divalent CI approach. Comparison of the CI, CI+MBPT, and CI+all-order values allows us to evaluate the importance of the various correlation corrections, therefore establishing the upper bound on the uncertainty of our calculations. Unless stated otherwise, we use atomic units (a.u.) for all matrix elements and polarizabilities throughout this paper.

Tables I and II presents the comparison of the experimental energies of Al^+, B^+, and In^+ with those calculated in the CI, CI+MBPT, and CI+all-order approximations. Relative differences of the B^+, Al^+, and In^+ calculated energy levels and NIST data (in %) are presented. CI, CI+MBPT, and CI+all-order data are listed in columns labeled "CI", "CI+MBPT", and "CI+All", respectively. Experimental energy levels are listed for reference in cm^{-1}. Two-electron binding energy is listed in the first row. All other levels are counted from the ground state. Significant improvement of the energy values is observed for Al^+ and In^+ with the CI+all-order method as expected owing to the more complete inclusion of the correlation corrections in comparison with the CI and CI+MBPT approaches. The CI+all-order energies are within a few cm^{-1} of the experimental values for B^+ and Al^+ for most of the levels. The accuracy of the In^+ energy levels is sufficient for the purposes of the present work, i.e. replacing our theoretical energies by the experimental values in the dominant polarizability contributions leads to only 1% change in the value of the BBR shift. The accuracy of the CI+MBPT method for B^+ is already at the level of our numerical precision for most of the transitions.

The breakdown of the contributions to the $ns^2\ ^1S_0$ and $nsnp\ ^3P_0$ polarizabilities α of B^+ ($n = 2$), Al^+ ($n = 3$), and In^+ ($n = 5$) is given in Table III. Absolute values of the corresponding reduced electric-dipole matrix elements

TABLE III

CONTRIBUTIONS TO THE $ns^2\ ^1S_0$ AND $nsnp\ ^3P_0$ POLARIZABILITIES α_0 OF B^+, Al^+, AND In^+ IN a_3^0. ABSOLUTE VALUES OF THE CORRESPONDING REDUCED ELECTRIC-DIPOLE MATRIX ELEMENTS ARE LISTED IN COLUMN LABELED "D" IN a_0e. FINAL POLARIZABILITY VALUES ARE LISTED IN ROWS LABELED "TOTAL".

Ion	State	Contr.	D	α_0
B^+	$2s^2\ ^1S_0$	$2s^2\ ^1S_0 - 2s2p\ ^1P_1$	2.118	8.918
		$2s^2\ ^1S_0 - 2s3p\ ^1P_1$	0.320	0.104
		Other		0.582
		Core		0.020
		VC		0.000
		Total		9.624
B^+	$2s2p\ ^3P_0$	$2s2p\ ^3P_0 - 2p^2\ ^3P_1$	1.354	3.216
		$2s2p\ ^3P_0 - 2s3s\ ^3S_1$	0.476	0.754
		$2s2p\ ^3P_0 - 2s3d\ ^3D_1$	1.175	1.517
		Other		2.267
		Core		0.020
		VC		-0.001
		Total		7.772
Al^+	$3s^2\ ^1S_0$	$3s^2\ ^1S_0 - 3s3p\ ^1P_1$	3.113	23.661
		$3s^2\ ^1S_0 - 3s4p\ ^1P_1$	0.045	0.003
		Other		0.138
		Core		0.265
		VC		-0.019
		Total		24.048
Al^+	$3s3p\ ^3P_0$	$3s3p\ ^3P_0 - 3s4s\ ^3S_1$	0.900	2.197
		$3s3p\ ^3P_0 - 3p^2\ ^3P_0$	1.836	8.687
		$3s3p\ ^3P_0 - 3s3d\ ^3D_1$	2.236	12.568
		Other		0.836
		Core		0.265
		VC		-0.010
		Total		24.543
In^+	$5s^2\ ^1S_0$	$5s^2\ ^1S_0 - 5s5p\ ^1P_1$	2.977	20.554
		$5s^2\ ^1S_0 - 5s6p\ ^1P_1$	0.123	0.020
		Othe		0.261
		Core		3.220
		VC		-0.041
		Total		24.014
In^+	$5s5p\ ^3P_0$	$5s5p\ ^3P_0 - 5s6d\ ^3S_1$	1.015	2.921
		$5s5p\ ^3P_0 - 5s5d\ ^3D_1$	2.189	11.755
		$5s5p\ ^3P_0 - 5p^2\ ^3P_1$	1.664	6.649
		Other		1.645
		Core		3.220
		VC		-0.170
		Total		26.019

are listed in column labeled "D" in a_0e. The ionic core polarizability and VC term that corrects it for the presence of the valence electrons are listed in rows labeled "Core" and "VC". Final polarizability values are listed in rows labeled "Total". We subtract the values of the terms listed separately in Table III from our total valence polarizability values to obtain the remaining contributions that are listed in the rows labeled "Other". Our dominant contributions for Al^+ are in excellent agreement with CI calculations with a semi-empirical core potential (CICP) of Mitroy *et al.* [7].

The oscillator strengths f_{gn} can be obtained from the reduced matrix elements by using

$$f_{gn} = \frac{2D_{gn}^2 \Delta E_{ng}}{3(2J_g+1)}, \tag{2}$$

where $\Delta E_{ng} = E_n - E_g$ and $J_g = 0$ for the present clock states.

We investigate the uncertainty due to the inclusion of the core excitations by comparing the difference $\Delta\alpha_0$ calculated in the CI, CI+MBPT, CI+all-order approximations. These results are summarized in Table III. We find that the entire contribution of core excitations to the BBR shift, estimated as the difference of the $\Delta\alpha_0$ CI+all-order and CI values is only 3%, 5%, and 16% for B^+, Al^+, and In^+, respectively. The difference between CI+MBPT and CI+all-order values is 0.4% for B^+ and Al^+, and 1.7% for In^+. Therefore, we place an upper bound on the uncertainty of our BBR values at 10% for all three cases.

III. BBR SHIFTS

Our final results are summarized in Table IV [4], where we list the polarizability difference $\Delta\alpha_0$, BBR shift at $T = 300K$, relative BBR shift $\Delta\nu_{\rm BBR}/\nu_0$, and the uncertainty in the relative BBR shift for B^+, Al^+, and In^+. We find that dynamic corrections are very small for both states and nearly cancel each other contributions. Their contributions to BBR shift are negligible for all three ions. Our BBR shift value in Al^+ $\Delta\nu_{\rm BBR} = -0.00426(43)$ Hz is in agreement with CICP value of Mitroy *et al.* [7] and recent coupled-cluster calculation [8]. It is also consistent with experimental measurement $\Delta\nu_{\rm BBR} = -0.008(3)$ Hz from Ref. [9]. The values of η for Al^+ are in agreement with [7].

TABLE IV

BBR SHIFTS AT $T = 300K$ IN B^+, Al^+, AND In^+. $\Delta\alpha_0$ IS GIVEN IN a_3^0; THE BBR SHIFTS $\Delta\nu_{\rm BBR}$ ARE GIVEN IN HZ.

Ion	$\Delta\alpha_0$	$\Delta\nu_{\rm BBR}$ (Hz)	$\lvert\Delta\nu_{\rm BBR}/\nu_0\rvert$	Uncertainty
B^+	-1.851	0.0159(16)	1.42×10^{-17}	1×10^{-18}
Al^+	0.495	-0.00426(43)	3.8×10^{-18}	4×10^{-19}
In^+	2.01	-0.0173(17)	1.36×10^{-17}	1×10^{-18}

IV. CONCLUSION

In summary, we have performed the calculations of the BBR shifts in B^+, Al^+, and In^+ that allowed to reduce the ultimate uncertainties due to this effect at room temperature to 10^{-18} level for B^+ and In^+ and to 4×10^{-19} for Al^+. These uncertainties approach recent estimates of the feasible precision of currently proposed optical atomic clocks [10]. This work demonstrated the success of the CI+all-order method in accurate determination of the polarizabilities and BBR shifts.

ACKNOWLEDGMENT

This research was performed under the sponsorship of the US Department of Commerce, National Institute of Standards

and Technology, and was supported by the National Science Foundation under Physics Frontiers Center Grant PHY-0822671. The work of MGK was supported in part by RFBR grant #11-02-00943.

REFERENCES

[1] C. W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband, "Frequency Comparison of Two High-Accuracy Al^+ Optical Clocks", *Phys. Rev. Lett.*, vol. 104, p. 070802, 2010.

[2] M. S. Safronova, M. G. Kozlov, W. R. Johnson, and D. Jiang, "Development of the configuration-interaction + all-order method for atomic calculations," *Phys. Rev. A*, vol. 80, p. 012516, 2009.

[3] S. G. Porsev and A. Derevianko, "Multipolar theory of blackbody radiation shift of atomic energy levels and its implications for optical lattice clocks," *Phys. Rev. A*, vol. 74, p. 020502(R), 2006.

[4] M. S. Safronova, M. G. Kozlov, and C. W. Clark, "Precision calculation of blackbody radiation shifts for metrology at the 18^{th} decimal place," submitted to *Phys. Rev. Lett.*, arXiv:1105.3233, 2011.

[5] M. G. Kozlov and S. G. Porsev, "Polarizabilities and hyperfine structure constants of the low-lying levels of barium", *Eur. Phys. J. D*, vol. 5, p. 59, 1999.

[6] V. A. Dzuba, V. V. Flambaum, and M. G. Kozlov, "Combination of the many-body perturbation theory with the configuration-interaction method," *Phys. Rev. A*, vol. 54, p. 3948, 1996.

[7] J. Mitroy, J. Y. Zhang, M. W. J. Bromley, and K. G. Rollin, "Blackbody radiation shift of the Al^+ clock transition", *Eur. Phys. J. D*, v. 53, p. 15, 2009.

[8] Mihály Kállay, H. S. Nataraj, B. K. Sahoo, B. P. Das, and Lucas Visscher, "Relativistic general-order coupled-cluster method for high-precision calculations: Application to the Al^+ atomic clock", *Phys. Rev. A*, 83, 030503(R), 2011.

[9] T. Rosenband, W. M. Itano, P. O. Schmidt, D. B. Hume, J. C. J. Koelemeij, J. C. Bergquist, and D. J. Wineland, "Blackbody radiation shift of the $^2 7Al^+$ $^1S_0 - ^3 P_0$ transition", *Proc. EFTF Conf.* p.289, 2006.

[10] T. Rosenband *et al.*, "Frequency ratio of Al^+ and Hg^+ single-ion optical clocks; metrology at the 17th decimal place", *Science*, vol. 319, p. 1808, 2008.

Ferroelectric PZT RF MEMS Resonators

Jeffrey S. Pulskamp, Sarah S. Bedair, Ronald G. Polcawich, Daniel Judy
US Army Research Laboratory
Adelphi, MD, USA

Sunil A. Bhave
School of Electrical and Computer Engineering
Cornell University
Ithaca, NY, USA

***Abstract*—This paper presents recent work on ferroelectric-piezoelectric lead zirconate titanate (PZT) RF MEMS resonators. Several research topics are introduced including design modifications that improve out-of-band rejection by 35dB, a demonstration of a high rejection (~117dB) parallel resonant mode with DC bias tunable rejection, a MEMS support design for the mitigation of anchor loss, and a numerical electrode shaping design technique that permits the excitation of arbitrary modes in arbitrary geometries.**

I. Introduction

Piezoelectric transduction via PZT offers a number of unique advantages for RF MEMS applications. It offers extremely low power consumption and low voltage device operation [1]. CMOS compatible voltages in RF MEMS switches have been previously demonstrated [2]. Due to the strong piezoelectric effect in PZT, actuators are capable of very large force generation and or large displacements. This provides advantages in devices like RF MEMS switches where large forces lead to low contact resistance/insertion loss and improved lifetime and large displacements lead to improved high frequency isolation. These actuation capabilities are a consequence of the large electromechanical coupling factors and piezoelectric coefficients of the material. PZT has coefficients that are generally an order of magnitude larger than those of AlN or ZnO [3]. PZT is not only a piezoelectric material but also possesses ferroelectric properties. The high electric field response in ferroelectrics permits the tuning of the dielectric, elastic, piezoelectric, and coupling factors of PZT [4]. This permits the advantageous DC-bias tuning of frequency and motional impedance in RF MEMS resonators based on ferroelectrics like PZT [5].

Thin film piezoelectrics tend to have much higher breakdown strengths than their bulk counterparts. This permits the safe application of actuation or bias fields that are well in excess of the coercive field of the PZT. Consequently, significant high electric field nonlinearity and hysteresis are observed in thin films that are rarely encountered in bulk PZT actuator applications. At these high electric fields, PZT exhibits the traditional ferroelectric hysteresis loop relating electric polarization and field. The dielectric and piezoelectric coefficients display a similar nonlinear response [6]. The combination of the piezoelectric and dielectric nonlinearity leads to tuning of the various quasi-static electromechanical coupling factors; such that high DC biasing provides significant improvements in electromechanical coupling.

Despite the long and successful history of bulk PZT resonators [7]; drawbacks to PZT, particularly at higher frequency operation, have discouraged the exploitation of this strong piezoelectric in thin films and MEMS. PZT does have high mechanical and dielectric loss and the temperature stability of the dielectric, ferroelectric, and piezoelectric properties can pose challenges. Despite these challenges, there exist several key and unique advantages for utilizing PZT in RF MEMS resonator applications. The strong piezoelectric coefficients and material coupling factors permit wide-bandwidth and low motional resistances in filters and high efficiency and voltage gain in electromechanical MEMS power transformers [8]. The high field ferroelectric response permits interesting opportunities to tune and/or trim resonators with DC bias. High levels of device integration have also been demonstrated in PZT RF MEMS, including one of the first demonstrations of monolithically integrated switches and filters [9]. Previous work demonstrated the advantages of integrating thin film PZT with high mechanical quality factor single crystal silicon, or PZT-on-Si resonator technology, to create an effective thin film composite piezoelectric material [5]. By utilizing the PZT as the transducer and the silicon as the primary resonator structure, favorable tradeoffs exist between the mechanical quality factor of the silicon and the effective coupling factor of the resonators. Despite the reduced effective coupling, the improvements in quality factor generally lead to not only higher Q but also to reduced motional resistance.

The remainder of the paper introduces some recent research in PZT-based RF MEMS resonators. Section II addresses the improvement of out-of-band rejection in two-port PZT resonators. Section III presents the demonstration of a high rejection (~117dB) and tunable parallel resonant mode. Section VI introduces a novel MEMS support design technique for the mitigation of anchor loss. Section V introduces a numerical electrode shaping design technique that permits the excitation of arbitrary modes in arbitrary geometries and Section VI concludes the paper.

II. Improving Out-of-Band Rejection

Improving out-of-band rejection in resonant filters is necessary for reducing the interference from unwanted signals. An illustration of a previously designed length extensional mode, PZT-on-silicon resonator design is shown in Fig. 1(a). A high isolation device type is also illustrated where the drive

978-1-61284-110-6/11/$26.00 ©2011 IEEE

and sense common bottom electrode is etched with the fabrication process previously developed [10]. This PZT/Pt etch feature is illustrated in Fig. 1(b). Equivalent circuit models, including the device parasitics, are shown in Fig. 2(a-b). These correspond to the illustrations in Fig. 1(a-b), respectively. The resistance, R_g, arising from the routing of the bottom platinum in the tethers to the ground plane is proportional to the tether length/width. It is important to note that this resistance has a significant effect on the stop-band performance of the resonant device. This R_g value is particularly important to consider with the high dielectric constant of PZT. At high frequencies, current can flow out of node A through the sense capacitance, due to the low impedance of C_{sense}, manifesting itself as "feed-through" current in the frequency response.

Three fabricated device types are under consideration. The first (type-I) is a high feed-through device with R_g = 1.7 Ω, the second (type-II) is a high feed-through device with R_g = 15 Ω, and the third (type-III) is a low feed-through, high isolation device with R_g = 1.7Ω. The length, L_{elec}, width, w_{elec}, of both devices' top electrodes (Fig. 1(b)), and gap in the PZT and platinum, w_{gap}, are 196 µm, 14 µm, and 4 µm, respectively. The device length, L_{tot}, and width, w_{tot}, are 200 µm and 40 µm, respectively. The platinum thickness is 100 nm, the device silicon is 10 µm, and the PZT is 0.5 µm. The measured S-parameter (S_{21}) frequency responses for the device types in the 1-200MHz range are shown in Fig. 3. The devices have been de-poled to illustrate the purely electrical feed-through. An improvement of 35dB (type-I versus type-III) in the stop-band is measured at 200MHz. A 16 dB degradation in the stop-band rejection, consistent with theory, results for the device with larger R_g (15 Ω versus 1.7 Ω). This confirms the importance of this parasitic as a relevant device parameter. Further minimizing this parasitic may be achieved through modified tether design (Section IV) and fabrication steps to improve both resonator insertion loss as well as high frequency stop-band rejection.

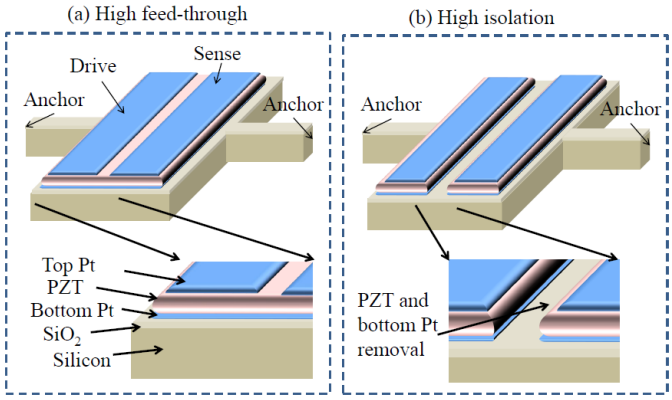


Figure 1: (a) 3D illustration of high feed-through, low isolation resonator device fabricated through a PZT-on-silicon process [10]. (b) Illustration of the high isolation device where PZT and the bottom platinum are etched in the region between the drive and sense electrode.

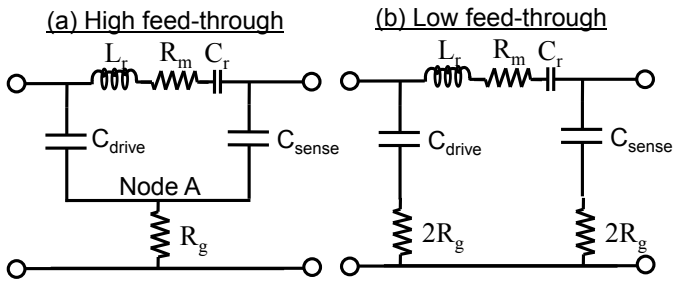


Figure 2: (a) Equivalent circuit of the two port resonant device of Fig.1(a) including the routing parasitics and electrically equivalent representation of R_m, C_r and L_r of the mechanical response (b) equivalent circuit of device with etch feature as illustrated in Fig. 1(b).

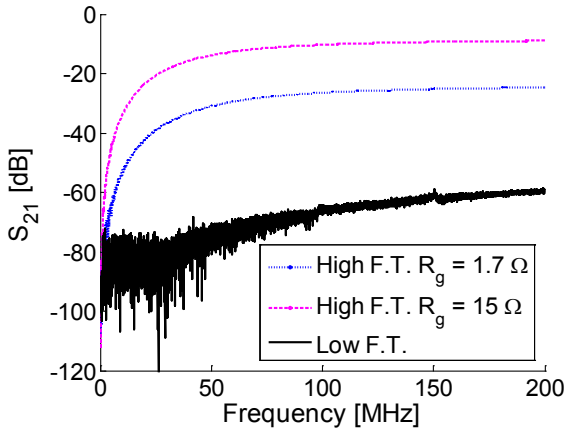


Figure 3: Frequency response of device with (low F.T.) and without (high F.T.) the PZT/Pt cut on the resonator.

III. High Q Parallel Resonances

In addition to the utility of RF MEMS resonators for band-pass filter applications; they may also be used as high rejection, parallel resonators, in band-stop or notch filter applications. The following illustrates that the existence of high rejection, parallel resonances are a consequence of the cancellation of motional and feed-through currents. For the current analysis, the following frequency dependent impedances will be used: $Z_r(s)$, the motional impedance of the resonator, $Z_d(s)$, the impedance of the drive capacitance, $Z_s(s)$, the impedance of the sense capacitance, and $Z_L(s)$ the load impedance of the measurement instrument (50 Ω). The drive and sense capacitances are illustrated in Fig. 1(a). The complex variable, s, is represented by $j\omega$, where ω is the frequency in [rad / sec]. Assuming $Z_r(s) >> Z_L(s)$, the ratio of the motional current, $i_r(s)$, to feed-through current, $i_f(s)$, current is

$$\frac{i_r(s)}{i_f(s)} \approx \frac{1}{Z_r(s)}\left(Z_d(s)\left(1+\frac{Z_s(s)+Z_L(s)}{R_g}\right)+1\right) \quad (1)$$

The conditions for parallel resonance are achieved with +/- 180° phase and unity magnitude for the above expression which leads to current cancellation of the motional with the feed-through current.

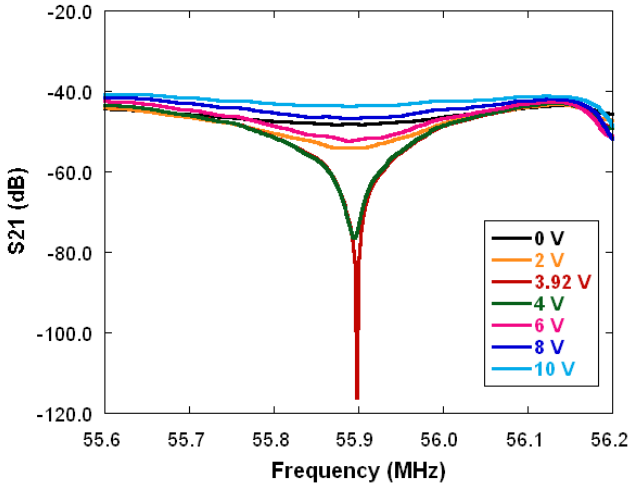


Figure 4: Frequency response of device with a high feed-through design under various DC bias conditions. High rejection parallel resonance is observed when the resonator motional exactly cancels the feed-through current.

For a device similar to that described above in Fig. 1(a), a high rejection parallel resonance mode was identified. The measured frequency response is shown in Fig. 4. A DC bias is superimposed on the AC signal at both the input and output ports using bias-tees. A high isolation of 117 dB is measured with a 3.92V DC bias at 55.9 MHz. The pass-band insertion loss is measured as 45dB at 55.6 MHz. Through DC bias tuning of the ferroelectric PZT properties, a quasi-infinite isolation can be achieved through proper design of the resonator motional impedance and the finite resistance to ground, R_g, which was introduced in Section II.

IV. Anchor Loss Reduction

Anchor loss has long been known to be a potential limiting mechanical loss mechanism. Integration of PZT with single crystal silicon enables dramatically higher quality factors than are achievable with PZT alone and anchor loss may eventually limit resonator quality factors. The common solution in MEMS resonators and older wire supported quartz and ceramic resonators is quarter-wave acoustic matched support design [11]. Quarter-wave support design views the support structure as a mechanical impedance transformer [12]. The support should accommodate some velocity at the attachment location between the tether and the resonator (drive point) and provide a large mechanical impedance at the attachment location with the substrate (anchor point). However, in an actual continuous elastic structure, there is always the potential for stress concentration at the anchor point. Therefore anchor loss can be further reduced by minimizing the amount of strain energy present near the anchor point that could potentially radiate into the substrate.

A conservative estimate of the anchor loss quality factor (anchor Q) may be obtained by assuming all of the strain energy present at the anchor point will be completely radiated during a given vibration cycle. This implies that no acoustic energy is reflected back into the resonator due to scattering of

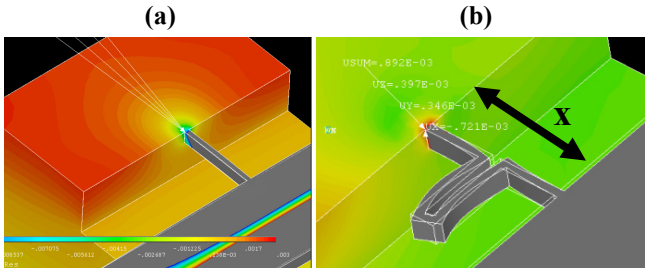


Figure 5: (a) Displacement modal analysis of a 200μm x 40μm free-free beam resonator operating at the fundamental length extensional mode (23.1 MHz). The resonator is supported by a straight 20μm long x 2μm wide tether. (b) Modal analysis of the same resonator as in (a) but with an alternative tether design. The beam resonator is visible in the lower right of both figures. The plotting contours highlight the response in the substrate.

acoustic waves radiated during prior vibration cycles. This assumption will result in predicting a lower bound to the quality factor. The anchor Q is merely the ratio of 2π times the stored mechanical energy in the resonator to the acoustic energy radiated into the substrate per period. As strain energy is proportional to the square of the strain, the conservative estimate for the anchor Q is proportional to the inverse of the square of the strain at the anchor point.

In contrast to modeling anchor loss via harmonic analysis [13], modal analysis is far less computationally demanding for the same physical model but only provides relative kinematic and dynamic data. However dissimilar support designs can be compared if the modal data of the two designs are scaled with a consistent normalization scheme; for example by normalizing both modeshapes to unity. The ratio of the squares of the modal strains of the two designs will be equal to the ratio of their anchor Qs (lower bound).

Fig. 5(a) illustrates a typical MEMS straight tether support, and Fig. 5(b) illustrates a potentially more compliant support. The results are from an ANSYS 12.0 FEA modal analysis of the same resonator at its fundamental length extensional mode of 23.1 MHz. As can be seen in Fig.s 5(a-b), the different tethers are experiencing different "local" vibrations, at the identical resonator mode, which provide different anchor loss conditions. The tether in Fig. 5(b) actually has an average absolute value of the relative Von Mises stress at the surface between the tether and the substrate that is 89% lower than that of the straight tether design in Fig. 5(a).

This difference in these tether responses can be understood by considering the frequency response and modal data of the isolated tether. The tether structure behaves essentially as its own resonator with dynamics, partly determined by the boundary conditions imposed by the substrate and resonator, and excitation imposed by the motion of the resonator at the drive point. At a given frequency, the motion of the tether will be some linear combination of the "local" tether modes of vibration. Tether modes with low mechanical impedances at the drive point, subjected to large relevant forces by the motion of the resonator, will be the dominant tether modes governing the dynamics of the tethers, at any given frequency. The goal of any design approach is to determine

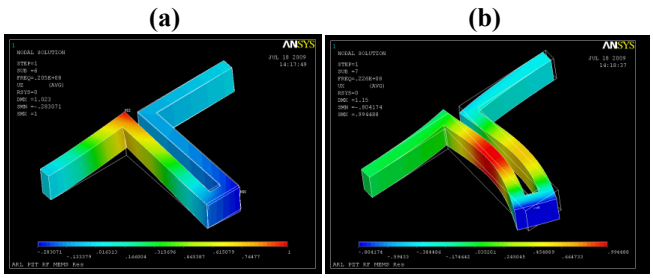


Figure 6: Modal analysis of isolated tether under boundary conditions consistent with substrate and resonator attachments impedances. The plotted response is the vector-sum of displacements (unity normalized).(a) 21.2 MHz tether mode (b) 22.6 MHz tether mode.

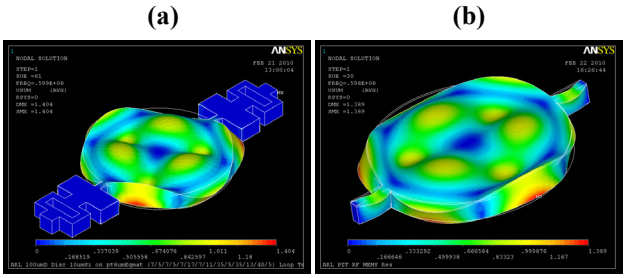


Figure 7: Modal analysis of 100µm diameter PZT-on-Si disc resonator operating at 60 MHz high order disc flexure mode. The plotted response is the vector-sum of unity normalized displacements (blue identifies near zero relative displacements). (a) New transformer design (b) 20µm long x 5µm wide straight tether design.

what tether modes will dominant the dynamics of the tethers at the desired resonator operating frequency. Generally, the supports can be designed by considering the isolated tether clamped at the anchor point and then by defining mechanical boundary conditions at the drive point that are compatible with the relevant mechanical impedances presented by the resonator.

For example, the drive point of the length extensional mode free-free beam resonator of Fig.s 5(a-b), present x-direction finite displacements to the tethers. Consequently, the relevant mechanical impedance is that of the ratio of the x-direction translational velocity to axial forces in the same direction. The isolated tether was simulated with a clamped boundary condition at the anchor point and the drive point was constrained to only permit x-direction displacements.

Fig. 6(a) and Fig. 6(b) illustrate two modeshapes of the tether from Fig. 5(b) that have resonant frequencies within a few MHz of the main free-free resonator frequency (23.1 MHz). The mode in Fig. 6(a) displays a much higher mechanical impedance (less motion at the drive point – right side of tethers in Fig. 6) than the mode in Fig. 6(b). Comparison of Fig. 6(b) with Fig. 5(b) reveals that the 22.6 MHz isolated tether mode is the dominant "local" tether mode present in the tether at the resonators resonant length extensional mode of 23.1 MHz.

Utilizing this approach for design involves determining a desired resonator geometry and mode, determining the relevant mechanical impedances of the tether attachment locations, and simulating a given tether design with the appropriate boundary conditions. The average values of the stresses on the surface between the substrate and the tether (anchor surface) are calculated and compared for the various tether modes. Once a low average stress mode is selected, the tether dimensions are resized so that it is the dominant local tether mode at the desired resonator frequency.

This technique was used to design the resonator support seen in Fig. 7(a). To compare this design with a typical MEMS straight tether design, a 20µm x 5µm straight tether (Fig. 7(b)) was simulated with a 100µm diameter silicon disc resonator with a similarly sized new mechanical impedance transformer design. The tether dimensions were fixed without constantly matching for each resonator mode to assess the bandwidth sensitivity of the designs. The average

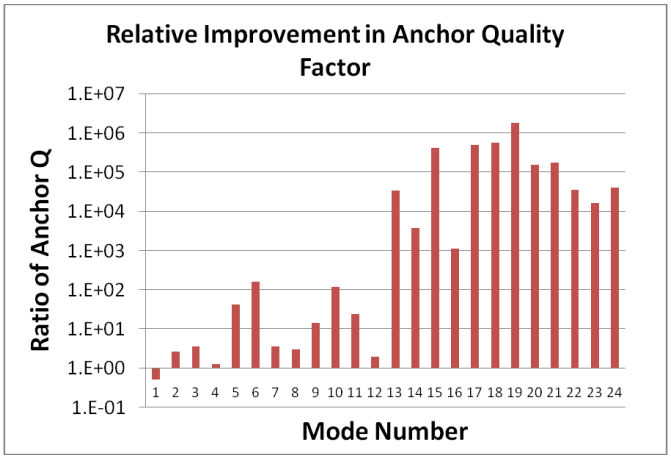


Figure 8: Plot of anchor Q ratio (mechanical transformer design Q / straight tether design Q) for the first 24 disc modes. The anchor Q ratio is the ratio of the squares of the average anchor surface stresses.

of the absolute value of the Von Mises stress on the anchor surface was calculated for each mode in each case. Fig. 8 plots the ratio of the anchor Q (lower bound) for the straight and transformer tether designs for the first 24 disc modes. In all but the first case, there are multiple order of magnitude improvements in the anchor Qs between the two designs, particularly for higher frequency modes. Fig. 7(a-b) plots the modal displacements for an example 60 MHz disc mode. The improved theoretical performance of the transformer tether design is apparent between Fig. 7(a) and 7(b) where the transformer tether exhibits virtually no displacements. Nearly all of the disc modes from about 15 MHz through ~200 MHz (upper frequency limit of simulation) display similar anchor performance as Fig. 7(a). The predicted performance difference suggests that the transformer based design dramatically reduces the anchor loss, is less sensitive to the matching requirements of quarter-wave designs (superior performance over a wide range of frequencies without resizing), and permits larger geometries that can reduce the tether parasitics described in Section II.

V. ELECTRODE SHAPING

The drive to achieve insertion losses and quality factors consistent with commercializing MEMS resonator technologies have led to a research focus on minimizing motional resistances and maximizing quality factors. In addition to achieving further improvements in these metrics, there remains a need to improve the suppression of spurious vibrational modes, increase resonant frequencies, improve power handling and linearity, and achieve wider bandwidth resonators for band select filter applications. The proper design of the resonator electrodes to selectively excite and detect desired modes, or "electrode-shaping", can provide improvements in these performance metrics. There has been an extensive body of research in electrode-shaped transducers, particularly for applications in active vibration control of structures [14].

Optimal excitation and detection electrodes may be analytically derived for simple one-dimensional modes by utilizing the orthogonality property of the vibrational modeshapes [15]. However, extending this technique to more complicated modes is difficult but extending the theory to higher dimensional modes, such as those existing in discs, poses considerable complications. In addition, to determine electrode-shapes in arbitrary cases, numerical techniques are required to treat issues such as nontrivial geometries, modes not conveniently described by analytical models, deformation coupling effects in asymmetric composites, anisotropic material properties, and the influence of non-optimal tether conditions on the modeshapes.

An alternative to the traditional approach is to determine the local suitability of placing an electrode on the resonator. The electrode shapes are then ascertained from maps of this local determination of electrode placement. For excitation, maximizing the local strain leads to maximizing the local contribution to the lumped equivalent excitation force, or modal force. Therefore, we may determine the local suitability of electrode placement by considering strain compatibility between the excited piezoelectric and the desired mode at each location on the resonator. The mode is excited locally when the piezoelectrically induced stresses are consistent with the modal stresses; for example when the piezoelectric material is assumed to be in compression, it will excite regions of the resonator for modes where that location is also simultaneously in a state of compression.

This technique does not necessarily maximally excite the desired mode or maximally suppress unwanted modes, but rather determines electrode shapes that ensure the desired mode is excited and detected. For this scenario, the excitation electrode shape for a particular mode may be determined by finding those locations where the e_{31} and e_{32} piezoelectric stress constants both excite or inhibit the mode. For PZT, these are the locations where the two modal principle stress components have the same sign. For locations where the two modal principle stress components have dissimilar signs, some degree of force cancellation occurs and these locations are neglected. The determination of the detection electrode shape consists of solving the constitutive equation for the direct piezoelectric effect locally for the differential of the induced charge.

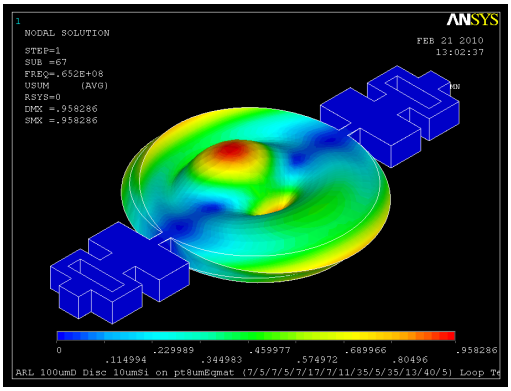


Figure 9: FEA modal analysis of an example electrode-shaped device, a high order flexure disc mode (100μm diameter). The plotted response is the vector-sum of displacements (unity normalized).

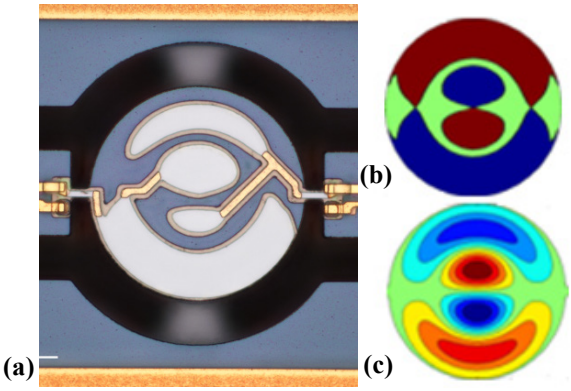


Figure 10: (a) Micrograph of the fabricated two-port PZT-on-Si disc resonator electrode-shaped for the mode illustrated in Fig. 9. The intended excitation port is on the left side of the image (b) Model result for excitation electrode shape. (c) Model result for detection electrode shape.

This technique was implemented with numerical modal analysis in ANSYS 12.0, MATLAB data processing and scripted automatic CAD generation. To validate the technique, sixteen different disc modes were designed as two-port devices and characterized. The device fabrication and layer thicknesses were previously described in Section II. The fabricated resonators were tested in an RF probe station at atmospheric pressure. Full S-parameter matrices were extracted for each resonator using a Rohde & Schwarz ZVB vector network analyzer.

Fig. 9 illustrates the modeshape of an example device, a high order flexure mode. Fig. 10(a) is a micrograph of the fabricated device; Fig. 10(b) displays the results of the electrode-shaping technique for the excitation electrode; and Fig. 10(c) displays the results of the technique for the detection electrode for the two-port devices. The red and blue regions of Fig. 10(b) correspond to the two possible excitation electrode locations. These locations correspond to either tensile or compression regions of the modal analysis simulation and either location may be chosen. The red and blue graded regions visible in Fig. 10(c) display the locations for the detection electrode. These locations correspond to either positive or negative charge regions of the modal analysis simulation and again, either may be chosen. However, in both the excitation and detection cases, electrodes

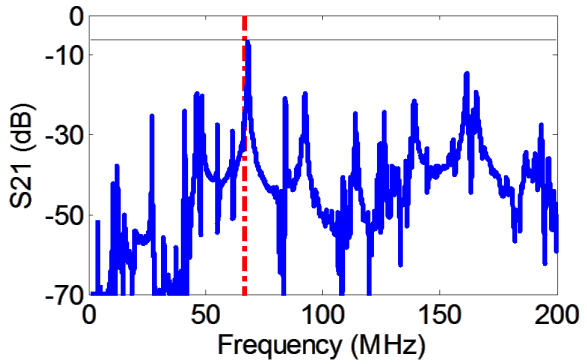


Figure 11: Experimental results for the transmission frequency response of the device shown in Fig. 10(a). The dashed red line identifies the design frequency.

that overlap the two different regions will experience some degree of force or charge cancellation, respectively. Fig. 11 displays the frequency response of the measured forward transmission coefficient S_{21} for the device shown in Fig. 10(a), displaying an unmatched (50Ω terminated) insertion loss of 6.6 dB. The dashed red line indicates the design resonant frequency for the intended mode at 68 MHz. The proximity of the lowest loss mode to the design frequency suggests the electrode shape is successfully exciting the desired mode.

To more fully characterize the electrode-shaped resonators, the measured S_{21} peak values within ±4% of *each* design frequency were determined for *each* resonator. The difference between the peak S_{21} of a given design and the average of all of the other resonators, within the same ±4% search window, provides a measure of the how well a particular electrode-shape is performing. Fig. 12 is a stem plot of this difference for each of the 16 different electrode designs for the disc resonators. The average across the different disc designs yields an average improvement in S_{21} for the shaped designs of 18.1 dB and a maximum improvement of 44.3 dB.

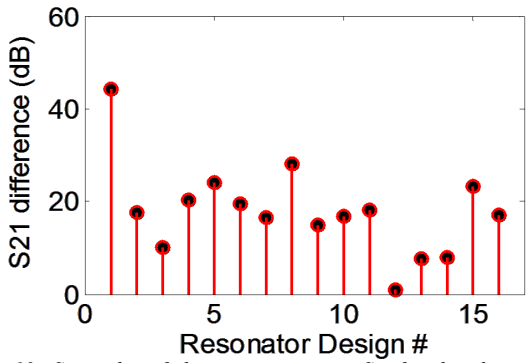


Figure 12: Stem plot of the improvement in S_{21} for the electrode-shaped designs, relative to all other electrode designs (same resonator geometry). Each value represents the improvement in insertion loss of that mode relative to the average of the other designs within ±4% of the design frequency.

VI. Conclusions

In summary, this paper introduced several research topics related to piezoelectric-ferroelectric PZT RF MEMS resonators. Significant improvements in out-of-band rejection (>35 dB), consistent with theory, were demonstrated by isolating the bottom electrodes of the excitation and detection transducers of two-port piezoelectric resonators. A high rejection (~117dB) parallel resonant mode was demonstrated with the benefit of the tunable ferroelectric properties of PZT-on-silicon resonators via high field DC bias. A MEMS support design technique, utilizing modal analysis, for the mitigation of anchor loss was presented. A theoretical comparison illustrated mechanical transformer-based tether designs derived from this technique, can provide orders of magnitude lower anchor loss than a more traditional MEMS straight support design. A numerical electrode shaping design technique was demonstrated to permit the excitation of arbitrary modes in disc geometries. The average improvement in S_{21} for the electrode-shaped designs was 18.1 dB relative to all other electrode designs for the same resonator geometry and a maximum improvement of 44.3 dB were demonstrated.

Acknowledgment

The authors wish to thank Brian Power and Joel Martin (Army Research Lab) for the fabrication.

References

[1] S. Trolier-McKinstry and P. Muralt, "Thin Film Piezoelectrics for MEMS," J. Electroceramics, vol. 12, pp. 7 – 17, 2004.

[2] R. G. Polcawich, J. S. Pulskamp, D. Judy, P. Ranade, S. Trolier-McKinstry, and M. Dubey, "Surface micromachined microelectromechanical ohmic series switch using thin-film piezoelectric actuators," IEEE Transactions Microwave Theory and Techniques, vol. 55, no. 12, pp. 2642 – 2654, 2007.

[3] R. Ghodssi, Reza and P. Lin, *MEMS Materials and Processes Handbook.* New York: Springer, 2011.

[4] A. Tagantsev, V. Sherman, K. Astafiev, J. Venkatesh, and N. Setter, "Ferroelectric Materials for Microwave Tunable Applications," J. of Electroceramics, vol. 11, pp. 5–66, 2003.

[5] H. Chandrahalim, S. A. Bhave, R. Polcawich, J. Pulskamp, D. Judy, R. Kaul, and M. Dubey, "Performance comparison of $Pb(Zr0.52Ti0.48)O3$-only and $Pb(Zr0.52Ti0.48)O3$ -on-silicon resonators," Applied Physics Letters 93(23), p. 233504, 2008.

[6] A. Kholkin, E. Akdogan, A. Safari, P.Chauvy and N. Setter, "Characterization of the effective electrostriction coefficients in ferroelectric thin films," J. Appl. Phys. vol. 89, no. 12, pp. 8066-8073, 2001.

[7] S. Fujishima, "The History of Ceramic Filters", IEEE Trans. UFFC, vol. 47, no. 1, pp. 1-7, January 2000.

[8] S. Bedair, J. Pulskamp, B. Morgan, and R. Polcawich, "Performance model of electrode tailored thin film piezoelectric transformers for high frequency switched mode power supplies," Power MEMS 2009, Silver Spring, Maryland, pp. 435-438, Dec. 1 - 4, 2009.

[9] J. Pulskamp et al., "Monolithically integrated piezoMEMS SP2T switch and contour-mode filters," MEMS 2009, pp. 900-903.

[10] H. Chandrahalim, et. al., "Influence of silicon on quality factor, motional impedance, and tuning range of PZT-transduced resonators," *2008 Solid State Sensor, Actuator and Microsystems Workshop*, Hilton Head Island, SC, pp. 360-363, 2008.

[11] R. A. Johnson, *Mechanical Filters in Electronics*. New York: Wiley, 1983.

[12] K. Wang, Y. Yu, A. Wong and C. Nguyen, "VHF free–free beam high-Q micromechanical resonators," IEEE/ASME J. Microelectromech. Syst., vol. 9, no. 3, pp. 347–360, Sep. 2000.

[13] M. Pandey, R. Reichenbach, A. Zehnder, A. Lal and H. Craighead, "Reducing anchor loss in MEMS resonators using mesa isolation," JMEMS, vol. 18, no. 4, pp. 836-844, August 2009.

[14] M. Frecker, "Recent advances in optimization of smart structures and actuators," J. Intell.Mater. Syst. Struct. 14 207–16, 2003.

[15] A. Prak, M. Elwenspoek, and J. Fluitman, "Selective mode excitation and detection of micromachined resonators," J. Microelectromech. Syst., vol. 1, no. 4, pp. 179–186, Dec. 1992.

Ultra High Frequency Air/Aluminum Nitride Fractal Phononic Crystals

Nai-Kuei Kuo, Gianluca Piazza
Department of Electrical and Systems Engineering
University of Pennsylvania
Philadelphia, USA
kuo1@seas.upenn.edu; piazza@seas.upenn.edu

***Abstract*—This paper presents the first design and experimental demonstration of a fractal phononic band gap (PBG) structure in air/aluminum nitride (AlN). The PBG exhibits two frequency band gaps for symmetric Lamb waves centered around 900 MHz and 1.10 GHz, respectively with bandwidths of 11.1% and 9.09% and attenuation as high as 40 dB. The unit cell of the fractal design consists of a center square air scatterer with four scaled-down side squares repeating at its corners. The presence of the side squares effectively shortens the scattering distance between each unit cell and permits the opening of band gaps for higher vibration modes, *i.e.* higher frequencies, for a given minimum feature size. AlN Lamb wave transducers were integrated with the PBG in order to launch acoustic waves into the structure and characterize its performance. This demonstration not only offers the evidence that it is possible to synthesize GHz PBG structures with relatively large feature sizes, but also shows that these devices can be integrated with more established radio frequency microelectromechanical systems (RF MEMS) and ultimately lead to novel all-acoustic signal processors.**

I. INTRODUCTION

Phononic crystals (PnCs) have recently become a very interesting research topic because of the possibility to synthesize novel acoustic functionalities (*e.g.* acoustic isolation and acoustic multiplexing) in a small form factor. Additionally, the use of microfabrication techniques for the making of PnCs has enabled the substantial miniaturization and mass manufacturability of these structures [1,2]. Nonetheless, the goal of utilizing these devices for the synthesis of compact radio frequency (RF) signal processors eventually capable of on-chip acoustic multiplexing and channel selection requires noteworthy advancement of the state-of-the-art. Some PBG-based acoustic devices such as acoustic cavities [3, 4, 5] waveguiding, and focusing [6] have been demonstrated, but very few have been extended to the ultra high frequency range (UHF). Operation in this region of the frequency spectrum is required in order to devise PBG structures applicable to RF systems. Although UHF PBG structures have been previously demonstrated [5], these solutions require extremely small features (< 500 nm), which challenge the reproducible synthesis of high frequency PBGs.

The fractal unit cell design is introduced in this work in an AlN/air structure to address this challenge. A fractal structure is a geometry that repeats its original shape with a scaling factor in a specific pattern. Because of the particular arrangement of the repeating elements, it is possible to reduce the scattering length between adjacent unit cells and effectively attain operation at higher frequency for a larger feature size than a conventional circular PBG topology. In this paper, the operation of an AlN/air fractal PBG around 1 GHz is demonstrated experimentally with a minimum lithographically defined feature size of 700 nm.

II. DESIGN OF THE FRACTAL PBG STRUCTURE

A. *Principle of Operation of the Fractal PBG*

Fig. 1 shows the example of a conventional "T-square" fractal structure and its pattern formation (up to 3 repetitions).

Figure 1. The patterns of a conventional T-square fractal with side square extension up to 3 times.

In this work, the T-square fractal design is employed for the implementation of the PBG structure. However, in order to attain a sufficiently wide frequency band gap, the area of each repeating square is not kept constant, but the center square is made smaller than the conventional T-square design (area scaling of 0.16 instead of 0.25) and the corner squares are scaled to each occupy 0.09 of the unit cell area (the largest square). Only the first pattern repetition is implemented experimentally, since further extension would require the definition of very small features. The PBG structure array is defined in aluminum nitride (AlN), so that direct in-plane

This work was funded by the DARPA CSSA program.

integration with piezoelectric Lamb wave transducers is possible.

In the conventional PBG structure design, the frequency of operation of the band gap is determined by the acoustic wavelength associated with the phonon scattering within the periodic array and is proportional to the unit cell lattice constant, a (Fig. 2). Differently, the presence of the side squares in the fractal PBG structure essentially forms sub-scatterers in each unit cell, which decrease the acoustic wavelength of phonon scattering (Fig. 2).

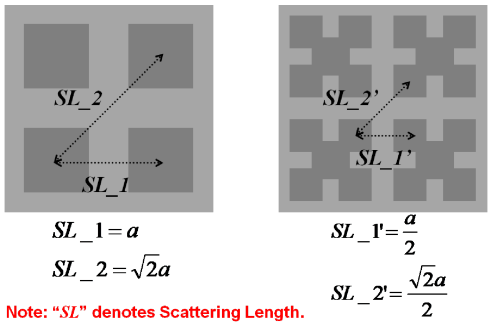


Figure 2. Comparison between the scattering length of a conventional square lattice PBG structure and that of a fractal PBG structure.

Thus, phonon scattering occurs for higher vibrational modes and results in higher frequency band gaps. Because of the extended side squares, the design space of the fractal PBG structure is more complex than then one of a simple circle and it is set by the following parameters: the lattice constant, a, the center square length, c, the side square length, s, and the film thickness, d (Fig. 3).

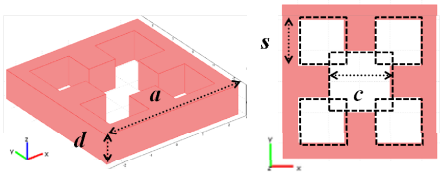


Figure 3. Unit cell of a fractal PBG structure. The main physical parameters that characterize the unit cell are: the lattice constant, a, the center square length, c, the side square length, s, and the thickness, d.

In order to understand the dependence of the fractal PBG on these various parameters, COMSOL® finite element method (FEM) simulation software is utilized to derive the operating frequency and bandwidth of the PBG [7]. The FEM solver is used to derive the eigenfrequencies of the PBG along the first symmetric Brillouin zone of the desired unit cell design. By applying periodic boundary conditions to the unit cell design, it is possible to construct a dispersion relation between the vibration modes and the reciprocal space, which further effectively reveals where frequency band gaps exist (no vibration modes present). An example of such analysis is given in Fig. 4, where the frequency dispersion plot of a conventional square lattice and a fractal PBG structure in AlN with the same periodicity (5 µm) are compared. In Fig. 4, the conventional square lattice with a circular scatterer PBG structure opens a frequency band gap centered about 3,700 m/s and bounded by the 6th and 7th vibration modes. In contrast, the fractal PBG structure operates between the 16th and 17th modes and exhibits a complete frequency band gap at approximately 5,300 m/s, which is almost 1.4 times higher than what extracted for the square lattice. This comparison confirms that the fractal design has the ability to operate at higher frequencies than a conventional circular design for a given lattice size.

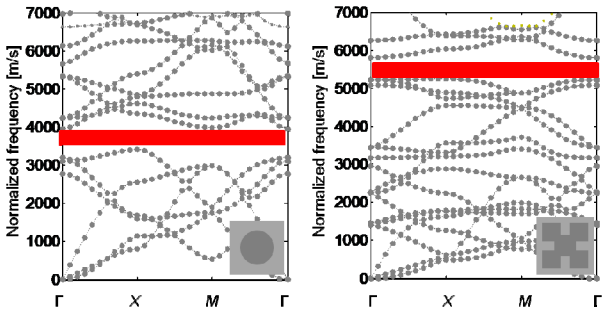


Figure 4. Comparison between the dispersion relationship of a square lattice PBG and a fractal PBG having the same lattice size. The red highlighted area is the complete frequency band gap.

In addition, in the minimum feature size associated with this PBG operating at higher frequency is much larger than the one required by the circular geometry (800 vs. 500 nm). This preliminary analysis clearly shows that the fractal design is beneficial in alleviating the manufacturing challenges generally associated with high frequency PBGs.

B. Frequency Band Gap Design and Dependence on Geometrical Parameters

For operation at 1 GHz, the lattice constant is set to be 5 µm (λ ~ 10 um in the AlN/air compound). The center square length, c, also plays an important role in the definition of the center frequency (Fig. 5). When varying the c-to-a ratio from 0.3 to 0.5, a general increase in the center frequency of the band gap is observed. Many frequency gaps are present (Fig. 5), but several are available around 1 GHz (5,000 m/s in terms of the normalized frequency) for a c/a of 0.4. For this reason, the center square length is set to be 2 µm.

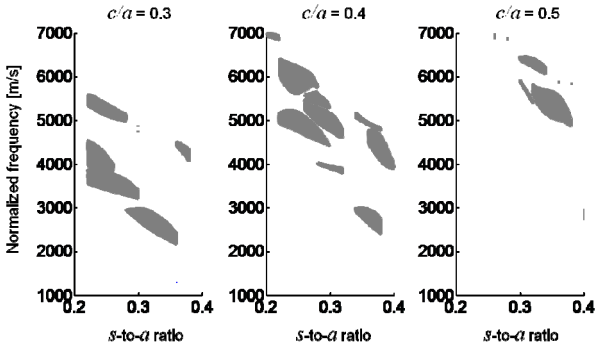


Figure 5. Map of frequency band gaps with respect to the s-to-a ratio for various center square length (c/a varying from 0.3 to 0.5). d/a is fixed at 0.2.

Moreover, to attain sufficient frequency band gaps without extreme miniaturization, the side square length, s, is defined to be 1.5 μm (s/a = 0.3), so that the minimum feature on the mask is kept at approximately 700 nm (the diagonal distance between the two neighboring cantilevers where center and side squares overlap), which is also the minimum feature that our lithography equipment allows us to define.

Furthermore, the frequency band gap sensitivity with respect to changes in the minimum feature size was also investigated. Fig. 6 is a clear evidence of the insensitivity of the band gap width to small variations in the s-to-a ratio. When the side square length decreases by 5% of the original value of 1.5 μm, the frequency band gap width shifts from 8.64% to 8.4%. On the other hand, when the side square length increases by 5%, the frequency band gap changes from 8.64% to 8.12%. Therefore, these variations in the minimum feature size are tolerable.

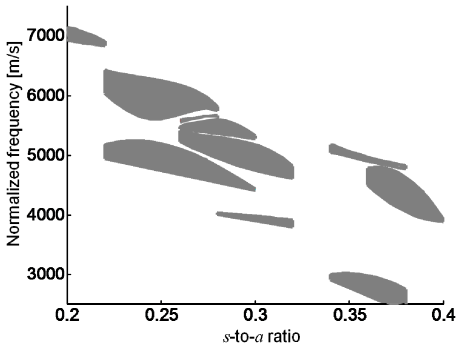


Figure 6. Map of frequency band gaps with respect to the variation in the s-to-a ratio from 0.2 to 0.4.

Finally, the film thickness is primarily fixed by the need of integration of the PBG with the AlN Lamb wave transducers. In order to attain an efficient electromechanical coupling (k_t^2) at 1 GHz, the AlN film thickness is set at 1 μm [8].

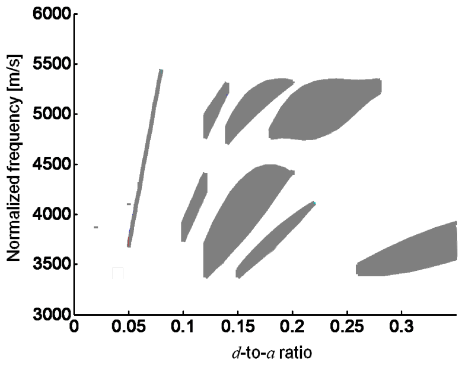


Figure 7. Map of frequency band gaps with respect to variations in the d-to-a ratio from 0 to 0.35.

Fig. 7 shows the behavior of the frequency band gaps with respect to the change in the film thickness while all other parameters are kept constant. The analysis shows that for the d/a ratio < 0.2, the frequency band gaps defined by the lower modes tend to widen. Nonetheless, small variations of d/a from 0.2 result in acceptable (almost negligible) changes in the PBG bandwidth. For example, if the thickness increases by 1% of the original value of 1 μm, the bandwidth of the stop band changes from 8.64% to 8.71%; on the other hand, if the thickness decreases by 1%, the bandwidth changes from 8.64% to 7.8%. These data clearly indicates that small variations in thickness will not have a dramatic impact on the desired frequency band gap.

Therefore, the desired fractal PBG structure of this work is defined by the following parameters: a = 5 μm, c = 2μm, s = 1.5 μm, and d = 1 μm. This PBG exhibits two complete frequency band gaps: one from 960 MHz to 1.03 GHz, another from 1.05 GHz to 1.07 GHz (Fig. 8).

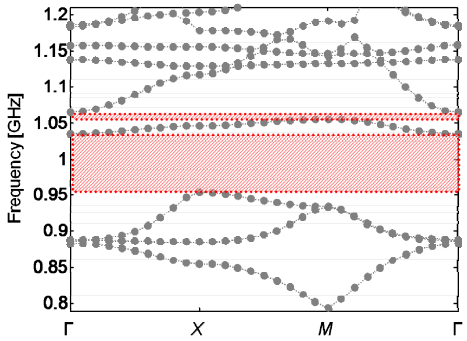


Figure 8. Dispersion relationship for the designed fractal PBG structure with physical key parameters: a = 5 μm, c = 2μm, s = 1.5 μm, and d = 1 μm. The red shaded areas are the regions where the complete frequency band gaps exist

III. EXPERIMENTAL RESULTS AND DISCUSSION

The designed fractal PBG structure was microfabricated in a four mask process. Fig. 9 presents the scanning electron microscope (SEM) images of the fabricated structure and devices.

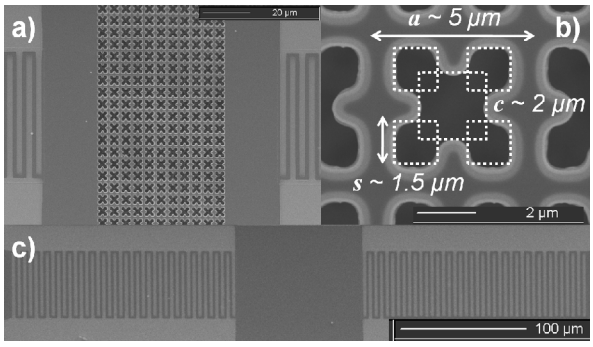


Figure 9. SEM images of the micro-fabricated a) fractal PBG array sandwiched by the Lamb wave transducers with b) the zoomed-in view of the fractal unit cell with key parameters labeled. c) The piezoelectric acoustic delay-line device used as the acoustic transmission reference.

In order to verify the PBG performance over a large frequency range (0.8 to 1.2 GHz), 14 sets of AlN piezoelectric transducers centered at various frequencies were integrated with the fractal PBG. Along with the PBG devices, 14 sets of the piezoelectric acoustic delay lines (Fig. 9c) matching the frequencies of the PBGs were fabricated simultaneously and served as a reference to which the PBG acoustic transmission was normalized. Approximately 4-5% bandwidth of the center

frequency of each of the 14 transducers was taken and summed up to realize the response over the entire frequency range. Each response was normalized to the corresponding acoustic transmission line to report the pure PBG response (independent of electrical and acoustic mismatches between the transducers and the delay line element).

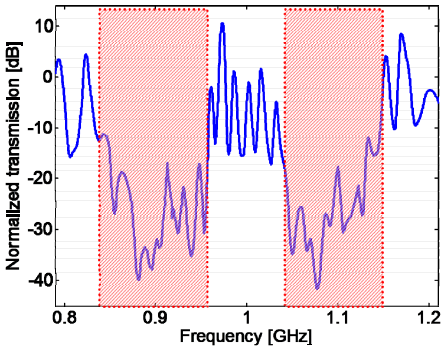


Figure 10. Experimental frequency response of the microfabricated fractal PBG device. The two main frequency band gaps are highlighted in red.

The experimental result indicates that there are two frequency stop bands: one from 850 to 950 MHz (with 11.1% gap-to-midgap ratio); and another from 1.05 to 1.15 GHz (9.09% gap-to-midgap ratio) with maximum attenuation of 40 dB. Although two frequency stop bands around 1GHz were attained, these are not the two complete band gaps for which the PBG was designed. According to the simulation (Fig. 11), the designed band gaps have shifted to higher frequencies because of fabrication imperfections that resulted in side squares with rounded edges.

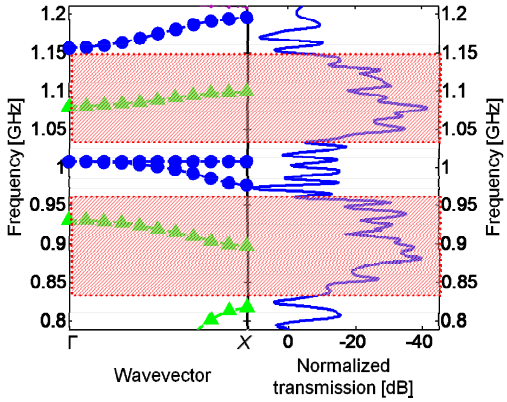


Figure 11. Dispersion curve obtained from COMSOL® simulations (blue dots represent the symmetric Lamb waves; and the green triangles represent the Asymmetric Lamb waves) are compared the the experimental data. The effect of rounded edges in the unit cell have been taken into account and justify the shift in frequency of the acoustic band gaps with respect to the design values.

The rounded edges (Fig. 9b) effectively change the acoustic velocity of each mode, and cause an upper shift in the frequency of operation of the PBG. Nonetheless, the fabricated PBG exhibits two large band gaps because the AlN lamb wave transducers mainly launch symmetric modes in the Γ-*X* direction and exclude the propagation of asymmetric modes. Although a non-complete frequency band gap was demonstrated, the experiment shows that the integration of the fractal design with same-plane AlN transducers can yield wide regions of little acoustic transmission. Edge rounding appear to have a more pronounced effect on the center frequency of operation of the PBG than the regular dimensional scaling that was accounted for in the design section. Future efforts will be devoted to eliminate the design sensitivity to edge rounding.

IV. Conclusion

In this paper, a novel fractal PBG design in AlN has been studied via FEM analysis and experimentally demonstrated to work in the UHF range when integrated in the same plane with AlN lamb-wave transducers. This work is a clear evidence that GHz operation in PCs can be achieved without drastic PBG miniaturization (minimum feature of 700 nm). As GHz operation was attained, it is now possible to consider utilizing these PBG structures for the making of radio frequency (RF) devices such as resonators or filters. Simultaneously, modified fractal designs will be explored in order to relax the dependence of the PC acoustic characteristics on manufacturing variations.

Acknowledgment

This work was funded by DARPA under the MTO/CSSA program. The authors would also like to thank the staff of Wolf NanoFabrication (WNF) facility at the University of Pennsylvania.

References

[1] S, Mohammadi, A. A. Eftekhar, A. Khelif, W. D. Hunt, and A. Adibi, "Evidence of large high frequency complete phononic band gaps in silicon phononic crystal plates," Appl. Phys. Lett., vol. 92, 221905, June 2008.

[2] R Olsson III and I El-Kady, "Microfabricated phononic crystal devices and applications," Meas. Sci. Technol., vol. 20, 012002, November 2008.

[3] S. Mohammadi, A. A. Eftekhar, W. D. Hunt, A. Adibi, "High-Q micro-mechanical resonators in a two-dimensional phononic crystal slab," Appl., Phys. Lett., vol. 94, 051906, January 2009.

[4] C.-Y. Huang, J.-H. Sun, T.-T. Wu, , "A twp-port ZnO/silicon Lamb wave resonator using phononic crystals," Appl. Phys. Lett. vol. 97, 031913, July 2010.

[5] M. Ziaei-Moayyed, M. F. Su, C. Reinke, I. F. El-Kady, R. H. Olsson III, "SILICON CARBIDE PHONONIC CRYSTAL CAVITIES FOR MICROMECHANICAL RESONATORS," IEEE MEMS 2011 Conference, Cancun, Mexico, January 2011, pp. 1377-1381.

[6] N. Kuo, G. Piazza, "Evidence of Acoustic Wave Focusing in a Microscle 630 Mhz Aluminum Nitride Phononic Crystal Waveguide," IEEE Frequency Control Symposium, Newport Beach, CA, USA, June 2010, pp. 530-533.

[7] T. Gorishnyy, C. K. Ullal, M. Maldovan, G. Fytas, E. L. Thomas, "Hypersonic Phononic Crystals," Phys. Rev. Lett., vol. 94, issue 11, 115501, 2005.

[8] J. H. Kuypers, C.-M. Lin, G. Vigevani, A. P. Pisano, "Intrinsic temperature compenstation of aluminum nitride Lamb wave resonators for multiple-frequency references," IEEE Frequency Control Symposium, Honolulu, HI, USA, May 2008, pp. 240.

A new acoustic resonator concept based on acoustic waveguides using Silicon/Periodically Poled Transducer/Silicon structures for RF applications

Florent Bassignot, Gwenn Ulliac, Jean-Marc Lesage*, Julien Garcia, Emilie Courjon, Bruno François, Sylvain Ballandras

Institut FEMTO-ST, Besançon, France
*CELAR, DGA, Rennes, France
Email: florent.bassignot@femto-st.fr

Abstract: We propose a new acoustic resonator concept based on a periodically poled transducer (PPT) in a piezoelectric substrate ($LiNbO_3$ or $LiTaO_3$), embedded between two guiding substrates in order to create an acoustic waveguide. A resonator operating at 131MHz have been successfully fabricated and used in order to stabilize an oscillator. However the fabricated resonator presents a significant thermal sensitivity. The following experiments have consisted in studying a Si/thinned PPT layer/Si in order to reduce the thermal sensitivity.

I. INTRODUCTION

Surface Acoustic Waves resonators are usually used for the frequency stabilization of electrical oscillators in the range of 50 MHz to 2 GHz. We have developed an isolated wave guide allowing for the propagation of acoustic waves within a PPT plate ($LiNbO_3$ or $LiTaO_3$) in between two silicon guiding substrates, yielding advanced packaging opportunities. In the first part, we present the simulation, the fabrication and the characterization of a PPT-based resonator operating at 131 MHz with a Silicon(380µm)/$LiNbO_3$ PPT(500µm)/Silicon(380µm) structure. We present the resonator oscillation characteristics, more particularly the phase noise and the frequency stability. In a second part, we exhibit the advantages of thinning down the piezoelectric transducer in showing the dispersion properties analysis. We describe the fabrication process and the characterization of resonators based on thinned PPLN or PPLT layers. The phase noise of an oscillator using a Si/PPLN/Si resonator, exhibiting a 50µm period and a 500µm thick piezoelectric layer, has pointed out the concept feasibility and answers to scope statement. In fact, these resonators, with low insertion losses and a phase noise at 100 kHz of the carrier wave controlled, less than 160dBc/Hz, are well-suited for radar applications. Moreover, for each PPT period, the dispersion properties analysis has shown an optimum thickness of the piezoelectric plate, corresponding to the maximum electromechanical coupling coefficient. Si/PPT/Si test vehicles, exhibiting a 50µm period and a thick piezoelectric layer around 30µm, have pointed out the existence of an isolated mode operating at frequencies near 110MHz for a $LiNbO_3$ PPT and near 90MHz with $LiTaO_3$ PPT with low thermal sensitivities.

II. BASIC PRINCIPLES OF THE PPT-BASED RESONATOR

We propose a new guiding concept using interface ultrasonic waves. We suggest a compact structure based on poled ferroelectric layer bonded between two guiding substrates (fig.1). An acoustic wave is excited by a periodically poled transducer, and the waves then are trapped within a buried guiding structure. $LiNbO_3$ or $LiTaO_3$ as piezoelectric and ferroelectric materials are exploited to fabricate the PPT's [1] and silicon substrates are used to guide the acoustic wave and allow for a low cost packaging [2]. These components have to operate at high frequencies (RF until X band domain) and need spectral purity, short-term stability, low phase noise around the carrier wave and thermal insensitivity notably for radar applications. PPTs have been investigated recently as an alternative to standard periodic inter-digital

978-1-61284-110-6/11/$26.00 ©2011 IEEE

transducers (IDTs) for the excitation and detection of guided acoustic waves [2] and consist of a periodic alternation of oppositely poled ferroelectric domains included between two electrodes. The main advantages of PPTs compared to standard IDTs are the robustness of the excitation versus defects or surface contamination, the opportunity to excite fundamental waves exhibiting an operating frequency twice higher and the possibility to excite several waves at the same time.

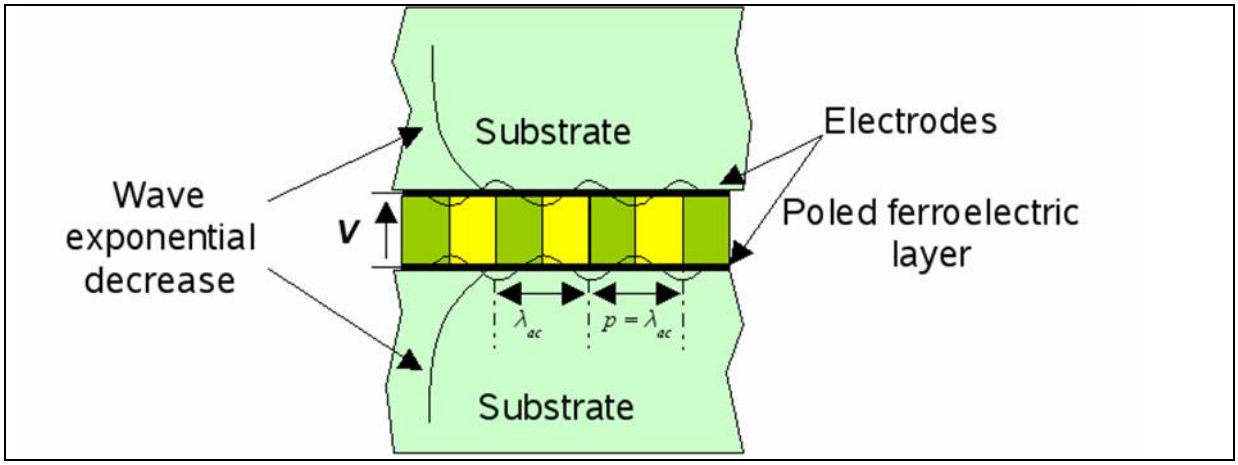


Fig. 1. Principle of the PPT isolated wave transducer

III. Si/thick PPT/Si- BASED WAVEGUIDE, RESONATOR AND OSCILLATOR

The fabrication of acoustic waveguides based on PPTs consists in bonding a silicon wafer on each side of the periodically poled wafer. $LiNbO_3$ or $LiTaO_3$ as piezoelectric and ferroelectric materials are good candidates for the fabrication of the PPTs. In fact, these materials have a spontaneous polarization which can be inverted by applying high voltage thanks to our home-made poling technique [3]. The voltage needed to invert the domains is around 11kV for a 500µm-thick wafer. The different steps of fabrication are detailed in fig. 2 and reported in reference [4].

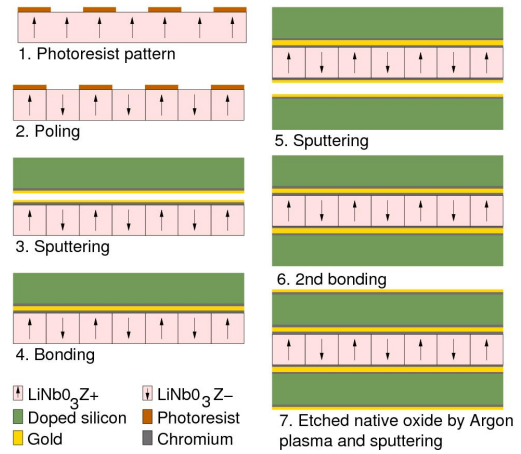


Fig.2. Flow-chart of fabrication

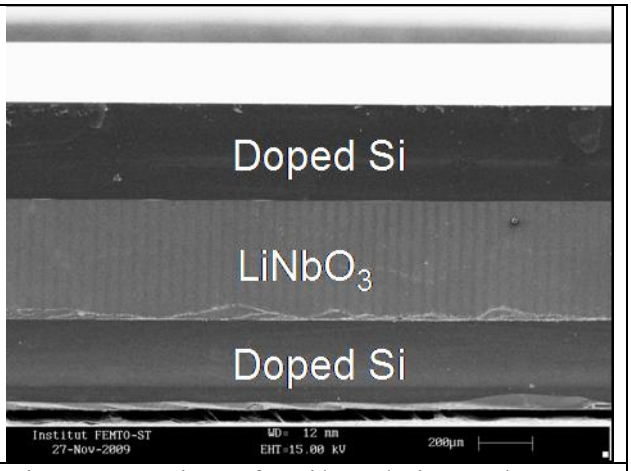


Fig.3. SEM view of a Si/PPT/ Si transducer

On figure 3, we can see a SEM observation of poled domains after HF wet etching for a 50µm period of this Si/500µm PPT/Si structure. The different domains confined between the 2 doped silicon layers are clearly observed. Operational test vehicles with a 50µm period have been achieved using doped silicon wafers in order to ease the electrical contact and tested by means of a spectrum analyzer yielding Si/PPT/Si devices responses.

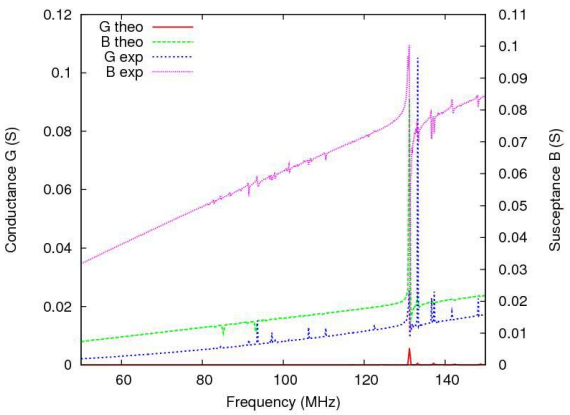


Fig.4. Theoretical and experimental admittances of a Si/$LiNbO_3$(500µm)/Si structure

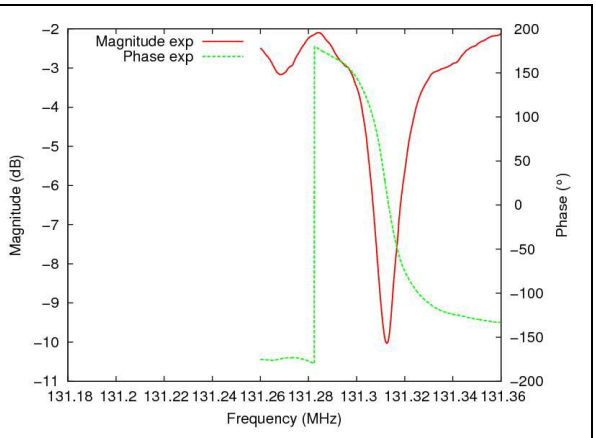


Fig.5. Evolution of the dynamic of the coefficient S_{11} and the phase rotation versus frequency for the longitudinal mode

Theoretical and measured electrical admittances agree well and allow for identifying a main contribution corresponding to a guided longitudinal mode at 131 MHz (fig.4). The corresponding phase velocity is very close to the one of the PPT alone (i.e 6500 $m.s^{-1}$) [3].
This resonator operating near 131MHz exhibits a quality factor of 13000 and an electromechanical coupling ks^2 equal to 0.25% (twice higher than a SAW quartz resonator). The corresponding phase rotation (320°) and the dynamic of its electrical reflection coefficient S_{11} equal to -8dB reveal compatible with oscillator applications (fig.5).

A specific printed circuit has been implemented for the phase noise tests with our Si/500µm PPT/Si resonator and the Colpitts oscillator (fig.6). The phase noise of the oscillator at 100kHz from the carrier shows a value less than -160dBc/Hz (fig.7), which can be honestly compared with other acoustic wave oscillators at such frequency, accounting for the fact that the device was excited with a quite low signal level (-6dBm). Therefore, increasing the excitation should allow for a significant reduction of the noise floor and then advantageously compete with standard solutions.

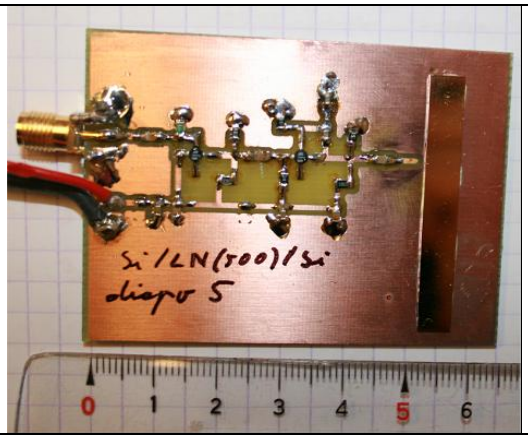


Fig.6. The oscillator board implemented for phase noise tests

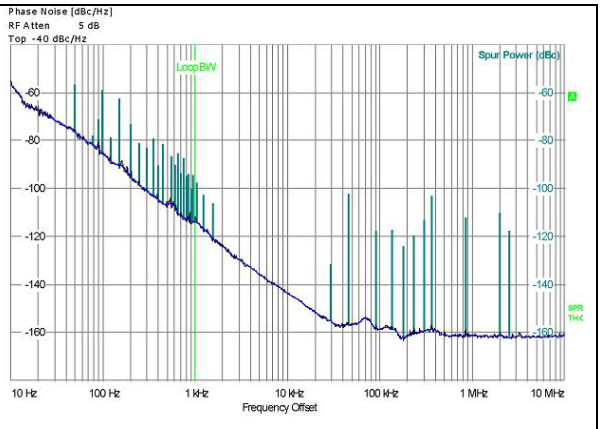


Fig.7. Phase noise of the 131 MHz oscillator stabilized

We have also studied the stability of this 131MHz oscillator, and we have obtained a frequency stability equal to 10^{-9} at less than 1s (fig.8). This experiment has been made at room temperature without any thermal compensation.

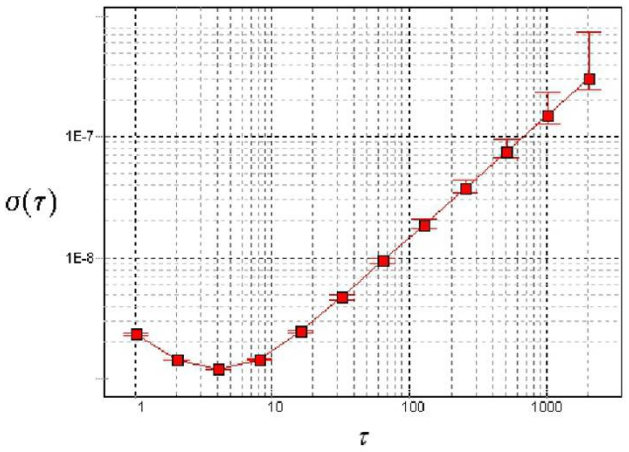


Fig.8. Frequency stability curve of 131 MHz oscillator

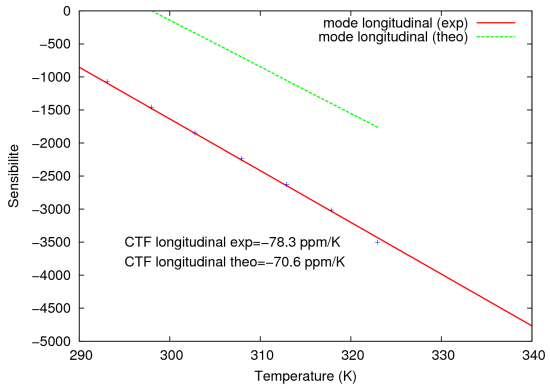


Fig.9. Frequency drift versus temperature for the longitudinal mode

The frequency drift of this longitudinal mode at 131MHz versus the temperature is reported in fig.9. We have found frequency variations of -78ppm/K for the longitudinal mode for a thick $LiNbO_3$ layer bonded between two Silicon layers. This value is sensibly the same than the one of isolated $LiNbO_3$ devices. We can conclude that we obtain a significant thermal sensitivity for Si/thick $LiNbO_3$ layer/Si and this value depends essentially of the material used for the PPT layer and cannot be reduced by the two guiding substrates. The further experiments consists in studying a Si/thinned PPT layer/Si in order to obtain a monomode acoustic wave device with an isolated mode with a low thermal sensitivity.

IV. Si/thinned PPT/Si- BASED WAVEGUIDES

The study of the dispersion properties enables to define a specific configuration using a thinned PPT layer [4]. Interesting guidance properties have been obtained for a ratio between the thickness of the layer and the period of inverted domains lower than 1 and as a result the maximum of the electromechanical coupling is obtained for a 25-35µm thick $LiNbO_3$/$LiTaO_3$ layer for a 50µm period. The global flow-chart to fabricate Si/thinned PPT layer/Si acoustic waveguides is approximately the same than the process used for Si/thick PPT/Si (fig.2). The supplementary step consists in lapping and polishing the PPT layer in order to obtain a thinned PPT layer of about 25-35µm, thanks to our lapping machine (fig.10). Finally this material stack is bonded on a second doped silicon wafer in order to define the aimed compact structure Si/PPT/Si. On figure 11, we can see a SEM observation obtained for a 25µm thick $LiNbO_3$ layer confined between the two silicon substrates for a 50µm period.

Fig.10. Photograph of the equipment used for lapping/polishing operations

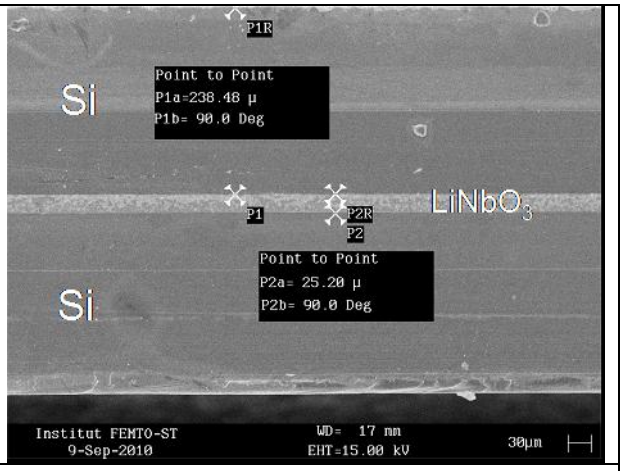
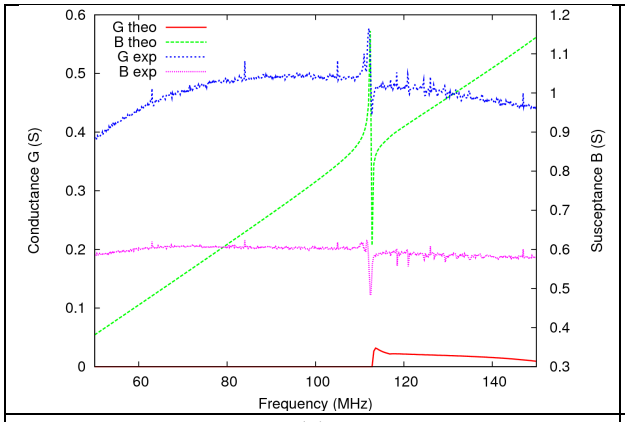
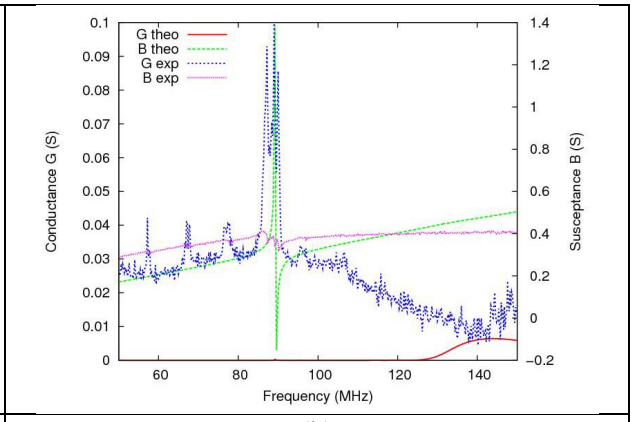


Fig.11. SEM view of a Si/thinned PPT/Si transducer

Operational test vehicles have been achieved using $LiNbO_3$ and $LiTaO_3$ wafers as poled ferroelectric layer. Experimental results of the Si/thinned PPT layer/Si devices with a 50µm period are reported in fig.12. For the two materials, only one contribution is obtained: at 110MHz for a 25µm thick $LiNbO_3$ layer and at 88MHz for a 35µm thick $LiTaO_3$ layer. Unfortunately, these monomode acoustic devices, due to a non homogenous thickness of the PPT layer along the several millimeter long ferroelectric gratings, don't exhibit properties compatible with oscillator applications.



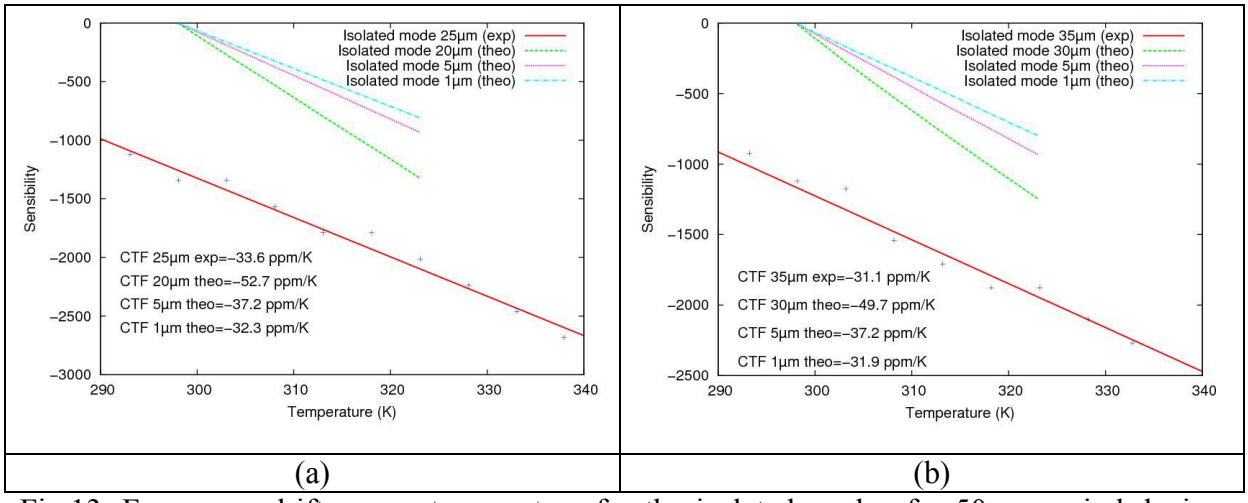
(a)



(b)

Fig.12. Theory/experitment assessment for a Si/thinned PPT/Si device with 50µm of poling period built on a Z-cut $LiNbO_3$ plate (25µm) (a) and a Z-cut $LiTaO_3$ plate (35µm) (b)

We have also studied the frequency drift of the isolated mode versus the temperature for a 50µm period structure (fig.13). We have found frequency variations of -33ppm/K for the isolated mode for a 25µm thick $LiNbO_3$ layer and of -31ppm/K for a 35µm thick $LiTaO_3$ layer. So we can conclude that bonding onto two silicon wafers and lapping/polishing the $LiNbO_3$ layer have compensated this frequency drift versus temperature by a factor 2. Moreover the temperature coefficient of frequency (TCF) value of the compact structure is close to the TCF value of the silicon guiding substrate. So we can conclude that if we choose an appropriate insensitive material as guiding substrate, we can obtain an insensitive resonator versus temperature.



(a) (b)

Fig.13. Frequency drift versus temperature for the isolated mode of a 50µm period device-theory/experiment assessment on a Z-cut $LiNbO_3$ plate (25µm) (a) and a Z-cut $LiTaO_3$ plate (35µm) (b)

V. CONCLUSION

We have proposed a new acoustic resonator concept based on a periodically poled transducer in a piezoelectric substrate ($LiNbO_3$ or $LiTaO_3$) embedded between two silicon guiding substrates in order to create an acoustic waveguide. A resonator operating at 131MHz have been successfully fabricated. This resonator presents the required properties to stabilize an oscillator at this frequency but however the developed resonator presents a significant thermal sensitivity (depending mainly on PPT substrate used). The following experiments have consisted in studying a Si/thinned PPT layer/Si: monomode acoustic wave devices with an isolated mode have been successfully fabricated with a low thermal sensitivity. A thermal insensitivity is possible with an appropriate choice of guiding susbtrate. However these monomode devices don't exhibit properties compatible with oscillator requirements due mainly to technological limits to be solved. Future works consist in improving the lapping process in order to obtain a homogeneous thickness for the compact structure, studying the possibility to reduce the noise floor of the oscillator with increasing the input power of the resonator at 131MHz and trying to invert micrometric periods in order to achieve a 2GHz frequency resonator.

ACKNOWLEDGMENT

This work is supported by the French DGA (Délégation Générale de l'Armement) under grant #0734020.

REFERENCES

[1] E. Courjon et al, "Pure longitudinal plate mode excited by poled domains transducers on $LiNbO_3$", Proc. of the IEEE IFCS, 2007
[2] E. Courjon et al, "Fabrication of periodically poled domains transducers on $LiNbO_3$", Proc. of the IEEE IFCS, 2006
[3] E. Courjon et al, "Lamb wave transducers built on periodically poled Z-cut $LiNbO_3$ wafers", Journal of Applied Physics, 102, 114107, 2007
[4] E. Courjon et al, "An acoustic wave guide based on doubly-bonded silicon/PPT/silicon structures", Proc. of the joint IEEE IFCS-EFTF, 2009

High quality factor lithium niobate electrostrictive HBAR resonators

M. Pijolat, C. Deguet, C. Billard, D. Mercier, A. Reinhardt, M. Aïd, E. Defay
CEA LETI Minatec Campus
Grenoble, France
edefay@cea.fr

S. Ballandras
FEMTO-ST UMR CNRS-UFC-ENSMM-UTBM 6174
Besançon, France

***Abstract*— Electrostrictive acoustic resonators have been widely investigated during the last 8 years. Despite their interesting dc-controlled electromechanical coupling and switchable properties, the main drawback remains their poor quality factor. Here we show that HBAR LiNbO3 single crystals can be useful to act as high quality factor switchable resonators thanks to their electrostrictive properties and to their intrinsic low acoustic losses. For a resonance frequency at 1.95GHz, HBAR Q reaches 9000. This outstanding behaviour could be of interest for RF applications.**

I. INTRODUCTION

Since 2003, several groups have been working on electrostrictive acoustic resonators (EARs) as an alternative to the well-established piezoelectric resonator [1,2,3]. The latter is used as the building block for RF filters for telecom purpose [4,5]. Although the commercial success of these devices is huge, it is highly desired to add controllable abilities to resonators in order to simplify RF architectures [6]. A solution would be using electrostrictive materials instead of piezoelectric ones. The best identified materials belong to the perovskite family, including $SrTiO_3$ (STO), $Pb(Zr,Ti)O_3$ (PZT) and $(Ba,Sr)TiO_3$ (BST), the last being the most studied one. Indeed, using such materials gives the opportunity 1) to switch on/off the resonator as desired simply by adding a dc voltage superimposed to the RF signal as pointed out by Volatier et al. [7] 2) to tune the resonance frequencies by adjusting the dc voltage due to linear and non linear electrostriction as described by Vendik in 2008 [2]. Although such properties might deeply change future RF systems architectures, the main EARs drawback today is their low resonance quality factor Q, which is a figure of merit of the resonance and antiresonance peak sharpness. The majority of the papers dealing with EAR reports Q values between 50 and 150 [1,2,3,7], ten times lower than the required values for filtering devices, fulfilled with AlN resonators for which quality factors reach 1000 or even 2000 [4]. To be attractive, EARs need to overcome this issue. The physical explanation of the low Q factor of EARs is manifold. The global material quality is obviously involved as for BST and STO films we are far from the AlN thin films quality in terms of density, orientation control and phase purity. Recently, Gevorgian's team published Q values reaching 350 for hot sputtered BST thin films which is a significant breakthrough [8], even though still far from the targeted value exhibited by pure piezoelectric thin films. Besides, the domain wall boundaries motion in lead based materials like PZT is an intrinsic reason for the limited Q factors [9]. Identified some years ago by Arlt [10] it was pointed out more recently by Muralt's team [11]. Avoiding domains is why people are keener in developing non ferroelectric materials as STO and BST instead of PZT. Another possible reason is the film roughness, which is related to the material quality as well [13]. To overcome these three main drawbacks, we propose hereafter to use single crystals $LiNbO_3$ (LNO) thin films. Indeed, LNO belongs to the ilmenite family and exhibits only one possibility for the polarization direction, avoiding non-180° domains and consequently energy dissipation due to domain wall motion. Moreover, single crystals exhibit a prefect structure, without any grains boundaries and misorientation. Single crystals thin films can be obtained by using a transfer process based on ion implantation / molecular bonding on host substrate / blistering by heating. This technique, called Smart Cut™ has been recently adapted from Silicon On Insulator (SOI) purposes to LNO thin films onto LNO carrier substrate in collaboration with SOITEC [12]. Another specificity of this study is that the carrier substrate is kept during all experiments. Indeed, one should consider this study as a first step towards the released single crystal LNO thin film which represents a tremendous technological challenge. . Therefore, in the following, we will focused on the dc-controlled High overtone Bulk Acoustic Resonator (so-called HBAR) behavior of LNO single crystals thin films transferred on LNO substrate.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

II. EXPERIMENT

A. *Technology*

The technology used to perform the sample is described elsewhere [12]. Both thin films and substrate are X-Cut LNO. The transferred film is 800nm thick. Handle and donor wafers are both full 3-inch LNO wafers. The bottom electrode (Pt 100nm) was sputtered on the handle wafer. Bonding was achieved between two SiO_2 layers deposited respectively on the metallic layer and the handle wafer. The 100nm-thick sputtered top electrode is made of Pt and patterned by lift off technique. The LNO transferred layer is indeed a single crystal exhibiting an electromechanical coupling coefficient due to piezoelectricity as high as 40%, which is equivalent to bulk values [12].

B. *Characterization*

The device under test corresponds to a Metal Insulator Metal (MIM) configuration. RF characterization is performed by using a Vector Network Analyser (VNA) in the 2GHz range. A dc field is superimposed to rf signal.

III. RESULTS

Figure 1 shows the real part of the MIM capacitance impedance as obtained versus different dc voltages between 1942 and 1958MHz. One can observe two types of resonances: P (piezoelectric) and E (Electrostrictive). On one hand, the highest peaks (P) are related to the piezoelectric shear waves always appearing whatever the dc voltage applied. Interestingly, on the other hand, another peaks type (E) appears only when dc is on, as zoomed on inset b), Figure 1. Its height depends on the dc voltage amplitude. Peaks repetition is due to the substrate. They correspond to harmonics of the fundamental substrate longitudinal thickness mode of vibration. Although the nature of the P peaks is well identified (shear wave in X-LNO), E type is not common. The frequency distance between two substrate peaks (Δf) is related to the sound velocity v and the substrate thickness t_s by $v=2 t_s \Delta f$.

Therefore, if one assumes the wave propagation and the resonance frequencies are mainly influenced by the substrate, v_E extracted from E type gives $v_E=6550 m.s^{-1}$.

As $v = \sqrt{c^D / \rho}$ (with c^D the stiffness coefficient and ρ the density (4630 kg/m^3 for LNO)), one deduces $c^D=1.99.\ 10^{11}$ N/m^2. This value corresponds to $c_{33}{}^D$ in X-cut. Therefore, this sound velocity is the evidence that dc voltage in LNO induces a longitudinal acoustic wave which is not coupled in X-cut while its corresponding piezoelectric constant is zero. The longitudinal coupling is possible here as dc voltage is combined to the second order electromechanical coupling called electrostriction.

Figure 2 is a zoomed part of Figure 1 focusing on one electrostrictive peak. One can clearly observe that at 0V dc, no resonance is visible. As soon as dc voltage is applied, resonance (ReY) and antiresonance (ReZ) peaks appear. The frequency positions of these peaks depend on the dc voltage value, which is typical for electrostriction. It is essential noting that the quality factor Q of the highest peak reaches 9000 at 1955MHz, inducing Qf factor reaching 1.7 10^{13} Hz.

One should bear in mind that this very high Q is mainly influenced by the LNO substrate which exhibits lower acoustic losses than silicon. As the acoustic wave spend much more time in the substrate than in the thin film, its influence on Q is huge. It is also interesting to compare this Q with the one exhibited by the shear wave resonance named *P* (as Piezoelectric) as observed in Figure 1. Indeed, this last reaches 45,000 as already reported [13]. This is 5 times higher than the *E* resonance. The first reason is that shear waves are always less lossy than longitudinal waves. Moreover, one should consider the accuracy a of the applied dc voltage which is typically 0.02% as given by the supplier. If one consider that the corresponding quality factor Q_S of this voltage supply is roughly $1/a$, than $Q_S=5000$. Therefore, it is likely that the resonator Q is not intrinsically limited by electrostriction on the current characterization.

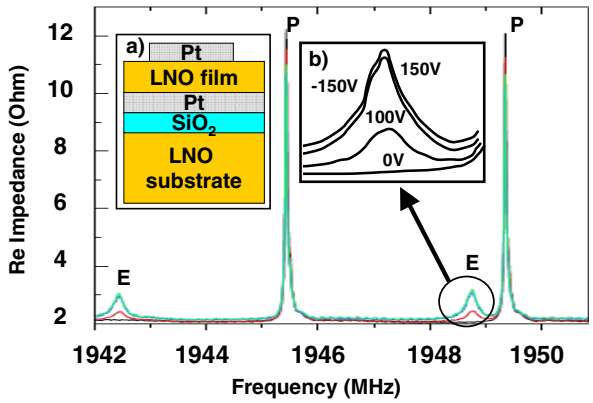


Figure 1. real part of the impedance of LNO MIM capacitor in HBAR configuration. Inset (a): schematic of the cross section of the HBAR resonator. Inset (b): zoom of the dc voltage sensitive peak E, as a result of the electrostrictive effect

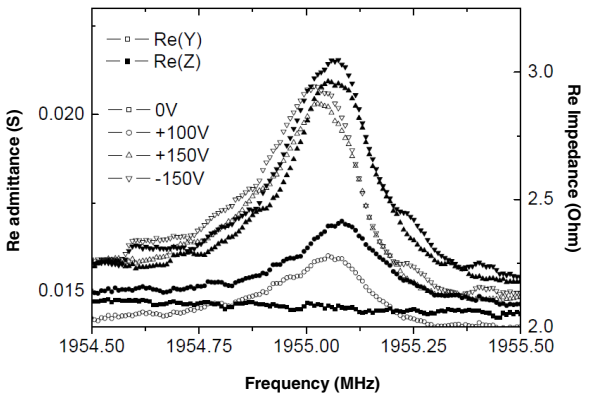


Figure 2. real part of admittance and impedance of the resonator versus frequency with respect to applied dc voltage. Resonances appear when dc voltage is different from zero.

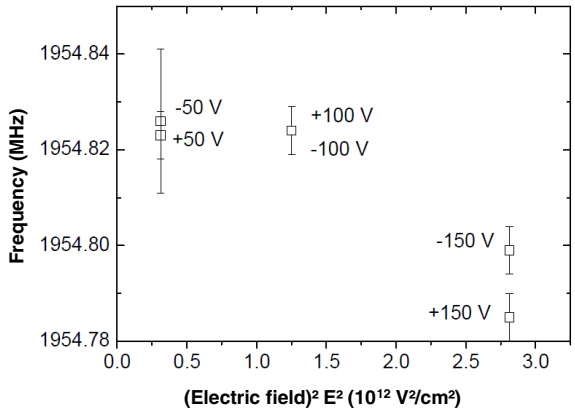


Figure 3. antiresonance frequency extracted from Re(Z) of the resonator versus the square of the dc electric field applied

Apart from Q, one can estimate the tuneable properties of the resonances. An antiresonance frequency f_p corresponding to harmonic at 1954.84 MHz is reported in Figure 3 versus the square of the applied dc electric field. Unlikely one can observe in high-K perovskite electrostrictive materials such as BST [3], the variation of f_p is not linear with E^2. However, this variation seems to be always symmetrical versus the dc electric field E_{dc} as f_p is roughly identical for a given E_{dc} whatever its sign. This suggests how to manage the free energy form in order to explain this variation compared to what was proposed for electrostrictive materials [2]. The idea is to add a fourth order term in the stiffness coefficient, leading to a six order term in the free energy, as reported in the following equation.

$$F - F_0 = \frac{1}{2}c_0^D S^2 + \frac{1}{2}\beta^S D^2 + \frac{1}{4}\gamma^S D^4 + GD^2 S + \frac{1}{2}MS^2 D^2 + \frac{1}{2}M' S^2 D^4 \quad (1)$$

For this equation, we use the following parameters - *S:* strain, T : stress, D : electric displacement, E : electric field, c_0^D : constant value of stiffness, β^S: impermittivity, γ^S: non linear impermittivity, G: first order electrostriction, M: non linear electrostriction, M': fourth order non linear electrostriction. Only *M'* term was added to the standard model. *S* is used as the mechanical variable as it permits working with a 1D model as the substrate avoids all the in-plane strain. Besides, *D* (electrical displacement) is used as electric variable as the antiresonance frequency does not depend on any coupling, but just on c^D and density [2]. For the following, we focus on the stiffness term only. Therefore, after differentiating *F* twice with respect to *S*, one obtains c^D:

$$c^D = c_0^D + MD^2 + M'D^4 \quad (2)$$

This expression of LNO layer stiffness is then introduced in the standard acoustic model of HBAR developed by Lakin [14]. Consequently, one can extract c^D variation from f_p variation experienced by the HBAR when the dc field is applied. It is worth noting that it would be more accurate and convenient to extract c^D variation from a released resonator (Film Bulk Acoustic Resonator (FBAR) configuration). But important technological issues have not been overcome yet. By using f_p variation between 50 and 150V from Figure 3, the extracted stiffness variation for the LNO layer is -120ppm/V. By supposing that the D^4 term is the most prominent one in equation (2), one can deduce that *M'* is -2.5 10^{14} N.m^6/C^4. This relative variation of stiffness can be used to calculate what would be the frequency variation of LNO thin films alone, standing as FBAR resonator. It can be compared to other reported stiffness variations of other materials as AlN [15] and BST. For sake of clarity, a correct comparison figure is *R* the relative antiresonance frequency variation versus applied electric field ($R=\Delta f_p/f_p.1/E$). One obtains R_{AlN}=-$2.6.10^{-11}$m/V, R_{LNO}=-$4.8.10^{-11}$m/V and R_{BST}=-$1.0.10^{-10}$m/V. The LNO frequency variability is around 1% for the maximum applied field which is the same order of magnitude exhibited at the same time by AlN and BST. This rather equivalent variability is due to different breakdown electric fields of these three materials (typically 1MV/cm, 2.5MV/cm and 6MV/cm respectively for BST, LNO and AlN).

IV. CONCLUSION

In this paper, we showed that $LiNbO_3$ single crystals HBAR resonators can be useful to act as high quality factor switchable resonators thanks to their electrostrictive properties and to their intrinsic low acoustic losses. For a resonance frequency at 1.95GHz, Q reaches 9000. This outstanding behavior could be of interest for RF applications.

REFERENCES

[1] S. Gevorgian, A. Vorobiev, T. Lewin, J. Appl. Phys., 99, 124112 (2006)

[2] Irina B. Vendika, Pavel A. Turalchuk, Orest G. Vendik, John Berge, J. Appl. Phys., 103, 014107 (2008)

[3] A. Noeth, T Yamada, V O. Sherman, P Muralt, A K. Tagantsev, N Setter, J. Appl. Phys 102, 114110 (2007)

[4] R. Ruby, IEEE Radio Frequency Integrated Circuits Symposium, 2009. 7-9 June 2009

[5] Dubois MA, Carpentier JF, Vincent P, Billard C, Parat G, Muller C, Ancey P, Conti P, IEEE Journal Of Solid-State Circuits 41, p. 7 (2006)

[6] Robert Aigner, Proceedings of the 40th European Microwave Conference, P. 787

[7] A. Volatier, E. Defay, M. Aid, A. Nhari, P. Ancey; B. Dubus, Appl. Phys. Lett. 92, 032906 (2008)

[8] S. Gevorgian, A.Vorobiev, Appl. Phys. Lett., 96, 212904 (2010)

[9] J. Conde, P. Muralt, IEEE trans. on ultrasonics, ferrœlectrics, and frequency control, 55(6), p. 1373 (2008)

[10] G. Arlt, U. Böttger, and S. Witte, Appl. Phys. Lett. 63, 602 (1993).

[11] P. Muralt, J. Antifakos, M. Cantoni, R. Lanz, F. Martin, IEEE Ultrasonics Symposium, 315 (2005)

[12] J. S. Moulet, M. Pijolat, J. Dechamp, F. Mazen, A. Tauzin, F. Rieutord, A.Reinhardt, E. Defaÿ, C. Deguet, B. Ghyselen, L. Clavelier, M. Aïd, S.Ballandras, and C. Mazuré, Proceedings of the International Electron Devices Meeting _IEEE, New York, 2008

[13] M. Pijolat, J. S. Moulet, A. Reinhardt, E. Defaÿ, C. Deguet, D. Gachon, B.Ghyselen, M. Aïd, and S. Ballandras, Proceedings of the International Ultrasonic Symposium, 201 _IEEE, Beijing,2008

[14] K.M. Lakin, G.R. Kline, K.T. Mc Carron, IEEE transactions on microwave theory and techniques, 41, p. 2139 (1993)

[15] Ben-Hassine-N; Mercier-D; Renaux-P; Chappaz-C; Basrour-S; Defay-E, J. Of Appl. Phys. 104 (3), 034110 (2008)

Motional Impedance of Resonators in the Nonlinear Regime

Hyung Kyu Lee, Renata Melamud, Saurabh Chandorkar, Yu Qiao Qu, James Salvia, and Thomas W. Kenny
Departments of Mechanical and Electrical Engineering
Stanford University, Stanford, California 94305, USA
Email: hyungkyu@stanford.edu

***Abstract*—We developed a model for the motional impedance Z_m of electrostatic resonators that operate in a nonlinear regime. The model predicts that the Z_m at resonance has the same expression as it has in the linear case. In addition, we developed a new measurement setup, which is based on a variable-phase feedback circuit. This setup enables the measurement of Z_m at operating points that are unobservable with a conventional impedance analyzer measurement. The estimation based on our model matches the measurement with errors less than 10% in the worst case.**

I. Introduction

Micromechanical resonator-based oscillators are considered by many to be the next generation of frequency references [1]. Frequency references are an essential component of electronic devices, and quartz crystal-resonator-based references have dominated the market. Although quartz references have many advantages, they have become a bottleneck in the miniaturization of electronic devices since their manufacturing process is not CMOS-compatible. In contrast, micromechanical resonators can be batch fabricated within a standard CMOS process. They thus have the potential to provide low-cost micro-scale frequency references.

Motional impedance Z_m of electrostatic micromechanical resonators describes the relation between a resonator's input drive voltage, $v_{ac}(t) = V_{ac}\cos(\Omega t)$, and output current, $i_{out}(t) = I_{out}\cos(\Omega t + \theta)$. It is called the "motional" impedance because it is the impedance that stems from mechanical motion of resonators. Estimating and measuring Z_m is important because it is a critical parameter to consider in the design of sustaining feedback amplifiers in a closed-loop oscillator. Furthermore, precise estimation of Z_m is necessary to engineer the phase-noise performance of resonator-based oscillators [2].

Because of its importance, there have been numerous studies on Z_m. Researchers have developed models that allow precise estimation of the Z_m of resonators operating in a linear regime [2]–[5]. However, most micromechanical resonators have nonlinearities. These nonlinearities can be effectively modeled with additional higher-order terms of mechanical displacement in the governing equation [5]–[7]; hence, their effect is negligible when the drive amplitude V_{ac}—which is proportional to the mechanical displacement—is small, but it must be accounted for when V_{ac} becomes larger. In resonator-based oscillators, large V_{ac} is preferred for better signal-to-noise ratio. Therefore, micromechanical resonators in oscillator applications often operate in the nonlinear regime, where the nonlinear effect is not negligible. Furthermore, a conventional measurement method cannot observe some operating points when a resonator is in the nonlinear regime because instability arises. To address all of these issues, a new measurement method must be developed for observing the nonlinear Z_m characteristic.

In this research, we first developed a nonlinear Z_m model. We then demonstrated an all-new closed-loop measurement setup. The measurement results from the setup match the developed model.

II. Theory

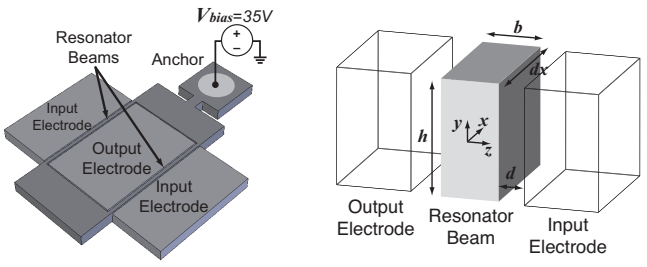
A. Nonlinear mechanical amplitude

The modeling starts with the partial differential equation model of a single-anchored double-ended tuning fork (SA-DETF) resonator, shown in Fig. 1(a). The Euler-Bernoulli equation for the differential beam, Fig. 1(b), is

$$\frac{\partial^2 w}{\partial t^2} + \frac{\hat{c}}{\mu}\frac{\partial w}{\partial t} + \frac{EI}{\mu}\frac{\partial^4 w}{\partial x^4} = \frac{\hat{q}(t)}{\mu}, \tag{1}$$

where $w(x,t)$ is the beam displacement in the $z-$direction, $\hat{c}$ is the damping coefficient per unit length, μ is the mass per unit length, E is Young's modulus, $I = \frac{hb^3}{12}$ is the second moment of area, and $\hat{q}$ is the net electrostatic force per unit length. For a two-port electrode configuration,

$$\hat{q}(t) = \frac{\varepsilon_0 h}{2}\left[\left(\frac{V_{bias} - v_{ac}}{d - w}\right)^2 - \left(\frac{V_{bias}}{d + w}\right)^2\right], \tag{2}$$



(a) SA-DETF resonator diagram (b) differential beam element

Fig. 1. Single-anchored double-ended tuning fork (SA-DETF) resonator and its differential element

where ε_0 is the dielectric constant of vacuum, V_{bias} is a dc-polarization voltage applied to a resonator beam, $v_{ac}(t)$ is an ac-stimulus signal with amplitude V_{ac}, and d is the transduction gap. Applying the Taylor series expansion to Eqn. (2) with respect to w with the assumption of $V_{ac} \ll V_{bias}$ and $w \ll d$, we get

$$\hat{q}(t) \approx -\frac{\varepsilon_0 h V_{bias}}{d^2}\left(v_{ac} - 2V_{bias}\frac{w}{d} - 4V_{bias}\frac{w^3}{d^3}\right). \quad (3)$$

Plugging Eqn. (3) into Eqn. (1), we obtain the governing equation:

$$\frac{\partial^2 w}{\partial t^2} + c\frac{\partial w}{\partial t} + \alpha_0\frac{\partial^4 w}{\partial x^4} - \frac{\beta_1}{\mu}w - \frac{\beta_3}{\mu}w^3 = \frac{\beta_0}{\mu}v_{ac}, \quad (4)$$

where $c = \frac{\hat{c}}{\mu}$, $\alpha_0 = \frac{EI}{\mu}$, $\beta_0 = -\frac{\varepsilon_0 h V_{bias}}{d^2}$, $\beta_1 = \frac{2\varepsilon_0 h V_{bias}^2}{d^3}$, and $\beta_3 = \frac{4\varepsilon_0 h V_{bias}^2}{d^5}$ are defined to make expression concise. Note that the w^3 term from Eqn. (3) introduces nonlinearity.

Preparing modal analysis, we obtain the mode shape functions $\phi_n(x)$: $\phi_n(x)$ is the nth mode Eigenfunction for the Eigenvalue problem

$$\mathcal{L}\left[\phi_n(x)\right] = \omega_n^2\phi_n(x) \quad (5)$$

with a clamped-clamped boundary condition [8], where

$$\mathcal{L} \triangleq \alpha_0\frac{\partial^4}{\partial x^4} - \frac{\beta_1}{\mu}. \quad (6)$$

Solving Eqn. (5), we get $\phi_n(x) = \cosh(k_n x) - \cos(k_n x) + \theta_n\left(\sinh(k_n x) - \sin(k_n x)\right)$, where $k_n l$ is a solution of the characteristic equation $\cos(kl)\cosh(kl) = 1$ and

$$\theta_n = \frac{\cos(k_n l) - \cosh(k_n l)}{\sinh(k_n l) - \sin(k_n l)}. \quad (7)$$

The characteristic equation can only be solved numerically, and it has infinitely many solutions of $k_n l$: for instance, $k_1 l \approx 4.7300$ for the fundamental mode, and $k_n l \approx 7.8532$, 10.9956, 14.1372, 17.2788... for higher modes. We then assume that

$$w(x,t) \approx \phi_1(x)\xi_1(t) \quad (8)$$

because in the general equation, $w(x,t) = \sum_{n=1}^{\infty}\phi_n(x)\xi_n(t)$ —where $\xi_n(t)$ are the generalized coordinates— ξ_1 prevails over the other ξ_i when the driving frequency Ω is near the fundamental natural frequency $\omega_1(\Omega \approx \omega_1)$ [9]. We now apply modal analysis to Eqn. (4) by plugging Eqn. (8) into Eqn. (4) and integrating along the beam length l. Using the orthogonality of the Eigenfunctions $\frac{1}{l}\int_0^l \phi_n\phi_m dx = \delta_{nm}$ (the Kronecker delta) [9] and the relation of $c = \frac{\omega_1}{Q}$ and $v_{ac}(t) = V_{ac}\cos(\Omega t)$, we obtain

$$\frac{d^2\xi_1}{dt^2} + \frac{\omega_1}{Q}\frac{d\xi_1}{dt} + \omega_1^2\xi_1 - \gamma\xi_1^3 = q\cos(\Omega t), \quad (9)$$

where Q is the quality factor, $\gamma = \frac{\beta_3}{\mu l}\int_0^l \phi_1(x)^4 dx$ is the nonlinearity parameter, and $q = \frac{\beta_0 V_{ac}}{\mu l}\int_0^l \phi_1(x)dx$ is the forcing parameter. Eqn. (9) describes the forced Duffing system [10], and it is widely applicable to other types of resonators because we will obtain an ordinary differential equation similar to Eqn. (9) if we apply modal analysis to the model for other types of resonators.

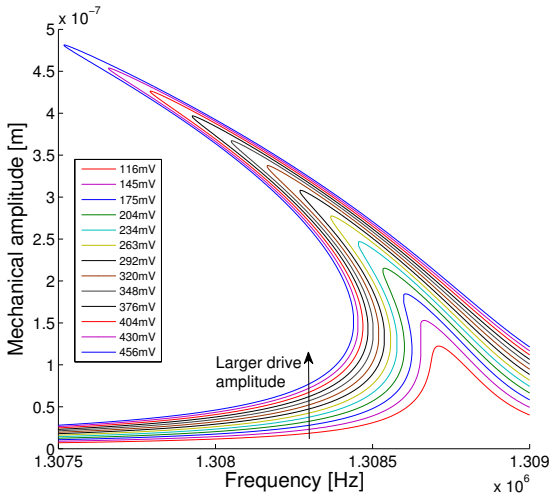


Fig. 2. Amplitude of the generalized coordinate, A in Eqn. (10), at various V_{ac} levels, shown in the legend. For each V_{ac}, we used Eqn. (11) to estimate A for a given range of a driving frequency Ω.

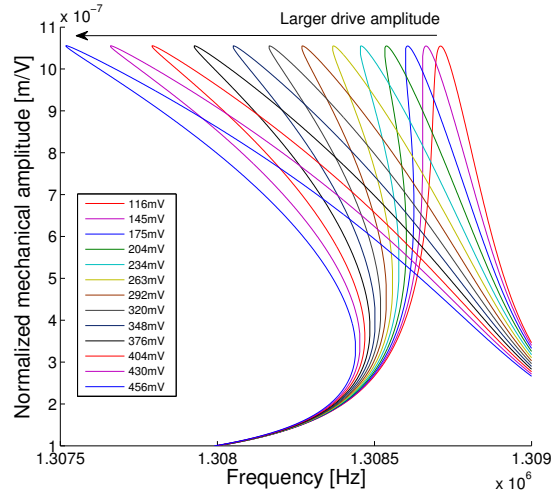


Fig. 3. Normalized mechanical amplitudes ($\frac{A}{V_{ac}}$) at various V_{ac} levels, shown in the legend. For each V_{ac}, we normalized the result in Fig. 2 with V_{ac}.

We obtain an approximate solution of Eqn. (9) by applying the method of multiple scales with the assumption of a weakly nonlinear system and a resonant excitation $(\Omega \approx \omega_1)$ [10]. The steady-state solution of Eqn. (9) is

$$\xi_1(t) \approx A\cos(\Omega t - \psi), \quad (10)$$

where A is given by the solution of the frequency response equation

$$A^2\left[\frac{\omega_1^4}{Q^2} + \left(\frac{3}{4}\gamma A^2 - 2(\Omega - \omega_1)\omega_1^2\right)^2\right] = q^2 \quad (11)$$

TABLE I
PARAMETERS FOR NUMERICAL ESTIMATIONS

beam length: l	200μm
beam width: b	6μm
beam height: h	20μm
transduction gap: d	1.5μm
mass density: ρ	2329 kg/m^3
Young's moduls: E	169 GPa
Polarization voltage: V_{bias}	35 V
Quality factor: Q	8710
Feedthrough capacitance: C_{feed}	40.36 fF

and ψ is

$$\psi = \arctan\left[\frac{-\omega_1}{2Q\left(\Omega - \omega_1 - (3\gamma/8\omega_1)A^2\right)}\right]. \quad (12)$$

Fig. 2 shows estimated A at various V_{ac} levels. The V_{ac} levels are selected to match the V_{ac} levels used for actual measurements shown in Fig. 8. Table I summarizes the parameters for the estimation.

By taking the derivative of Eqn. (11) with respect to Ω and setting $\frac{dA}{d\Omega} = 0$, we can find the resonant frequency Ω_r, which results in the maximum amplitude A_{max}. The Ω_r results in $\left(\frac{3}{4}\gamma A^2 - 2(\Omega_r - \omega_1)\omega_1^2\right)$ in Eqn. (11) being zero [6], [10]; hence, A_{max} becomes

$$A_{max} = \frac{|q|\,Q}{\omega_1^2} = \frac{\varepsilon_0 h V_{bias} V_{ac} Q}{\omega_1^2 d^2 m}\int_0^l \phi_1(x)dx, \quad (13)$$

where $m = \mu l$ is the mass of the beam. Surprisingly, A_{max} is not a function of nonlinearities, but is proportional to $V_{bias}V_{ac}Q$, as in the linear case [4], [11]. Fig. 3 shows the normalized mechanical amplitude ($\frac{A}{V_{ac}}$) estimated with Eqn. (11) at many driving frequency Ω: it shows that $\frac{A_{max}}{V_{ac}}$ is not affected by the level of V_{ac} as Eqn. (13) predicts, even though the resonant frequency Ω_r changes.

B. Nonlinear motional impedance

To estimate Z_m, we start from the total surface charge Q_s at the two output electrodes of a DETF resonator:

$$Q_s = -2V_{bias}\varepsilon_0 h \int_0^l \frac{1}{d-w}dx. \quad (14)$$

Applying the Taylor series expansion to the above equation with respect to w, we obtain

$$Q_s \approx -\frac{2V_{bias}\varepsilon_0 h}{d}\int_0^l \left(1 + \frac{w}{d}\right)dx. \quad (15)$$

Using the single mode assumption in Eqn. (8), we get

$$i_{motion} = -\frac{dQ_s}{dt} \approx \frac{2V_{bias}\varepsilon_0 h}{d^2}\frac{d\xi_1}{dt}\int_0^l \phi_1(x)dx. \quad (16)$$

Plugging Eqn. (10) into Eqn. (16) and defining the non-dimensional parameter $L_n = \frac{1}{l}\int_0^l \phi_n(x)dx$, we obtain

$$i_{motion} \approx \frac{2A\Omega V_{bias}\varepsilon_0 h l L_1}{d^2}\cos(\Omega t + \frac{\pi}{2} - \psi). \quad (17)$$

To estimate the maximum amplitude of i_{motion}, we substitute Eqn. (13) into Eqn. (17) and use the definitions of q and β_0 with the assumption of a resonant excitation ($\Omega \approx \omega_1$). We then get

$$I_{motion,max} \approx 2\left[\frac{V_{bias}\varepsilon_0 h l L_1}{d^2}\right]^2 \frac{V_{ac}Q}{m\omega_1}. \quad (18)$$

Note that $I_{motion,max}$ is not a function of nonlinearities but is proportional to $V_{bias}^2 V_{ac}Q$, as in the linear case [4]. In addition to i_{motion}, the feedthrough current

$$i_{feed} = C_{feed}\frac{dv_{ac}}{dt} = -C_{feed}V_{ac}\Omega\sin(\Omega t) \quad (19)$$

also contributes to the output current $i_{out} = i_{motion} + i_{feed}$. However, when the driving frequency Ω is near the resonant frequency Ω_r, we can often assume that $|i_{feed}| \ll |i_{motion}|$; i_{out} is thus approximated with i_{motion} in this case.

The motional impedance Z_m is generally given by

$$Z_m = \frac{v_{ac}}{i_{out}} = \frac{v_{ac}}{i_{motion} + i_{feed}}, \quad (20)$$

but by using the above the assumption of $\Omega \approx \Omega_r$, $|i_{feed}| \ll |i_{motion}|$, and we can simplify the expression for the minimum motional impedance by using Eqn. (18):

$$Z_{m,min} \approx \frac{V_{ac}}{I_{motion,max}} \approx \frac{1}{2}\left[\frac{d^2}{V_{bias}\varepsilon_0 h l L_1}\right]^2 \frac{m\omega_1}{Q}. \quad (21)$$

$Z_{m,min}$ is often called the motional resistance R_m since Z_m gets to the minimum when the motional inductance L_m and the motional capacitance C_m cancel each other out. Comparing R_m from the linear models to Eqn. (21), we find that nonlinear $Z_{m,min}$ is not affected by nonlinearities and is similar to R_m from the linear models [3]–[5].

Fig. 4 shows estimated Z_m: for the estimation, we plug the solutions of Eqn. (11) into Eqn. (10) and then into Eqn. (17) to estimate i_{motion}; we use Eqn. (19) to estimate i_{feed}; we then estimate Z_m by using Eqn. (20). The result shows that $Z_{m,min}$ is not affected by nonlinearities (or the level of V_{ac}), as Eqn. (21) predicts. This result implies that the assumption of $|i_{feed}| \ll |i_{motion}|$ is valid. Table I summarizes the parameters for the estimation.

III. EXPERIMENTAL SETUPS

A. Conventional open-loop measurement setup

A conventional way to measure Z_m is by using an impedance analyzer (Fig 5(a)). An impedance analyzer provides input voltage signal v_{ac} and measures the output current signal i_{out}. By measuring this open-loop I-V characteristic, the analyzer can measure the Z_m in a given frequency range. This impedance analyzer method is good for resonators that do not have nonlinearities, but not for resonators with nonlinearities.

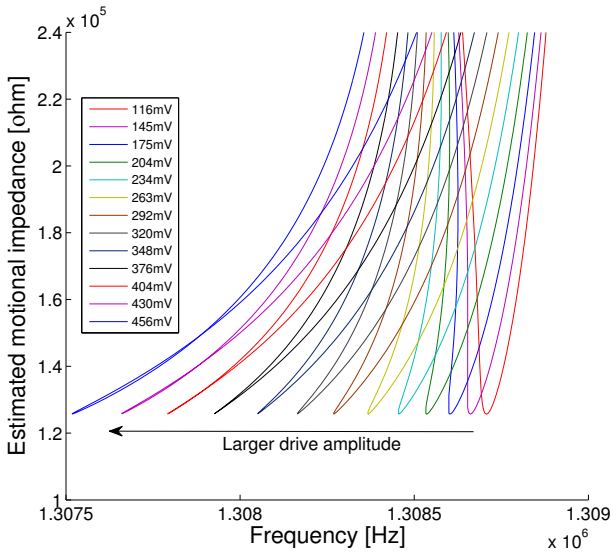


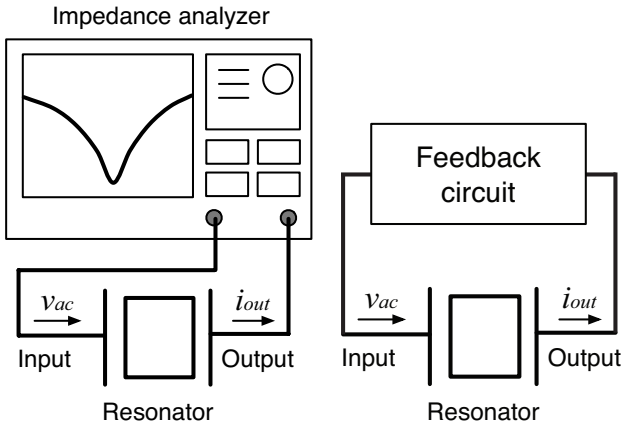
Fig. 4. Estimated $|Z_m|$ at various V_{ac} levels, shown in the legend, for a given range of the driving frequency Ω. For each V_{ac}, we use Eqn. (20) to estimate $|Z_m|$.

To understand this limitation, we must check the frequency-to-amplitude relation of a resonator in the nonlinear regime.

Nonlinearities in a resonator result in the frequency-to-amplitude relation exhibiting V_{ac}-dependency (Fig. 2) [3], [4], [6], [7], [12]–[14]. When V_{ac} is larger than a critical value ($\approx$ 145 mV in Fig. 2), the frequency-to-amplitude relation becomes indeterministic for some driving frequencies; hence, the amplitude of a resonator can be uncertain for some driving frequencies. This uncertainty leads to unstable states in the impedance analyzer measurement. In the actual measurement, the part of the Z_m response curve that corresponds to the unstable states becomes unobservable, though stable branches in the response curve can be measured by alternating the frequency sweep direction [4], [7], [12]–[15]. To summarize, an impedance analyzer cannot measure the full characteristic of Z_m when a resonator is in the nonlinear regime.

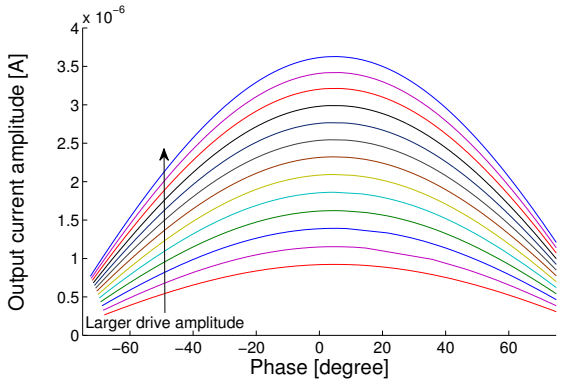
B. Novel closed-loop measurement setup

We propose a new Z_m measurement method by using a feedback circuit (Fig 5(b)). A feedback circuit provides gain for continuous operation of a resonator. By comparing amplitudes of an input voltage signal v_{ac} and an output current signal i_{out}, we can perform closed-loop I-V measurements. The major difference between using a feedback circuit and using an impedance analyzer is that a feedback circuit sets the phase rather than the driving frequency Ω; in other words, in this closed-loop method, phase is an independent variable. Unlike the frequency-to-other relations, the phase-to-amplitude and the phase-to-frequency relations are always deterministic in the closed-loop measurement setup, regardless of nonlinearities. (Fig. 6) [10], [16]. Since the amplitude and frequency are deterministic for any given phase, instability cannot occur and

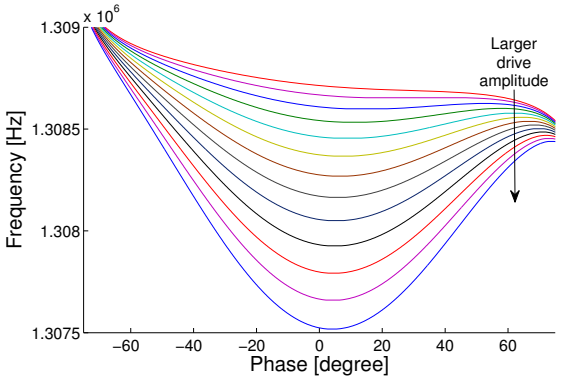


(a) Open-loop setup (b) Closed-loop setup

Fig. 5. Motional impedance measurement setups.



(a) Phase-to-amplitude relation



(b) Phase-to-frequency relation

Fig. 6. Phase-to-other relations at the V_{ac} levels that are used for the estimation in Fig. 2. Eqns. (17) and (19) are used to estimate the output current $i_{out} = i_{motion} + i_{feed}$; the magnitude of i_{out} is then used to estimate the output current amplitude; the phase difference between i_{out} and v_{ac} is used to estimate the phase. The peak amplitude occurs at non-zero phase because of the feedthrough current i_{feed}.

the measurement is not limited by nonlinearities unlike in the open-loop measurement setup [10], [16].

If a feedback circuit has a fixed phase, however, we can

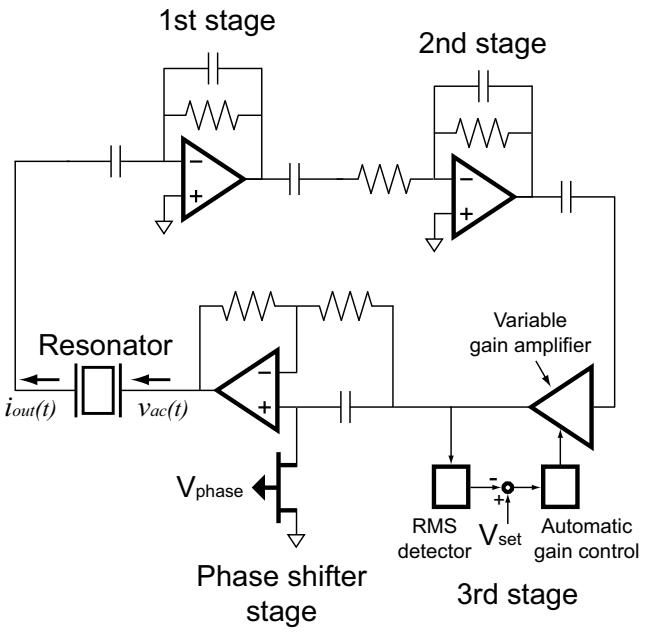


Fig. 7. Circuit diagram of a variable-phase closed-loop feedback circuit

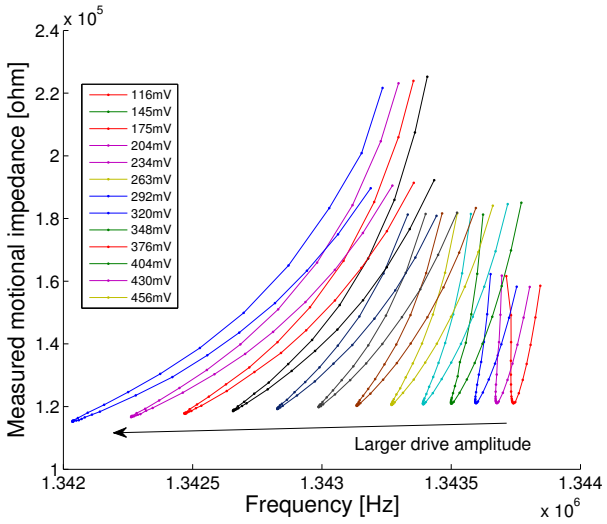


Fig. 8. Measured $|Z_m|$ at various V_{ac} levels, shown in the legend. For each V_{ac}, we used more than 25 V_{phase} values to measure nonlinear Z_m characteristics.

measure Z_m only at one point, which corresponds to the fixed phase, but not at many other points. Therefore, the feedback circuit must perform phase sweep, just as an impedance analyzer does frequency sweep, to measure many operating points. For this reason, we need a variable-phase feedback circuit. To implement the variable-phase feedback circuit (Fig. 7), we add a voltage-controlled phase shifter to an ordinary fixed-phase feedback circuit. The first and the second stages of the feedback circuit act as a trans-impedance amplifier (TIA), which converts current to voltage. The third stage employs an automatic level control, a variable gain amplifier, and a RMS detector to maintain V_{ac} at a desired level, which is externally set by V_{set}. To implement a voltage-controlled phase shifter, we use an all-pass filter with a voltage-controlled resistor, implemented by a JFET: the resistance of the JFET is externally set by V_{phase}. Then, this circuit is connected to a resonator to form a closed-loop oscillator.

IV. Experimental Procedure and Results

We used an SA-DETF type resonator for the experiment (Fig. 1(a)). This resonator was fabricated with the process detailed in [17]. The resonant frequency is about 1.3 MHz, and the Q is about 8710. This resonator experiences first-order electrical softening nonlinearity [13]. We connected a counter to the output of the TIA to monitor the resonant frequency. We also connected an oscilloscope to the output of the TIA and the input to the resonator to measure the amplitude of input and output signals (Fig. 7).

In order to measure Z_m at many phase values and V_{ac} levels, we first set the phase of the circuit; we then estimated Z_m by measuring the input and output amplitudes at many V_{ac} levels: although we measured the voltage amplitude at the output of the TIA, we can estimate I_{out} by using the known gain of the TIA. Repeating this for many phase values, we could measure Z_m at many operating points.

Fig. 8 shows the experimental result. We set more than 25 phase values and 13 V_{ac} levels. The $Z_{m,min}$ ranged from 121 kΩ to 115 kΩ. The curves show that our new method could measure nonlinear Z_m characteristics that an impedance analyzer cannot measure. Since an impedance analyzer cannot measure these characteristics, this is a remarkable result.

Our measured result for the nonlinear Z_m characteristics match well the estimation based on the model we developed. Our model over-estimates Z_m by less than 10% for the worst case. We speculate that the difference stems from uncertainties in dimensional and material parameters for the estimation. A more noticeable difference between the estimation and the measured data is the change of $Z_{m,min}$ with the change of V_{ac}. In the estimation, the $Z_{m,min}$ value is not a function of V_{ac}, and this is expected by the model in Eqn. (21): since we used constant Q for the estimation, $Z_{m,min}$ must be constant. In the measured data, however, $Z_{m,min}$ decreases by 5% as we increased V_{ac} for a given V_{ac} range. Since $Z_{m,min}$ is inversely proportional to Q, Eqn. (21), this result may imply that the Q of the resonator increases by 5% when V_{ac} increases for the given range. However, it may also be an experimental artifact or stem from assumptions we make when we derive the Z_m model in section II, such as a weakly nonlinear system. Therefore, further research is necessary to find the source of the $Z_{m,min}$ change.

V. Conclusion

We developed a nonlinear Z_m model and a new measurement setup, with which we can estimate and measure nonlinear Z_m characteristics of an electrostatic resonator. The Z_m model we developed can precisely estimate the Z_m of a resonator that

operates in the nonlinear regime with errors of less than 10% in the worst case. Although the model was developed for a specific type of resonators, the result of this study can be generally applied to other types of resonators because an ordinary differential equation similar to Eqn. (9) can be obtained after modal analysis is applied to the model for other resonators. In addition, the new closed-loop measurement setup can be also utilized to measure the nonlinear Z_m characteristics of other types of resonators. Therefore, this study is of broad relevance to the resonator community.

ACKNOWLEDGMENT

This work was supported by DARPA HERMIT (ONR-N66001-03-1-8942), the Bosch Palo Alto Research and Technology Center, the Samsung Scholarship, the National Science Foundation Graduate Fellowship, the Stanford Graduate Fellowship, and the National Users Network facilities funded by the National Science Foundation under award ECS-9731294. The authors would especially like to thank Clark Nguyen, Amit Lal, and Andrei Shkel for their DARPA HERMIT program management.

REFERENCES

[1] C. T.-C. Nguyen, "MEMS technology for timing and frequency control," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 54, no. 2, pp. 251–270, 2007.

[2] G. Ho, K. Sundaresan, S. Pourkamali, and F. Ayazi, "Low-motional-impedance highly-tunable I^2 resonators for temperature-compensated reference oscillators," in *Proc. IEEE The 18th International Conference on Micro Electro Mechanical Systems*, Miami beach, FL, USA, Jan.30–Feb.3, 2005, pp. 116– 120.

[3] T. Mattila, J. Kiihamäki, T. Lamminmäki, and O. Jaakkola, "A 12 MHz micromechanical bulk acoustic mode oscillator," *Sens. Actuators, A*, vol. 101, no. 1-2, pp. 1–9, Jan. 2002.

[4] S. Lee and C. T.-C. Nguyen, "Influence of automatic level control on micromechanical resonator oscillator phase noise," in *Proc. 2003 IEEE International Frequency Control Symposium*, Tampa, FL, USA, May4–8, 2003, pp. 341 – 349.

[5] M. Agarwal, H. Mehta, R. N. Candler, S. Chandorkar, B. Kim, M. A. Hopcroft, R. Melamud, G. Bahl, G. Yama, T. W. Kenny, and B. Murmann, "Scaling of amplitude-frequency-dependence nonlinearities in electrostatically transduced microresonators," *J. Appl. Phys.*, vol. 102, no. 7, p. 074903, Jan. 2007.

[6] M. Younis and A. Nayfeh, "A study of the nonlinear response of a resonant microbeam to an electric actuation," *Nonlinear Dyn.*, vol. 31, no. 1, pp. 91–117, Jan. 2003.

[7] V. Kaajakari, T. Mattila, A. Oja, and H. Seppä, "Nonlinear limits for single-crystal silicon microresonators," *J. Microelectromech. Syst.*, vol. 13, no. 5, pp. 715–724, 2004.

[8] L. Meirovitch, *Elements of vibration analysis*. New York: McGraw-Hill, 1986, vol. 2.

[9] H. A. C. Tilmans, "Equivalent circuit representation of electromechanical transducers: II. distributed-parameter systems," *J. Micromech. Microeng.*, vol. 7, no. 4, pp. 285–309, 1997.

[10] R. Mestrom, R. Fey, and H. Nijmeijer, "Phase feedback for nonlinear MEM resonators in oscillator circuits," *Mechatronics, IEEE/ASME Transactions on*, vol. 14, no. 4, pp. 423 – 433, Aug. 2009.

[11] M. A. Hopcroft, "Temperature-stabilized silicon resonators for frequency references," Ph.D. dissertation, Stanford University, Stanford, CA, USA, Sep. 2007. [Online]. Available: http://micromachine.stanford.edu/ hopcroft/

[12] Y.-W. Lin, S. Lee, S.-S. Li, Y. Xie, Z. Ren, and C. T.-C. Nguyen, "Series-resonant VHF micromechanical resonator reference oscillators," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2477– 2491, 2004.

[13] M. Agarwal, R. N. Candler, B. Kim, M. Hopcroft, R. Melamud, C. M. Jha, T. W. Kenny, and B. Murmann, "Optimal drive condition for nonlinearity reduction in electrostatic microresonators," *Appl. Phys. Lett.*, vol. 89, no. 21, p. 214105, Jan. 2006.

[14] L. Shao, M. Palaniapan, W. Tan, and L. Khine, "Nonlinearity in micromechanical free-free beam resonators: modeling and experimental verification," *J. Micromech. Microeng.*, vol. 18, no. 2, p. 025017, Jan. 2008.

[15] V. Kaajakari, T. Mattila, A. Lipsanen, and A. Oja, "Nonlinear mechanical effects in silicon longitudinal mode beam resonators," *Sens. Actuators, A*, vol. 120, no. 1, pp. 64–70, Jan. 2005.

[16] B. Yurke, D. S. Greywall, A. Pargellis, and P. Busch, "Theory of amplifier-noise evasion in an oscillator employing a nonlinear resonator," *Phys. Rev. A*, vol. 51, no. 5, pp. 4211–4229, May 1995.

[17] R. N. Candler, M. Hopcroft, B. Kim, W.-T. Park, R. Melamud, M. Agarwal, G. Yama, A. Partridge, M. Lutz, and T. W. Kenny, "Long-term and accelerated life testing of a novel single-wafer vacuum encapsulation for MEMS resonators," *J. Microelectromech. Syst.*, vol. 15, no. 6, pp. 1446 – 1456, Dec. 2006.

Time transfer by laser link - T2L2: Current status and future experiments

E. Samain[1], P. Exertier[1], Ph. Guillemot[2], Ph. Laurent[3], F. Pierron[1], D. Rovera[3], J.-M. Torre[1], M. Abgrall[3], J. Achkar[3], D. Albanese[1], C. Courde[1], K. Djeroud[3], M. Laaz Bourez[1], S. Leon[4], H. Mariey[1], G. Martinot-Lagarde[1], JL. Oneto[1], J. Paris[1], M. Pierron[1], H. Viot[1]
[1]OCA – Observatoire de la Côte d'Azur, Caussol, France
[2]CNES –French Space Agency, Toulouse, France
[3]OP –Observatoire de Paris, France
[4]CNES –French Space Agency, Paris, France

Email: etienne.samain@oca.eu

Abstract— **T2L2 (Time Transfer by Laser Link), developed by both CNES and OCA permits the synchronization of remote ultra stable clocks over intercontinental distances. The principle is derived from laser telemetry technology with dedicated space equipment deigned to record arrival time of laser pulses at the satellite. Using laser pulses instead of radio frequency signals, T2L2 permits to realize some links between distant clocks with a time stability of a few picoseconds and accuracy better than 100 ps. The T2L2 space instrument is in operation onboard the satellite Jason 2 since June 2008. Several campaigns were done to demonstrate both the ultimate time accuracy and time stability capabilities. It includes some experiments implemented in co-location to directly compare T2L2 time transfer residuals with the direct link between stations, and some ground to ground time transfer between ultra stable clocks. Important works have been done, between OCA and OP, to accurately compare T2L2 with microwave time transfer GPS and TWSTFT. These comparisons are based on laser station calibrations with a dedicated T2L2 calibration station designed to accurately set the optical reference of the laser station within the PPS reference of the microwave systems. Other experiments are also planned in the future: 3D localization with the lunar space vehicle LRO, T2L2 coverage extension over the Pacific Ocean (Tahiti), DORIS comparison and a third international campaign.**

I. INTRODUCTION

T2L2 is a time transfer technique based on laser ranging network at ground and a dedicated space segment. The project, developed by CNES (Centre National d'Etudes Spatiales) and OCA (Observatoire de la Côte d'Azur), has been launched in 2008 on the Jason-2 satellite for a 5 years mission [1], [2], [3].

T2L2 relies on the propagation of laser pulses between the clocks to be synchronized. It is an evolution of current radiofrequency time transfer systems profiting from advantages of the optical domain such as higher modulation bandwidth and insensitivity to ionosphere, and also from advantages of the two way technique on a mono-carrier scheme. All these elements permit to enhance the performance of time transfer by one or two order of magnitudes compared to existing microwave techniques such as GPS and Two-Way Satellite Time and Frequency Transfer (TWSTFT).

It provides the capability to compare today's most stable frequency standards with unprecedented stability and accuracy [4]. Ultimate expected T2L2 performances are in the 100 ps range for accuracy, with a time stability of about 1 ps over 1,000 s and 10 ps over one day [5].

II. PRINCIPLE

The T2L2 elementary block is a space to ground time transfer between the ground clock linked to the laser station and the space clock. The ground to ground time transfer between several remote clocks at ground is obtained through such individual space to ground time transfers. It can be obtained in a common view mode, when the distance between the laser stations is smaller than 5000 km, or in a non-common view mode when the distance is larger.

The laser station includes a telescope, a pulsed laser, some photo detection devices and two event timers connected to the ground clock, one to time tag the start event, the other for the return. The energy by pulse is between 400 µJ to 1 J and the pulse width is in the range of 20 ps up to 200 ps. The laser rate is between 5 Hz to 2 kHz. The laser stations range the satellite every 2 hours, 5 to 7 times a day when the weather conditions are favorable.

The space instrument (Fig. 1) includes a photo detector [6] and an event timer [7], [8] linked to the space clock and a laser ranging array (provided by the JPL) used to reflect the laser pulse toward the laser station. The space clock is an ultra-

stable oscillator (USO) fundamentally used by the Doris equipment.

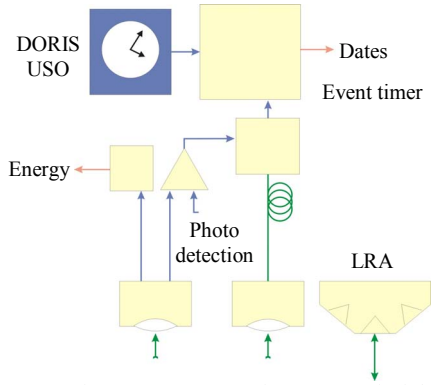


Figure 1. T2L2 Space instrument principle

For a given laser pulse emitted by the laser station, one get a start date and a return date at ground, and also an arrival date at the satellite. From these 3 dates (which are called a triplet) we can extract the time delay between the ground clock and the space clock. T2L2 relies on the laser ranging network which includes 40 international laser stations. Among them, 16 stations provide the full rate data needed for the triplets extraction.

The first step in the T2L2 data reduction is the determination of the ground to space time transfer for every passes acquired. Fig. 2 is an example of what is got for a pass obtained with the MeO Laser Station linked to an H-Maser at OCA.

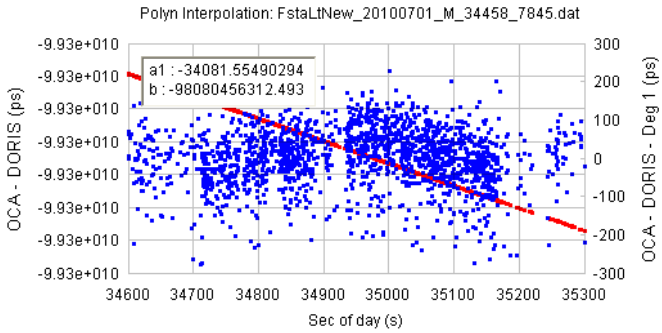


Figure 2. OCA – DORIS time transfer. The red line (left scale) is the direct time transfer. The global slope is coming from the large frequency offset between DORIS and OCA. The blue dots (right scale) are the residuals.

The corresponding time stability is illustrated in Fig. 3.

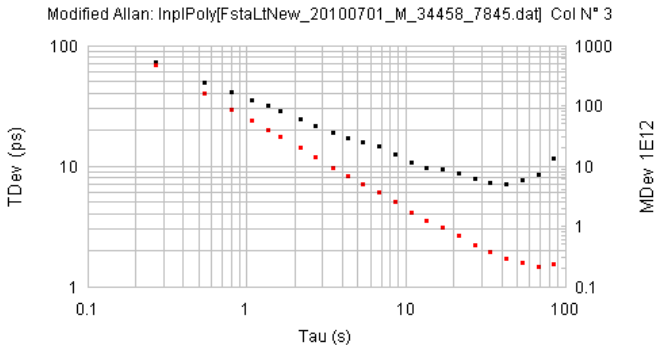


Figure 3. OCA – DORIS Time deviation (black, left scale) and Modified Allan deviation (red, right scale). Performances are in accordance with the DORIS USO.

Accumulation of several consecutive passes over the same station permits to give a mid-term time transfer comparison (Fig. 4). These mid-term monitoring will be used in the next reduction evolution to make time transfer in non common view configuration.

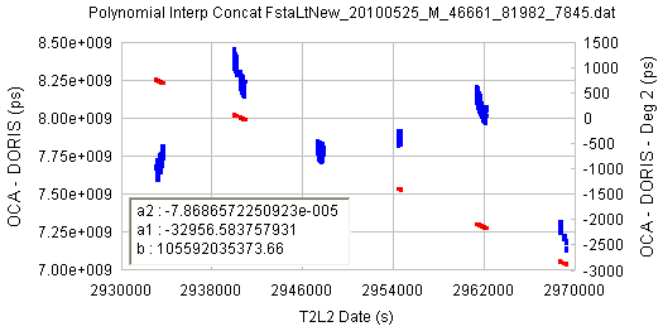


Figure 4. OCA – DORIS comparison over several consecutive passes (Direct comparison, red, left scale; residuals, blue, right scale). Up to 7 passes, giving an overall time comparison of 12 hours, can be acquired consecutively.

III. ACCURATE GROUND TO GROUND TIME TRANSFER

T2L2 has been designed to realize ground to ground time transfer with accuracy better than 100 ps. If one can neglect the noise coming from the individual ground to space time transfer, this accuracy requires the measurement of the laser pulse start times with accuracy in the range of 50 ps.

Usual laser stations are only designed to measure time difference between a start and a stop with an accuracy of the start pulse of typically 100 ns. These 50 ps can only be obtained with a special time calibration based on simultaneous measurements between the usual chronometry of the laser station and a dedicated calibration station. It permits to measure the delay between the optical pulse at the cross axis of the telescope and the electrical reference coming from the time and frequency laboratory. The reference point at the laser station level is the cross axes of the telescope which is also the space reference for laser ranging. The reference at the time and frequency laboratory is a given output of the PPS (Pulse Per Second) distribution unit.

The calibration station gather inside an unique equipment all the metrology required to perform measurements. The station includes a SigmaTime STX301Event Timer and an optical module to grab laser pulses from the laser station.

The event timer STX301 was developed in the T2L2 framework with 3 separate targets:

- to make the T2L2 calibration,
- to upgrade, if necessary, laser stations with no proper chronometry,
- to have a versatile high precision laboratory instrument.

The instrument has an internal frequency synthesis that can be slaved on an external 5, 10 or 100 MHz clock reference, two or four independent channels (electrical and optical: DC - 20 GHz), and a high stability pulse generator. All these metrology instruments are driven by an embedded Windows 7 computer able to control the hardware and to give to the user a friendly Human Machine Interface permitting to

measure both absolute start time and time interval, to sample the frequency, and to determine the shape of a given synchronous signal.

Figure 5. STX301-002R SigmaTime sub picosecond Event timer.

The global characteristics of the timer are given in table 1.

Parameters	Values
Number of independent channels	2 to 4
Time Stability @ 1000s	< 20 fs
Linearity	0.3 ps rms.
Thermal Sensitivity	< 200 fs/°C
Repeatability error Synchronous	< 600 fs rms
Repeatability error Random	700 fs rms
Dead time	130 ns
High speed Acquisition	500 kHz
Continuous rate	35 kHz
Input bandwidth	DC - 20 GHz

Table 1. Sigma Time STX301 main characteristics

Fig. 6 is a typical time deviation of the timer computed from the modified allan variance.

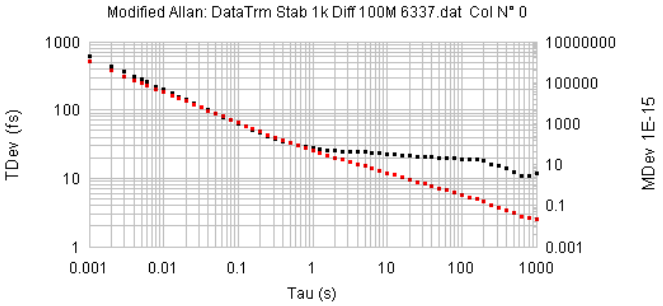


Figure 6. Time deviation (black, left scale) and Modified Allan deviation (red, right scale)of the SigmaTime STX301 event timer obtained from electrical pulses generated with an independent frequency synthesis slaved on a common clock

The optical module is made with a collimation optic connected to a 50 m mono mode optical fiber (532 nm). The delay inside the fiber is currently considered as a constant. In the next future a special instrumentation will permit to measure the variations that could be generated by some temperature changes or constraint variations. The optical module is placed in front of the telescope at a precise distant from the cross axes.

Time equation that permits to accurately time-tag laser pulses is given by:

$$\delta T = \delta_{cal} + \delta_{prg} = \delta_{cal} + \delta_{PPS} - (\delta_{ocx} + \delta_{ocf} + \delta_f + \delta_{det}),$$

where:

δ_{cal} is the difference between absolute measurement (calibration) and station measurement and δ_{prg} the global propagation between cross axes and the PPS unit (Fig. 7).

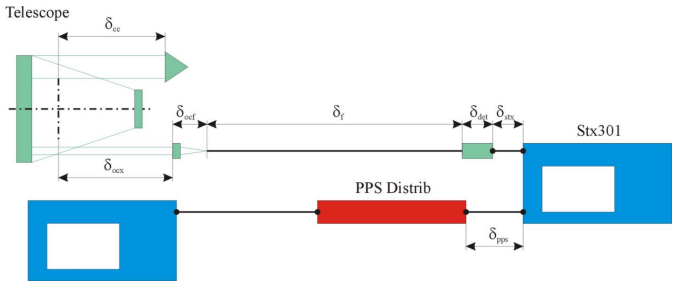


Figure 7. Definition of the propagation term through the telescope to the PPS distribution unit

The term δ_{cal} is determined in two steps:

- PPS Synchronization of the SigmaTime STX301 event timer
- Simultaneous acquisition of an ensemble of laser pulses from the laser station and the calibration station

The PPS synchronization is obtained from a Scan of the PPS signal by the event timer that permits to determine the reference threshold deduced from the inflexion point and the synchronization of the timer with this reference threshold.

The term δ_{prg} depends on three kinds of parameters:

- δ_{PPS} ; δ_f coming from propagation in cables and measured by the calibration station
- δ_{ocx} ; δ_{ocf} coming from propagation in free space and determined form the geometrical distance
- δ_{det} coming from propagation inside the detector (optical-electrical) and deduced form a propagation model (currently studied) [9].

The table 2 is an example of what we get for the MeO station

δ	Label/Ref	Value (ps)	Date
δ_{cc}	CC_{Lune}	12393	2010-07-01
δ_{PPS}	$T2L2CalC_2$	9408	2006-08-06
δ_{ocx}	Ref_{Axe}	12104	2010-07-01
δ_{ocf}	$T2L2MC_2$	58	2010-07-01
δ_f	$T2L2CalF_1$	248300	2006-08-06
δ_{stx}	STX301-001	0	2010-07-01
δ_{det}	NewFocus1454	660	2010-07-01
δ_{cal}	GioveB 100522	628721	2010-05-22
δ_T		377007	2010-07-01

Table 2. MeO Calibration

IV. T2L2 CO-LOCALISATION AT OCA

OCA has two laser stations (fixed: MeO and mobile: FTLRS) that can be used together in order to realize a time transfer in a co-location configuration (Fig. 8). Being both connected to the same H-Maser, this configuration permits to directly compare the time transfer deduced from the calibration process and time transfer measured by T2L2.

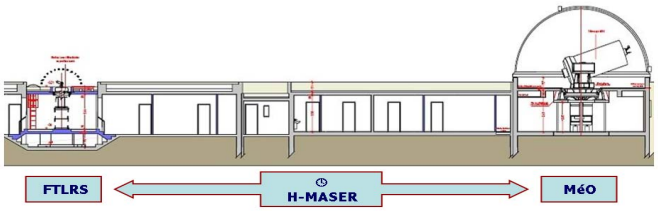


Figure 8. FTLRS (left) and MeO (right) connected to same HMaser reference in order to realize a co-localisation time transfer

13 common passes have been obtained in 2009 during a dedicated campaign (Fig. 9).

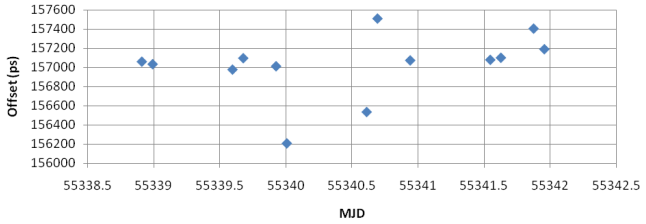


Figure 9. MeO – FTLRS T2L2 Time difference.

The time offset deduced from T2L2 Measurement is:

$$\delta_{TFtlrs} - \delta_{TMeo} = 157\ 019 \text{ ps} ; \sigma = 330 \text{ ps}$$

The time offset deduced from the T2L2 Calibration is:

$$\delta_{TFtlrs} - \delta_{TMeo} = 157\ 030 \text{ ps} ; s = 26 \text{ ps}$$

This gives a very good agreement within the given experimental standard deviation.

V. SECOND INTERNATIONAL CAMPAIGN

The T2L2 international campaigns aim to achieve the best possible time stability measurement between ultra stable clock and the best comparisons between microwave time transfer techniques [10],[11]. A first campaign was done in 2009 involving 4 laser stations. In 2010 one realizes a new 5 months campaign with 9 laboratories (Table 3).

Site	Clock	Time Transfer
Grasse-FRA	Fountain ; HM	GPS TWSTFT Eu
Paris-FRA	Fountain ; HM	GPS TWSTFT Eu
Borowiec-POL	HM	GPS TWSTFT Eu
Koganei-JPN	Fountain	GPS TWSTFT As
Simosato-JPN	Cs/Rb	GPS
Zimmerwald-CHE	Qx/GPS	GPS
Herstmonceux-GBR	HM	GPS
Matera-ITA	Cs ; HM	GPS
Wettzell-DEU	HM	GPS

Table 3. Laboratories global configuration

1155 passes was obtained during the campaign; 650 in common view configuration.

So far, only the link Grasse-Paris (Fig. 10) has been analyzed; the others will be reviewed in the next future.

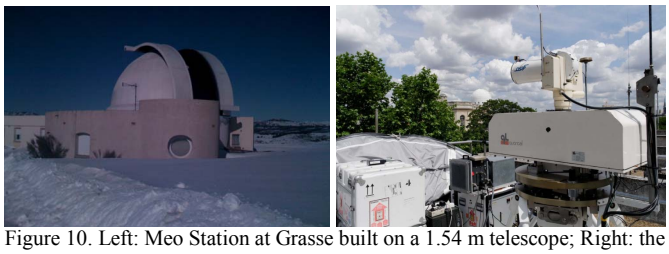


Figure 10. Left: Meo Station at Grasse built on a 1.54 m telescope; Right: the mobile FTLRS station at Observatoire de Paris.

Table 4 is a summary of the instrumental configuration of both laboratories for this Grasse-Paris link.

Location	Station	clocks	MW Time transfer
Grasse OCA	MeO	Fountain (OP) H-Maser	GPS TWSTFT
Paris OP	FTLRS (OCA)	Fountain H-Maser	GPS TWSTFT

Table 4. Clocks and Microwave time transfer systems

The OCA laser station FTLRS was installed on a dedicated platform at OP. Some special authorizations were obtained in order to be able to range with a laser in Paris. The OP mobile atomic fountain (FOM) was installed at OCA during the same period. At OP, FTLRS, the atomic fountain and the Microwave time transfer systems (GPS and TWSTFT) were connected to the same H-Maser. At OCA, MeO, the mobile atomic fountain and the time transfer systems were also connected to a common H-Maser. Both OP and OCA have been calibrated with the same T2L2 calibration station as described in chapter 3. 56 passes in 114 days was recorded. From each pass acquired, an equivalent point representing the average value of the time difference and the average value of the date is computed.

The first result extracted from these comparisons is the T2L2 time difference between H-Masers (Fig. 11, Fig. 12).

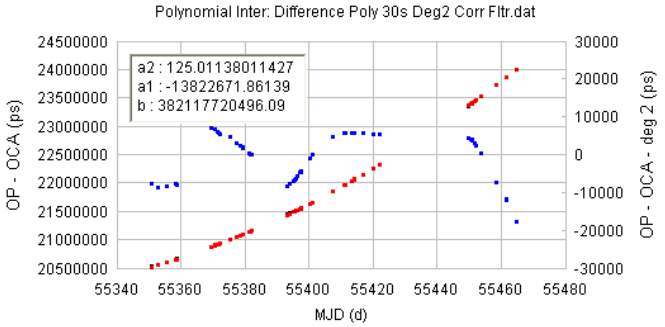


Figure 11. Direct T2L2 time differences between OP and OCA H-Masers (red, left) and difference – order 2 polynomial (blue, right).

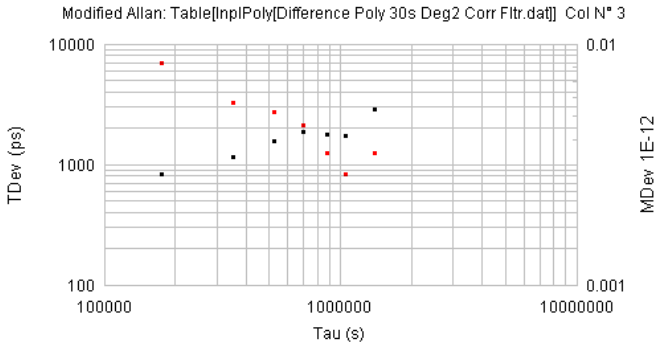


Figure 12. Time deviation (black, left scale) and Modified Allan deviation (red, right scale) of the T2L2 H-Maser differences

With a modified Allan deviation in the range of 3x10-15 over 10 days, one concludes that the noise introduced by T2L2 over several days is significantly below the noise coming from the H-Masers themselves.

The second important result is illustrated in figure 13. The graph shows both the time transfer comparison between T2L2 – GPS (carrier phase based) – TWSTFT (code phase) and the differences between Atomic fountains. The GPS analysis is based on the PPP NRCan algorithm (carrier phase technique). This is a preliminary result in which the absolute calibration hasn't been taken into account: an offset between each curve was introduced to facilitate the graph reading. Regardless of these absolute aspects, the global drift on the Microwave-T2L2 time comparisons are better than 2 ns over 2 months which is in accordance with the classical long term stability of microwave time transfers.

The phase of the atomic fountain comparison is computed from the frequency information between the fountain interrogation and the H-Maser frequency reference. For some technical reasons, the atomic fountains weren't in operation continuously during the whole campaign. This has probably introduced a significant midterm noise: pink and red plots of fig. 13. A more detailed analysis taking into account the real operation period of the fountains, will be made in the future.

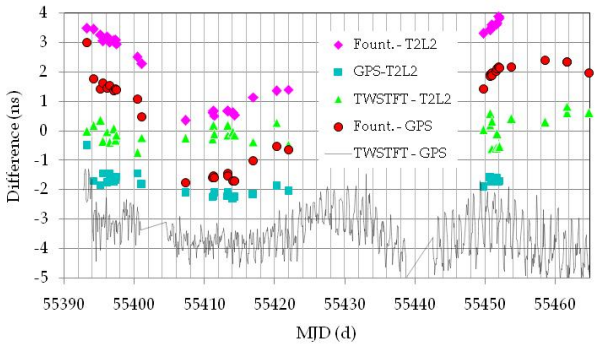


Figure 13. Microwave and T2L2 time transfer comparison and atomic fountain differences. A time offset between each plot is voluntary introduced to facilitate the reading of the graph.

VI. 3D LOCALIZATION WITH LUNAR RECONNAISSANCE ORBITER (LRO)

The primary objective of LRO (NASA project) is to conduct investigations in order to prepare for future lunar exploration. The orbiter has a laser altimeter called LOLA, fundamentally made to measure the relief of the moon surface. This altimeter has also been designed to be used from the earth with a 2 cm aperture telescope oriented toward the earth direction. This system permits to range the orbiter from the laser station network at ground. The range is deduced from the computation of the time difference between start time recorded by the laser stations and the arrival time recorded by altimeter. If one can range simultaneously LOLA with at least 2 laser stations one can resolve the 3D geometry from the differential measurement onboard LOLA (Fig. 14). Because the geometry is very sharp (distance between LRO >> distance between laser stations) the time synchronization between laser stations has to be very precise. Because T2L2 is based on the same laser ranging technologies which are also used for LOLA, T2L2 is perfectly well suited to perform this synchronization.

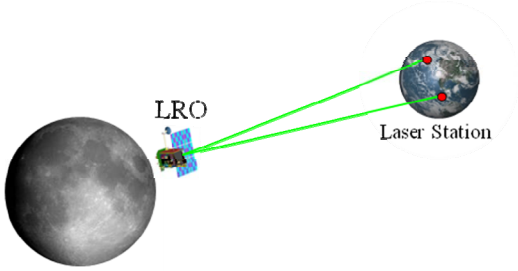


Figure 14. 3D localization of LRO with several laser stations at ground synchronized by T2L2

Several simultaneous observations in common view have already been acquired (Fig.15). So far, data hasn't been yet analyzed with T2L2: everything remains to be done on the subject.

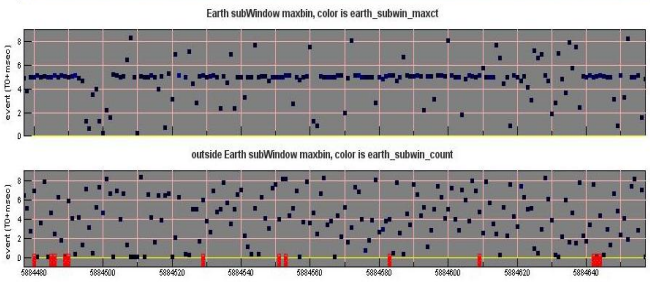


Figure 15. Simultaneous acquisition from Zimmerwal (CH) and MeO (FR)

VII. TAHITI CAMPAIGN

A new dedicated campaign was started in April 2011 on the OGT (Observatoire Geodésique de Tahiti) site (Fig. 16) where a USA Moblass laser station is installed. The campaign implies the installation of the mobile laser Station FTLRS, H-Maser, a devoted T2L2 GPS receiver and the T2L2 calibration station.

The objective is four fold:

- OneWay laser ranging on LRO between McDonald (USA) and Tahiti
- Remote control of the onboard DORIS oscillator over a region currently not observed
- T2L2 - DORIS Inter comparison
- T2L2 FTLRS -Moblass time transfer co-location

Figure 16. Mobile laser Station FTLRS at Tahiti

VIII. PROSPECTIVE-CONCLUSIONS

Four important developments are envisioned in the next T2L2 schedule:

- Ground to ground time transfer in a non common view configuration based on an onboard oscillator model fitted from an ensemble of referenced ultra stable clock at ground.
- International T2L2 calibration campaign. It includes the laser station calibration and also a relative GPS calibration from a mobile GPS Receiver
- A third international campaign. As compared to the standard laser station network, the campaign could be done with two more lab: OP with FTLRS and PTB with TROS (Chinese mobile laser station)
- DORIS comparison between ground to space time transfer deduced from the DORIS frequency measurement and T2L2

ACKNOWLEDGMENT

The authors wish to thanks ILRS and laser stations that have participate to the T2L2 project. They wish also to thanks the French Civil Aviation for laser security at Paris.

---

[1] P. Fridelance, E. Samain and C. Veillet, "*T2L2 - Time transfer by Laser link : A New Optical Time Transfer Generation*", Experimental Astronomy Vol.7 Num.3 Sept.97

[2] Ph. Guillemot, E. Samain, P. Vrancken, P. Exertier, S. Leon, "*Time transfer by laser link - T2L2 : An Opportunity to Calibrate RF Links*", Proceedings of the PTTI 2008, pp 95-106

[3] P. Exertier, E. Samain, P. Bonnefond, and P. Guillemot, 2010, "*Status of the T2L2/Jason2 Experiment,*" Journal Advances in Space Research, Doris Special Issue, 46, 1559-1565, doi:10.10.16/j.asr.2010.06.028

[4] P. Vrancken, "*Characterization of T2L2 (Time Transfer by Laser Link) on the Jason 2 Ocean Altimetry Satellite and Micrometric Laser Ranging*", Thesis, Université de Nice - Sophia Antipolis, 2008.

[5] Ph. Guillemot, P. Exertier, E. Samain, F. Pierron, J.-M. Torre, S. Leon, "*Time Transfer by Laser Link - T2L2 : Results of the First Year of Operation*", Proceedings of the PTTI 2009, pp 67-79

[6] E. Samain, "*Timing of optical pulses by photodiode in Geiger mode*", Applied Optics, 37, No 3, 1998, pp 502-506, 1998.

[7] E. Samain, "*An Ultra Stable Event Timer*", Proceedings of the 13th International Workshop on laser ranging instrum., 2002

[8] E. Samain, P. Fridelance, Ph. Guillemot, "*An Ultra Stable Event Timer Designed for T2L2*", Proceedings of the EFTF/IFCS 2010

[9] I. Prochazka, J. Blazej, J. Kodet, "Measurement of the optical to electrical detection delay in the detector for ground-to-space laser time transfer, Metrologia 48, 2011, L13-L16

[10] Ph. Guillemot, P. Exertier, E. Samain, F. Pierron, Ph. Laurent, M. Abgrall, J. Achkar, D. Rovera, K. Djerroud, S. Leon, "Time Transfer by laser Link – T2L2: First Results of the 2010 Campaign", Proceedings of the PTTI 2010, pp 397-412.

[11] J. Delporte, F. Mercier, D. Laurichesse, O. Galy, "*GPS carrier phase time transfer using single-difference integer ambiguities*", International Journal of Navigation and Observation, Hindawi Publishing Corporation, Volume 2008, Article ID 273785

Construction Progress in the Photon Counting Detector for the European Laser Timing Experiment

Jan Kodet, Ivan Prochazka, Josef Blazej
Czech Technical University in Prague
Prague 1, Czech Republic
kodet@fjfi.cvut.cz

Jan Brinek
CSRC
Karlova 7, Brno, Czech Republic

***Abstract*— We are presenting a progress in a construction and indoor tests of the photon counting detector for the European Laser Timing (ELT) experiment. ELT is an optical link prepared in the frame of the ESA mission "Atomic Clock Ensemble in Space" (ACES). The objective of this laser time transfer is the synchronization of the ground based clocks and the clock on board the space station with precision of the order of units of picoseconds and the accuracy of 50 ps. The requirements put on the detector package are quite high – temperature stability of the delay better than 20 ps peak to peak within one satellite orbit, operation within a broad temperature, absolute calibration of the photon to electrical signal delay with precision 25 ps and others.**

I. INTRODUCTION

The laser time transfer link is under construction for the European Space Agency (ESA) for its application in the experiment Atomic Clock Ensemble in Space (ACES). The device is expected to be launched toward the International Space Station (ISS) in 2014. The objective of this laser time transfer is the synchronization of the ground based clocks and the clock on board the space station with precision of a few picoseconds and the accuracy of 50 picoseconds [1].

Although the signal photon flux at the ISS orbit is of the order of 10^{13} photons per square meter per one laser shot and multi-photon signal strength may be obtained, the photon counting approach to the optical signal detection has been selected in order to reduce the systematic biases as much as possible.

The project is a spin-off of the existing projects of laser ranging to artificial Earth satellites [2] (SLR). The on-board hardware consists of a Comer Cube Retro-reflector (CCR), an optical receiver based on a Single Photon Avalanche Diode (SPAD) and an event timing device connected to the local time scale, see Fig. 1. The ultra short laser pulses fired towards the satellite by a ground laser ranging station will be time tagged with respect to the ground time scale ***T.*** They will be detected in space and time-tagged in the local time scale ***E***. At the same time, the CCR will re-direct the laser pulse towards the ground station providing precise ranging information ***D*** and hence providing the information about the ground-to-space signal propagation delay. This procedure should provide, among others, the time transfer ground to space with precision and accuracy outperforming the radiofrequency techniques.

Because of the ISS is at low orbiter (400-500 km) the signal photon flux at the detectors input optics is of the order of 10^{13} photons per square meter per one laser shot and multi-photon signal strength may be obtained, the photon counting approach to the optical signal detection has been selected in order to reduce the systematic errors as much as possible. Hence the requirements put on the detector optics are large attenuation of the receiving signal and the compensation of the radar equation. The input optics must not influence the timing and stability properties.

It is required that the detection timing resolution must be 25 ps with the timing stability of 1 ps per day. The entire ELT package must be able to operate at temperatures from -10 to +40 °C and the maximum mass can not exceed 0.5 kg with maximum power consumption 1 W. The ELT project is planned for three years mission in space and there will be no sun protection included.

The absolute calibration of the signal delays within the laser time transfer signal chain is very challenging. Among others the optical to electrical detection delay of the photon counting receiver itself should be calibrated with the uncertainty of 25 ps maximum.

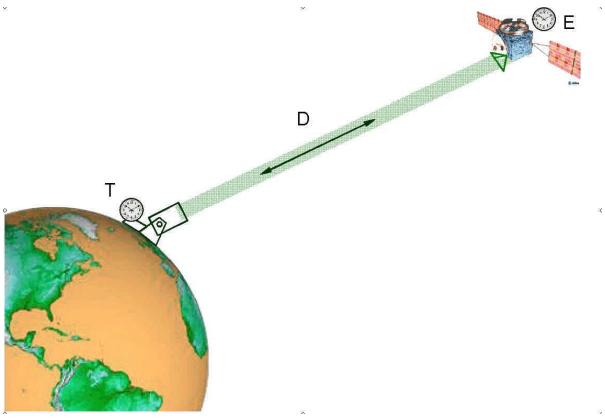


Figure 1. The principle of the laser time transfer ground to space and vice versa.

II. ELT CONSTRUCTION

The detector under construction is based on the Single Photon Avalanche Diode (SPAD) having an active area of 100 µm in diameter. The detection chip is manufactured by the K14 technology and is mounted in un-cooled socket and is operated in active quenching and gating mode. Designing the detector electronics with detection delay stability of 20 ps peak to peak (with the goal of 10 ps) over one satellite orbit, the operating temperature range of -10 °C to +40 °C and last but not least the radiation stability in a near Earth orbit environment were rather challenging. The classical schemes used until now could not be used. The space qualified electronic components had to be used whenever possible. The electronics block diagram is in Fig. 2.

Two linear stabilizers provide the biases +5.0 V and +3.3 V for the detector logic. The DC-DC converter provides +35 V, 10 mA for the SPAD bias stabilizer. The SPAD bias stabilizer is stabilizing the SPAD reverse bias according to the detector temperature hence the detector delay temperature dependence is compensated. The ultra-fast PECL comparator is sensing the SPAD break down after a photon is detected. The gate and quenching logic is based on LVDS and CMOS circuits. The circuit enables to operate the SPAD chip 0.1 V up to 2.0 V above its break down.

The photograph of the bread board (BB) version of the active quenching and gating circuit is in Fig. 3. The detection chip is connected in the centre from the bottom side of the printed circuit board. The power, gating signal, and test signals are connected via a multi-pin connector (left). The DC-DC converter providing the SPAD bias voltage is in the lower part, the SPAD bias stabilizing and temperature control circuit is in the left. The +5.0 V and +3.3 V power stabilizers are on the top, the fast comparator sensing the SPAD electrical output is in the center with the gating and level converting circuits on the right. The total board dimensions are 74 × 74 mm, its mass is 42 g.

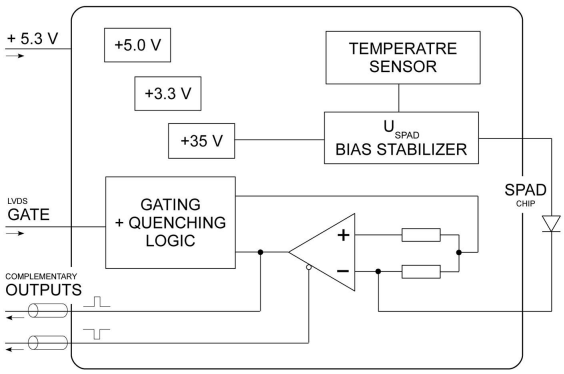


Figure 2. The ELT detector eletronics block diagram.

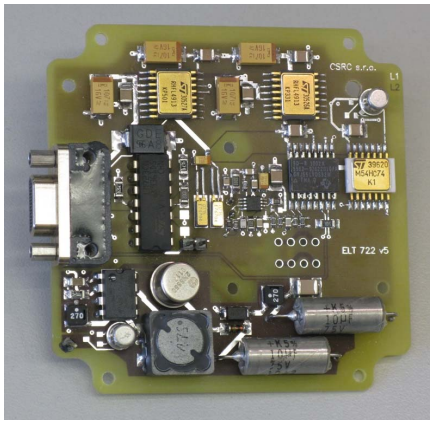


Figure 3. Bread board version of the active quenching and gating circuit. Entire components are space qualified or their military specified pin to pin compatible versions.

Entire components which are used for ESA ELT mission are space qualified and where approved by ESA except of the fast comparator sensing the SPAD breaks down. Due to extreme requirements on the propagation delay stability over a broad temperature range the Analog Devices ADCMP 553 was selected for our purpose. The manufacturer claims propagation delay of 500 ps with the temperature coefficient of 0.2 ps / K. We were allowed to test the comparator by ourselves by ESA approved radiation test procedure. The comparator survived three times larger radiation dose then it was prescribed and any of the comparator parameters had changed.

III. TEST PROCEDURE

The ELT detector package was tested in series of Time Correlated Photon Counting (TCPC) experiments. The Block diagram of the experimental setup is in Fig. 4. The gate pulse opens the detector detection window and consequently it is programmable delayed to trigger picoseconds laser. The picoseconds optical pulses were generated using Hamamatsu laser pulsar C4725, which provides optical pulses 42 ps long at the wavelength of 778 nm. The emitted picoseconds laser pulses are then spitted in to a strong signal and single photons pulses. The strong optical signal triggers photo-diode working in linear mode and the single photons are detected using ELT detector and the first photon detection closed the detection window. The time intervals between strong signal and single photos detections were measured by CTU sub-picoseconds Event Timer [3]. The recorded histogram corresponds to the optical waveform. The typical timing resolution of this setup was 18 ps rms.

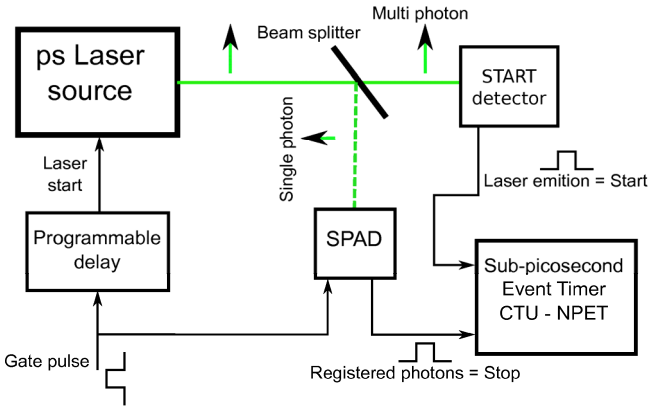


Figure 4. Time Correlated Single Photon Counting experimental setu.

IV. THERMAL PROPERTIES

The electronics together with the detection SPAD chip were installed into the housing of the engineering model. The thermal properties of the detector were tested in TCPC experiment where the internal temperature was monitored using the build in temperature sensor sensing the board temperature.

At first the temperature as a function after Power ON was measured. The heat exchange to the surrounding environment was minimized by thermal insulation of the entire package. From the measurement one can conclude, that the normal operation and high delay stability of the detector can be achieved as soon as 5 minutes after Power ON, because the delay stability is maintained for temperature range 6.5 °C peak-to-peak change. The detection delay in 80 minutes changed by 1 ps after 5 minutes after Power ON.

One of the most challenging requirements on the detector design was its long term stability and stability versus operating temperature changes. The detector is expected to be operated within the temperature range of -10°C to +40°C. The ESA requirements put on detection delay change is 20 ps with the goal of 10 ps for the temperature variation, which is expected during one ISS orbit. The estimation of the temperature change is 6.5°C. The temperature variations of the detection delay were tested in the TCPC experiment, where the detector was installed inside the thermostat. The detections delay was measured for a set of board temperatures, the results are in Fig. 5. The measured characteristics was fitted with a linear slope of 0.6 ps/K hence for 6.5 °C temperature range the maximum detection delay change is 3.9 ps. This value is well below the required 20 ps and much less to the 10 ps goal.

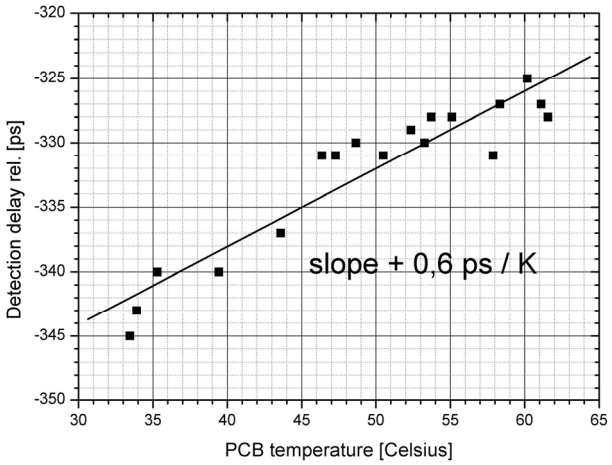


Figure 5. The detection delay temperature dependance, room temperature up to 60 °C. The detection delay coefficient is 0.6 ps/°C.

V. DETECTION LONG TERM STABILITY

The long term stability was tested in standard TCPC experiment. The experiment was carried out during the weekend in standard laboratory conditions where was no special intention put on temperature stabilization of the room temperature. The temperature changed during the experiment by -4 °C. The measured time delays were plotted with a linear fit with the slope of 280 fs/°C, see Fig. 6.

The experiment was repeated for different repetition rates of measurement (10, 30, 100 and 300 Hz) with typical data rates of 6 to 10 % and single shot jitter of 23 ps rms. The entire experiment temperature changed was ±2 °C. The resulting detection delay stability in the form of the time deviation T_{DEV} graph is plotted in Fig. 7. The detection timing stability for repetition rate of 30 Hz was measured to be 400 fs/day. It should be mentioned here, that the long term stability is influenced by a number of additional contributors: laser source, CTU event timing unit, coaxial cables and additional electronics stability.

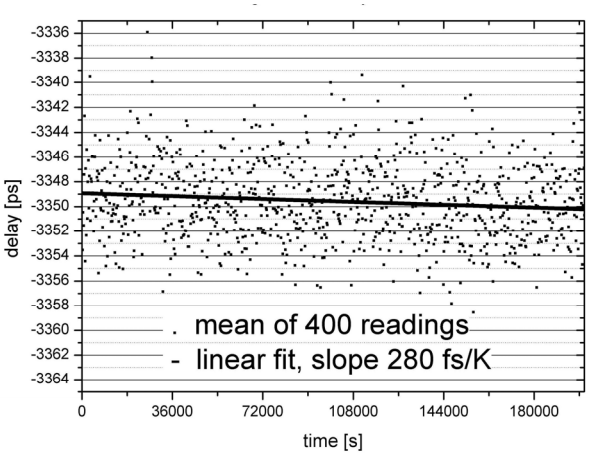


Figure 6. Detection delay long term stability test, the Time Correlated Photon Counting (TCPC) experiment was carried out during the weekend, temperature changed by -4 °C. The detection delay change has slope 280 fs/°C.

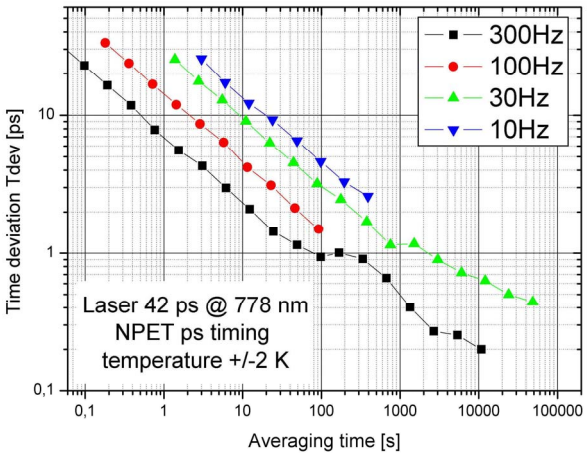


Figure 7. The ELT detector package timing delay stability. The standard TCPC experiment setup was used. The optical pulses 42 ps @778 nm, the sub-picoseconds timing system by CTU. The data rate was 6 - 10 %, the single shot timing jitter was 23 ps rms. The Tdev algorith Stable32 was used. The operating temperature changed ±2 K.

VI. Absolute Detection Timing Delay

We have developed the measurement technique and have measured the absolute value of the optical to electrical detection delay of the solid state photon counter. The individual contributions to the photon detection delay with rough estimate of their typical values are listed in Fig. 8. It should be mentioned that all the listed contributions are temperature dependent. The overall dependence of the change in detection delay with temperature was measured in an independent experiment, which was described in previous paragraph. The photon counting is a quantum mechanical process, the event of arrival of the photon of interest at the detector input can be determined only indirectly. The entire experiment is described in [4]. The absolute value of the time interval between the time of arrival of the signal photon onto detector input aperture and the time when the electrical output signal is exceeding the predefined level was determined with the precision of 10 to 12 ps. The acceptable mean values reproducibility and entire experiment stability was achieved. We have also demonstrated the dependence of the measured detection delay and hence the final result on the recording bandwidth.

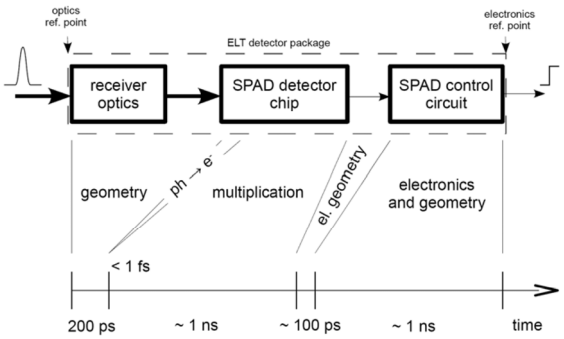


Figure 8. The individual contributions to the photon detection delay with rough estimate of the typical values.

VII. Receiver Optics

The ELT detector front-end consists of input aperture with diffuse polycarbonate 3 mm in diameter, narrow bandwidth interference filter, stack of pinholes and SPAD, as it is shown in Fig. 9. Because the entire experiment will work only with optical signal strength in single photon level, the impacting energy level must be kept as constant as it is possible for entire range of angles. It is not possible to range in angle interval form 0 to 5 degree, hence the detector should be blind in this angle interval otherwise the background photon noise scattered from the Earth will undesirably increase the detector effective dark count rate on the other side the angle interval is limited by the SLR station minimum ranging elevation, in most of the stations it is possible to range up to 60 degree from the zenith in all directions.

When the previous limitations will be taken on account, the detector optics Field of View (FoV) should be designed from 5 to 60 degree. The main advantage, which should be considered in designing the ELT optics, is that the signal must be attenuated by the optics. The expected signal strength is to be 10^{13} photons/m^2 in pulse length of 10 to 100 ps FWHM and the signal must be attenuated to the single photon level. The main attenuation is set by the distance between front aperture and SPAD.

The narrow bandwidth interference filter is used to effectively filter the background photon noise scattered from the Earth and coming from the sunlight. SLR stations work with second harmonic Nd:Yag laser - 532 nm thus the interference filter is centered at 532 nm with bandwidth of 3 nm. It should be pointed out, for several tests at CTU laboratory picoseconds laser of wavelength 778 nm was used, hence the 532 nm interference filter was replaced by 778 nm filter. However the interference filter has narrow FoV in units of degrees. The stack of pinholes is used for blocking the signal coming from larger angels and which do not propagates alongside the optical axis. The diffuse polycarbonate is used to increase the detector FoV, but it still prefers the perpendicular light. To increase the signal to noise ration the steal shield is used to block out the scattered photons from the Earth atmosphere. To minimize the angle dependency for angels from 5 to 60 degrees the circular shield was improved with conical cuts. The relative intensity varies only from 0.05 to 0.8 through entire interested interval with en-counted radar equation. The construction overview of the ELT optics can be find in [5].

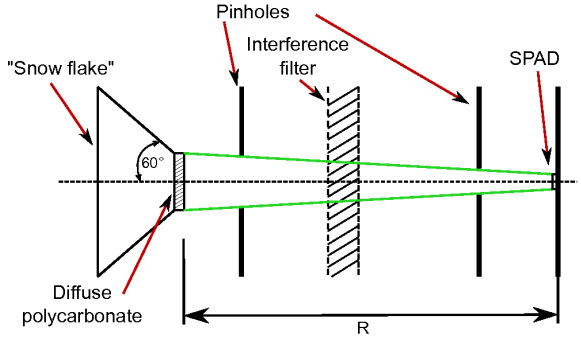


Figure 9. Schematic drawing of the receiver front-end optics.

VIII. CONCLUSION

We have developed and constructed a photon counting receiver, which is prepared for the European Laser Timing (ELT) experi-ment in space. The photon detection probability > 20 % at 532 nm wavelength, timing resolution better than 25 ps rms and the detection delay long term stability better than 1 ps were achieved. The detector package is capable to operate within a temperature range of -10 °C up to +40 °C, the detection delay change with the temperature is +0.6 ps/°C. The delay between the event of photon absorption and the appearance of the electrical pulse on the detector output was determined. The newly developed experimental technique permits to determine this detection delay within accuracy of 12 ps. The new configuration of the optical receiver should maintain uniform sensitivity over a wide field of view of ±5 to 60 degrees from nadir. Several radiation tests according ESA and ASTRIUM specifications were completed. The detector passed the ESA Critical Design Review; recently the engineering module is under construction.

ACKNOWLEDGMENT

This work has been carried out at the Czech Technical University in Prague. The research and development of solid state photon counting detectors and their applications is supported MSM6840770015.

REFERENCES

[1] Schreiber, U. - Procházka, I. - Lauber, P. - Hugentobler, U. - Schäfer, W. - et al.: Ground-Based Demonstration of the European Laser Timing (ELT) Experiment. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control (UFFC). 2010, vol. 57, no. 3, p. 728-737

[2] Procházka, I. - Yang, F.: Photon counting module for laser time transfer via Earth orbiting satellite. Journal of Modern Optics. 2009, vol. 56, no. 2&3, p. 253-257

[3] Pánek, P.; Procházka, I.: Time interval measurement device based on surface acoustic wave filter excitation, providing 1 ps precision and stability. Review of Scientific Instruments, vol. 78, no. 9, Sep 2007.

[4] Ivan Prochazka , Josef Blazej and Jan Kodet, "Measurement of the optical to electrical detection delay in the detector for ground-to-space laser time transfer", Metrologia 48 (L13), (2011); doi: 10.1088/0026-1394/48/3/L01

[5] Jan Kodet, Michael Vacek, Petr Fort, Ivan Prochazka and Josef Blazej, "Photon counting receiver for the laser time transfer, optical design, and construction", Proc. SPIE 8072, 80720A (2011); doi:10.1117/12.883701

The Dual Pseudo-Random Noise TWSTFT Time Transfer Experiment between NICT and TL

Yi-Jiun Huang, Wen-Hung Tseng*, Shinn-Yan Lin, Huang-Tien Lin and Chia-Shu Liao
Telecommunication Laboratories, Chunghwa Telecom
Taoyuan, Taiwan
dongua@cht.com.tw

Kai-Ming Feng
*Institute of Photonics Technologies, National Tsing Hua University
Hsinchu, Taiwan
whtseng@cht.com.tw

***Abstract*—A precise two-way satellite time and frequency transfer (TWSTFT) data by using the Dual Pseudo-random noise (DPN) signal was correspondingly compared with the regular TWSTFT and GPS Precise Point Positioning (GPSPPP) between the National Institute of Information and Communications Technology (NICT) and Telecommunication Laboratories (TL) in this study. Our result shows that the frequency stabilities of DPN are competitive with those of the GPSPPP. There was less obvious diurnal in the DPN time transfer data. Since the DPN system has advantages of higher precision and lower bandwidth cost when compared with the regular TWSTFT, it is one of the most potential methods to improve the international time transfer links.**

I. INTRODUCTION

International Atomic Time (TAI) and Coordinated Universal Time (UTC) are generated based on the comparison of atomic time scales between time-keeping labs around the world. The Two-Way Time and Frequency Transfer (TWSTFT) technique has been widely employed to achieve a high precision of time comparison [1]-[4]. Two labs perform TWSTFT for exchanging the time information by transmitting and receiving the time signals by using two earth stations and a limited bandwidth from a commercial satellite. Because the two-way signal paths are almost reciprocal, the 1 ns uncertainty of time accuracy and a few 10^{-15} frequency stability can be achieved at the averaging time of one day [3][4].

Based on code division multiple access (CDMA) architecture, the pseudo-random noise (PRN) signal has been employed for regular TWSTFT [5]. Recently, regular TWSTFT employs PRN signal with chip rate 2.5 Mcps or 1 Mcps [6]. Each time-keeping lab can receive the desired PRN signal from remote labs. In order to improve the precision, the possible way is to increase the chip rate of PRN signal. However, the link fee of bandwidth will be increased when increasing the chip rate. To improve TWSTFT precision without increasing the link fee, the National Institute of Information and Communications Technology (NICT) has developed the Dual PRN (DPN) signal [7], which is based on the Binary Offset Carrier (BOC) signal that will be further used on Galileo and GPS civil navigation systems. The first zero-baseline DPN test has been performed through 2006 [8]. It is a common clock test through the JCSAT-1B satellite. On this experiment, a software-defined radio technique was employed and the resulting data were post-processed. In 2008, another common clock DPN test with 500 m short baseline has been performed via IS-8 satellite [9]. In 2010, the first international DPN link was established by NICT and Telecommunication Laboratories (TL) [10]. The DPN TWSTFT experiment was performed to compare the UTC(NICT) and UTC(TL) with a baseline longer than 2000 km. It used parallel computing technique to achieve near-real-time data processing.

The aim of this paper is to demonstrate the high precision of the DPN data. We show the DPN data as well as those of the GPS Precise Point Positioning (GPSPPP) [11] in the short-term, and compare them with those of regular TWSTFT in the long-term. Since DPN system has not been fully developed, we list some challenges to be further overcome in performing DPN experiment, and propose some methods to improve its robustness and precision.

II. EXPERIMENT SETUP

The NICT in Tokyo and TL in Taoyuan have performed the first oversea international DPN experiment for a consecutive 9 months during 2010. The baseline between NICT and TL is about 2113 km, and both use H-Maser as their primary clock. This link was established via a geo-stationary satellite INTELSAT-8 (IS-8) as the signal relay. Meanwhile, the regular Asia-Pacific TWSTFT experiment was also conducted via the same satellite. The DPN modem was developed by NICT, employing the software-defined radio technology. Each lab had a DPN modem composed of an Arbitrary Waveform Generator (AWG), an Analog-to-Digital Sampler (SMP) and a personal computer. We used a customized AWG to generate the time signal. The AWG's memory can store any data pattern digitalized voltage level, and the data in memory can be periodically generated with a fixed rate 204.6 MHz. On this experiment, the time signal was

This work is support in part by the Bureau of Standards, Metrology & Inspection, Ministry of Economic Affairs, Taiwan under Grant number BSMI 100-1403-05-05-13.

configured as the combination of two replica PRN signals with a chip rate of 127.75 kcps and separated by 20 MHz. Fig. 1 shows the comparison of the signal spectra between the DPN and the conventional PRN signals. The conventional PRN signals with a chip rate of 2.5MHz occupied a bandwidth of 4 MHz. The occupied bandwidth of DPN was only a double of 250 KHz. But, the DPN's cover of spectra can be a bandwidth of 20 MHz.

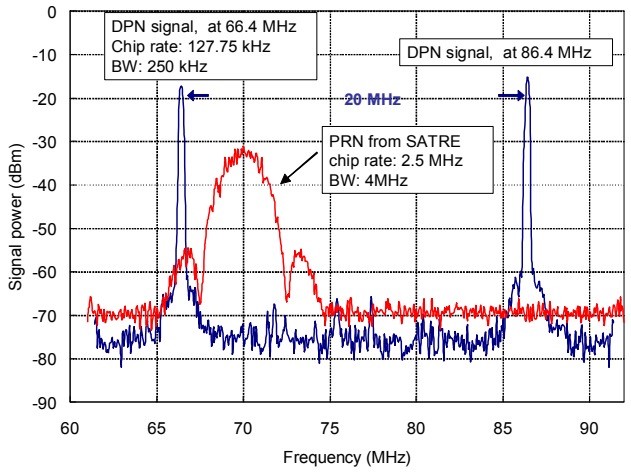


Figure 1. Power spectra of DPN (blue line) and regular TWSTFT (Red line).

For the receiving signal, we used the SMP to digitalize it and a series of sampled data stored in a buffer ring was fed. The quantization level and the sampling frequency were set to 4 bits and 64 Msps, respectively. It was necessary to set the first time tag carefully on the SMP so that all the samplings had a unique and correct time tag. The customized SMP needed manually set the first time tag, and the time tag can automatically count up when external 1PPS trigger came in.

All the samples were fed to a personal computer. The computer with a commercial GPU graphical card can provide Compute Unified Device Architecture (CUDA) library for parallel computation so as to reduce signal processing time. The DPN software was developed by using parallel computation to achieve near-real-time signal processing. NICT and TL exchanged 1-sec data through the internet and the time difference results can be derived.

III. RESULTS

Currently, two independent techniques, TWSTFT and GPSPPP, are widely used for comparing precise time difference between two time-keeping labs. TL performed the regular and DPN TWSTFT experiments last year. In the following, we demonstrate the experimental results of the regular TWSTFT, DPN and GPSPPP, to compare UTC(NICT) and UTC(TL).

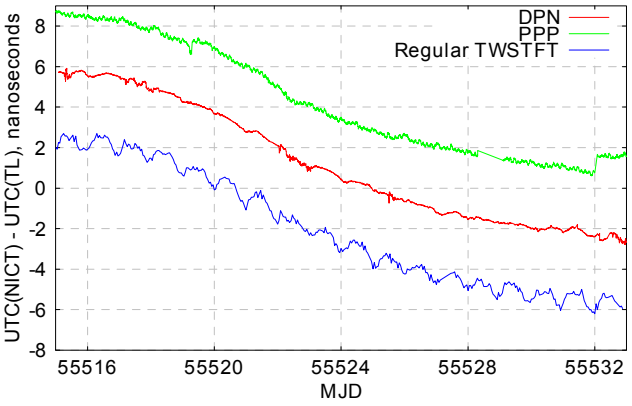


Figure 2. The time difference results of DPN (red line), GPSPPP (green line) and regular TWSTFT (blue line).

The time differences from MJD 55515 to 55532 are compared in the Fig. 2. Three curves are in a good agreement in a long term except a 0.9 ns jump on GPSPPP at MJD 55532. That jump reduces the performance, and it is regarded as a month boundary because GPSPPP's data are processed every month. The diurnal effect is about 0.85 ns on regular TWSTFT data, while that is 0.25 ns on DPN data.

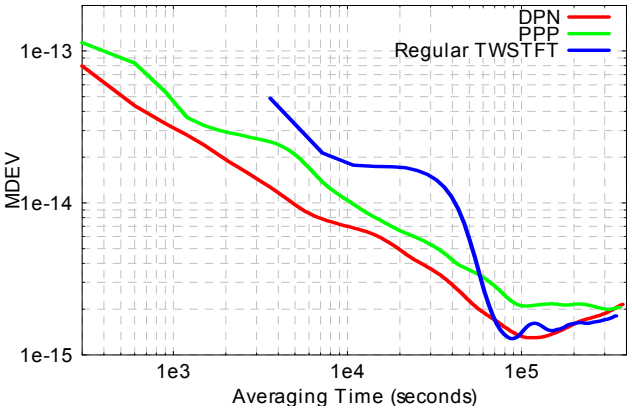


Figure 3. The modified Allan deviations of DPN (red line), GPSPPP (green line) and regular TWSTFT (blue line)

The modified Allan deviations (MDEV) of these experimental results are shown in Fig. 3. Among these results, DPN performs as well as GPSPPP in comparing the frequency stability. They could reach the 10^{-15} level for the averaging times longer than 10^4 seconds. In order to avoid the time scale discontinuity, the DPN data are aligned with the regular TWSTFT. Therefore, its frequency stability is steered by regular TWSTFT in long-term.

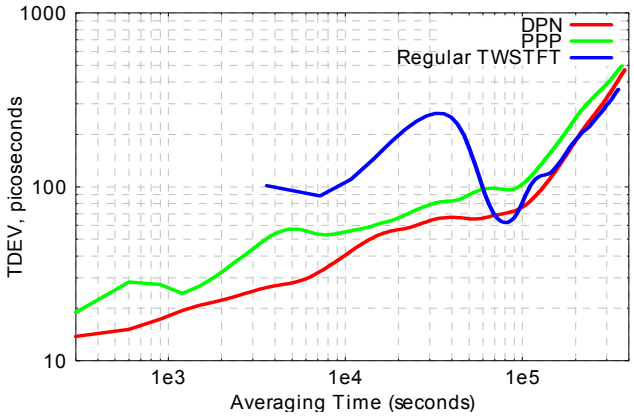


Figure 4. The time deviations of DPN (red line), GPSPPP (green line) and regular TWSTFT (blue line)

The time deviations (TDEV) are shown in Fig. 4. From this figure, the time stabilities of DPN and GPSPPP reach a level of 10 ps in the short term. On TWSTFT, we can find the diurnal exists both on regular TWSTFT and DPN results, for the curves are bumped around the averaging time of half a day.

The diurnal effect of TWSTFT leads a bad performance, and was discussed in many precious materials [12]-[15]. We found the DPN resulting data are much less diurnal than regular TWSTFT data. It can be well degraded by employing the DPN technology. Although the diurnal effect always exists in the TWSTFT system, its level should be affected by signal design and receiver methodology.

IV. CHALLENGES ON DPN

In this section, we would like to discuss some deficiencies that would affect the performance of the DPN experiment. In Fig. 5, we observe that the time difference data are somewhat noisy in the daytime. We found the AC power was not stable because of its overloading. A UPS was helpful to filter unstable AC power. After adding the UPS between the AC power and the input power supply, the time difference data became less noisy.

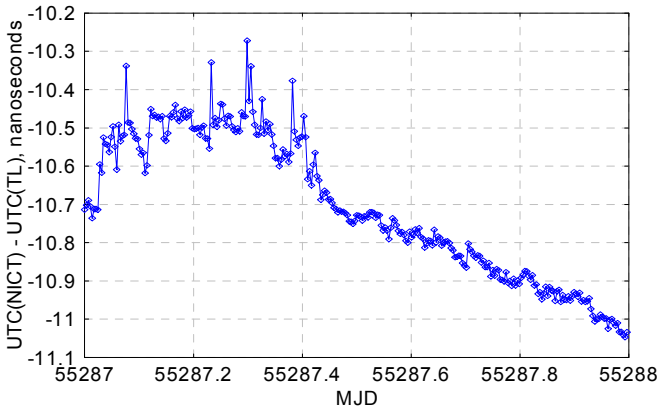


Figure 5. The unexpercted jumps occurred from MJD 55287 to MJD 55287.4 caused by the AC power instabilities

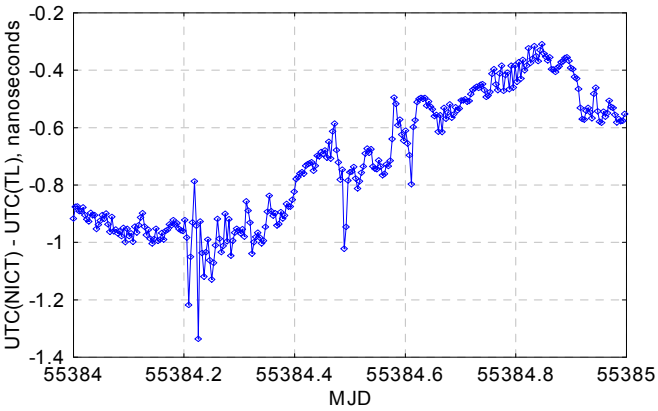


Figure 6. The unexpercted variation around MJD 55384.2 and MJD 55384.5 caused by rain fade

Sensitivity to the rain fade is another instable source on DPN experiment. The tracking process would be failed when outdoor environment was in high humidity. In Fig. 6, there are some the discontinuities of time difference data when the tracking process suspended. To improve data discontinuity due to rain fade, the integration time was increased from 4 ms to 20 ms.

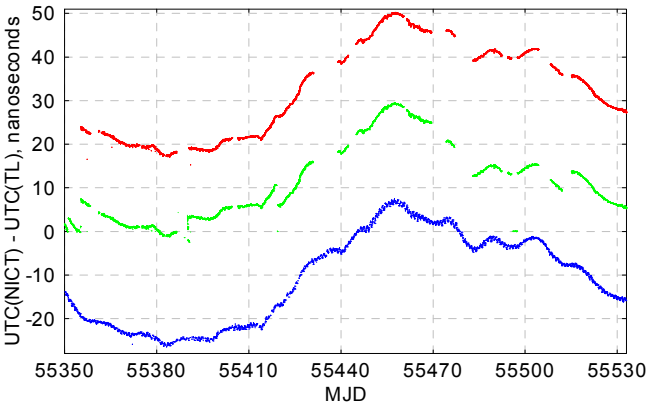


Figure 7. The time difference results of DPN with alignment (red line), without alignment (green line) and that of regular TWSTFT (blue line).

For power blackout, the DPN experiment had been restarted for some times. The on-off power switching made the time difference results piecewise continuous. The internal delays of AWG and SMP would change whenever rebooting their power. To avoid the discontinuity, internal delays of AWG and SMP need to be externally monitored in the future work. Nevertheless, regular TWSTFT data were helpful to compensate this discontinuity on DPN results. Since regular TWSTFT and DPN were performed on the same Asia-Pacific link, their results should be consistent in long-term. In Fig. 7, we aligned DPN data with regular TWSTFT in one-day average. Thus aligned DPN results can be seen as a combination whose performance was determined by DPN in short-term and by regular TWSTFT in long-term.

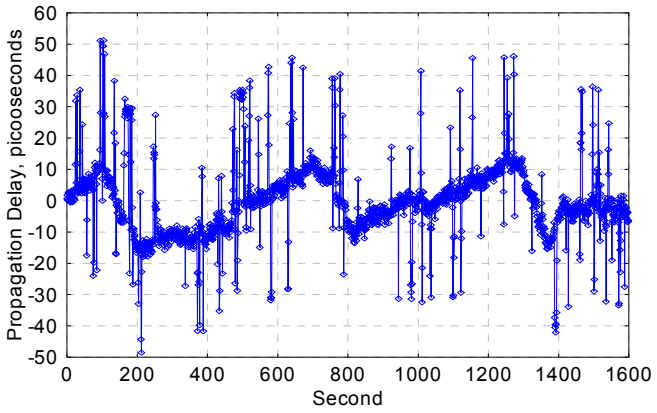


Figure 8. The spurious noise randomly occurred when measuring a propagation delay over a 10 km optical fiber by DPN mothod.

Finally, we carried out a loop-back test to check the basic performance of DPN. We measured the signal propagation delay over a 10 km optical fiber by using DPN technology. The carrier to noise ratio after passing the fiber is still larger than 80 dB/Hz, which is much higher than the 55 dB/Hz as in the satellite system. In Fig. 8, some spurious noise can be seen on the time difference data. With this spurious noise, the standard deviation of delay data is only 14 ps. After filtering out the noise, the standard deviation of 3 ps can be reached. This means that the short-term stability of DPN could be even better if the spurious noise could be avoided in the future design.

V. CONCLUSION

The DPN modem developed by NICT is composed of commercial devices and using software-defined radio technology. The first oversea DPN experiment has been successfully performed between NICT and TL in 2010. We compared the time transfer measurements among DPN, regular TWSTFT and GPSPPP for the consecutive 18 days. The DPN shows an outstanding performance which is competitive with that of GPSPPP in the short-term stabilities. And, it is clear that DPN has much less diurnal variation than the regular TWSTFT. Moreover, the main advantage of the DPN is its lower satellite bandwidth cost. In the future work, we hope to overcome some noise sensitivities of DPN to enhance its precision and keep using this technology for time comparisons.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Tadahiro Gotoh, Jun Amagai, Thomas Hobiger and other staff at NICT for developing the GPU-based DPN modem and a great deal of cooperation of these experiments.

REFERENCES

[1] D. Kirchner, "Two-Way Time Transfer via communication satellites," Proceedings of the IEEE, Vol. 79, No. 7, pp. 983-990, 1991.

[2] W.-H. Tseng, S.-Y. Lin, K.-M. Feng, M. Fujieda and H. Maeno, "Improving TWSTFT Short-Term Stability by Network Time Transfer," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 57, No. 1, pp. 161-167, 2010.

[3] D. Piester, A. Bauch, L. Breakiron, D. Matsakis, B. Blanzano, and O. Koudelka, "Time transfer with nanosecond accuracy for the realization of International Atomic Time," Metrologia, Vol. 45, No. 2, pp. 185-198, 2008.

[4] A. Bauch, J. Achkar, S. Bize, D. Calonico, R. Dach, R. Hlavac, L. Lorini, T. Parker, G. Petit, D. Piester, K. Szymaniec, P. Uhrich, "Comparison between frequency standards in Europe and the US at the 10^{-15} uncertainty level," Metrologia, vol. 43, no. 1, pp. 109-120, 2006.

[5] ITU Radiocommunication Sector, The operational use of two-way satellite time and frequency transfer employing PRN codes, 2003. Recommendation ITU-R TF.1153-2, Geneva.

[6] V. Zhang, T. Parker, J. Achkar and D. Rovera, "Transatlantic 2.5 MChip/S Two-Way Satellite Time and Frequency Transfer with Surface Acoustic Wave Filters," in Proc. 2011 Joint Conference of the IEEE International Frequency Control Symposium and European Frequency and Time Forum, 2011, to be published

[7] J. Amagai, "Two-Way Satellite Time and Frequency Transfer using a Pair of Pseudo Random Noises," in Proc. Asia-Pacific Workshop Time Freq., 2006.

[8] J. Amagai and T. Gotoh, "Current Status of Two-Way Time and Frequency Transfer via Satellite," in Proc. Asia-Pacific Workshop Time Freq., 2008.

[9] T. Gotoh, J. Amagai, "Two-Way Time Transfer with Dual Pseudo-Random Noise Code," in Proc. 40th Precise Time and Time Interval (PTTI) System and Applications Meeting, 2008, pp.459-466.

[10] T. Gotoh, J. Amagai, T. Hobiger, M. Fujieda, and M. Aida, "Development of a GPU based two-way time transfer modem," IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 7, 2011, to be published.

[11] T. Feldmann, A. Bauch, D. Piester, H. Esteban, J. Palacio, F. J. Galindo, T. Gotoh, H. Maeno, U. Weinbach, S. Schön, "GPS carrier phase and precise point positioning time scale comparisons using different software packages," in Proc. European Frequency and Time Forum – IEEE Frequency Control Symposium Joint Conference, 20-24 Apr 2009, Besançon, France, pp. 120-125, 2009.

[12] W.-H. Tseng, K.-M. Feng, S.-Y. Lin, H.-T. Lin, Y.-J. Huang, and C.-S. Liao, "Sagnac Effect and Diurnal Correction on Two-Way Satellite Time Transfer," IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 7, 2011, to be published.

[13] D. Piester, A. Bauch, M. Fujieda, T. Gotoh, M. Aida, H. Maeno, M. Hosokawa and S. H. Yang, "Studies on Instabilities in Long-Baseline Two-Way Satellite Time and Frequency Transfer (TWSTFT) Including a Troposphere Delay Model", in Proc. 38th Precise Time and Time Interval (PTTI) System and Applications Meeting, 2007, pp. 211-222.

[14] T. E. Parker and V. Zhang, "Sources of Instabilities in Two-Way Satellite Time Transfer," in Proc. 2005 IEEE International Frequency Control Symposium (IFCS), 2005, pp. 745-751.

[15] M. Fujieda, T. Gotoh, M. Aida, J. Amagai, H. Maeno, "Long-Baseline TWSTFT between Asia and Europe," in Proc. 38th Precise Time and Time Interval (PTTI) System and Applications Meeting, 2006, pp. 499-504.

Transatlantic 2.5 MChip/s Two-Way Satellite Time and Frequency Transfer with Surface Acoustic Wave Filters

Victor Zhang, Thomas E. Parker
Time and Frequency Division
National Institute of Standards and Technology (NIST)
Boulder, Colorado, U.S.A
vzhang@boulder.nist.gov, tparker@boulder.nist.gov

Joseph Achkar, Daniele Rovera
Time Metrology Group
LNE-SYRTE, Observatoire de Paris (OP)
Paris, France
joseph.achkar@obspm.fr, daniele.rovera@obspm.fr

***Abstract*—From August 2010 to April 2011, NIST and OP conducted an experiment of using surface acoustic wave (SAW) filters in transatlantic two-way satellite time and frequency transfer (TWSTFT) with 2.5 MChip/s pseudo-random noise (PRN) codes. The SAW filters used in the experiment are 2.5 MHz band-pass filters with a center frequency of 70 MHz. Instead of using 3.5 MHz bandwidth required for the 2.5 MChip/s signal, the use of SAW filters allows the 2.5 MChip/s TWSTFT to use only 2.5 MHz bandwidth on a satellite transponder. We evaluated the SAW filters. We compared the time transfer instability of the filtered 2.5 MChip/s TWSTFT to the 1 MChip/s TWSTFT with and without the SAW filters. This paper presents the experiment results.**

I. INTRODUCTION

The timing information of TWSTFT signals is carried by PRN codes [1]. The resolution of TWSTFT measurements is proportional to the chip-rate of the PRN code being used. The minimum bandwidth required for a TWSTFT signal is approximately 1.4 times the chip-rate of the PRN codes [2].

The transatlantic TWSTFT network used 2.5 MChip/s PRN codes with 3.5 MHz bandwidth, in Ku-band, from the early 1990s, when the network was established, until a satellite change in July 2009. The TWSTFT operation was then changed to use 1 MChip/s PRN codes with 2.5 MHz bandwidth. The chip-rate reduction was made because of the significant increase in satellite subscription cost, and our previous study [3] showed the time transfer instability, as measured by the Time deviation (TDEV), of the 1 MChip/s TWSTFT was about the same as the TDEV of the 2.5 MChip/s TWSTFT for averaging times of one day and longer. The SAW filters were not available at the time of our previous study. We acquired 2.5 MHz bandwidth on the new satellite so we could do experiments on the 2.5 MChip/s TWSTFT with SAW filters after the filters were made available in early 2010.

The concept of 2.5 MChip/s TWSTFT with SAW filters was proposed by colleagues at the National Institute of Information and Communication Technology (NICT) in Japan. Previous experiments were conducted in the Asian-Pacific TWSTFT network. The scheme takes advantage of the high measurement resolution of the 2.5 MChip/s codes, but limits the TWSTFT signal to 2.5 MHz bandwidth with SAW filters. To study the performance of the filtered 2.5 MChip/s TWSTFT with SAW filters across a transatlantic link, NIST and OP conducted an experiment from August 2010 to April 2011with the SAW filters kindly loaned to us by TimeTech GmbH* of Germany. We show characteristics of the SAW filters used in our experiment and the experiment configuration in Sections II and III. Analysis of the 2.5 MChip/s TWSTFT with SAW filters is presented in Section IV. Section V summarizes the experiment results.

II. CHARACTERISTICS OF THE 2.5MHZ BAND-PASS SAW FILTER

Figure 1 shows the spectrum of the 1 MChip/s code (in red), main lobe of the 2.5 MChip/s code (in green) and the

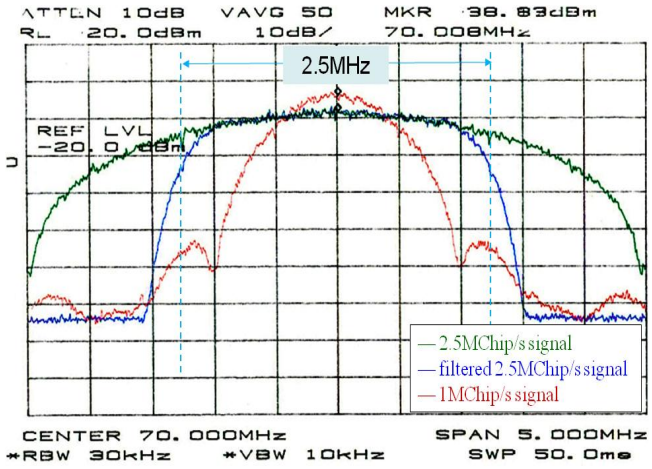


Figure 1. Spectrum of the 1 MChip/s (in red), main lobe of the 2.5 MChip/s (in green) and the filtered 2.5 MChip/s (in blue) TWSTFT signals.

This paper includes contributions from the U.S. government and is not subject to copyright.
*Commercial products and companies are identified for technical completeness only, and no endorsement by authors and organizations is implied

filtered 2.5 MChip/s code (in blue). When comparing the filtered 2.5 MChip/s signal to the 2.5 MChip/s signal, we see that the SAW filter provides unity gain to the 2.5 MChip/s signal in a range of about 2 MHz centered at 70 MHz. The 2.5 MChip/s signal is attenuated about 15 dB at the 2.5 MHz boundaries. Because the main lobe of the 1 MChip/s signal is bounded by 2 MHz, passing a 1 MChip/s signal through the SAW filter does not distort the main lobe.

We evaluated the SAW filter's characteristics by a loop test using a SATRE* modem manufactured by TimeTech and the SAW filter unit at NIST. The peak-to-peak variation of the one-second measurements increases by about 50 ps for each SAW filter added in the loop [4]. The use of the SAW filter inserts a large signal path delay of about 5.2 µs. This delay is sensitive to temperature variations, with a temperature coefficient estimated to be 200 ps/°C. However, the delay sensitivity to temperature can be compensated by use of SAW filters in both the transmit (TX) and receive (RX) paths. Figure 2 shows the SAW filter's delay response to frequency change. The measurements were made with SAW filters in both TX and RX paths. We shifted TX and RX frequencies from -10 kHz to +10 kHz in 2 kHz steps with respect to 70 MHz. The delay dependence on frequency varies between 184 ps/2 kHz and 289 ps/2 kHz.

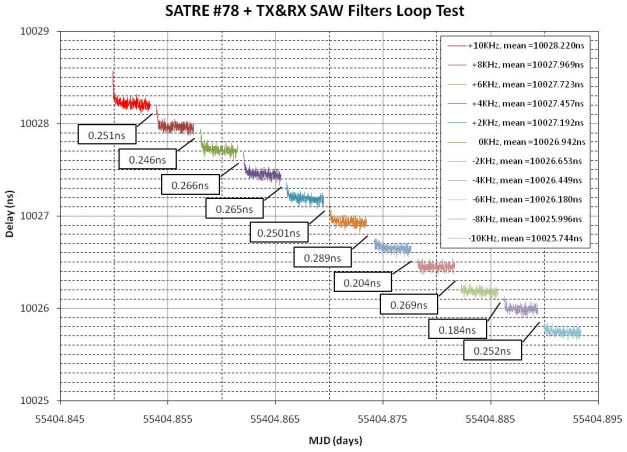


Figure 2. SAW filter's delay response to frequency change of ±10 kHz with respect to 70 MHz. Each step corresponds to the delay for a 2 kHz frequency change.

The SAW unit at OP was also evaluated by use of a vector analyzer. Figure 3 shows the measurement results of return loss (Trace 1, in yellow), insertion loss (Trace 2, in blue) and delay response to frequency change (Trace 3, in red). These characteristics were evaluated for frequencies from 68.9 MHz to 71.1 MHz. After correcting for the delays and for the losses from the cables and adaptors used in the measurements, the return loss is about 30 dB, which means the SAW filters have a very good impedance match. The insertion loss is about 0.4 dB for 70 ± 1 MHz. The delay through a SAW filter is 4.842 µs at 70 MHz, which is about 360 ns different than the SAW filter unit at NIST. Figure 3 also shows the SAW filters' delay response to frequency change is not linear.

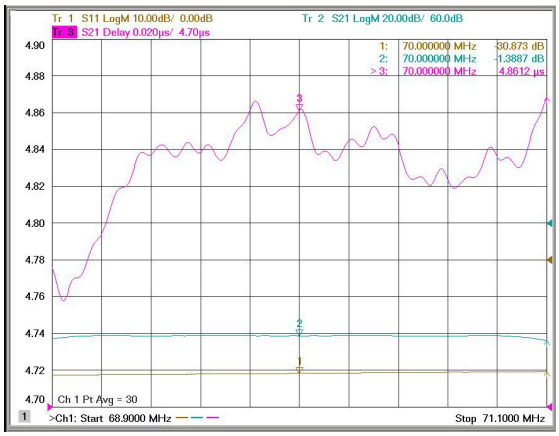


Figure 3. SAW filter's return loss (Trace 1, in yellow), insertion loss (Trace 2, in blue) and delay response to frequency change (Trace 3, in red) measured by a vector analyzer

Although the large constant delay introduced by SAW filters will be mostly canceled in the TWSTFT difference, TWSTFT links using SAW filters should be recalibrated to correct for the residual of delay difference of the two directions. The frequency of the received TWSTFT signal varies due to the daily satellite motion, or due to free running internal oscillators in some transceivers. The SAW filter's delay response to frequency change will increase the uncertainty of TWSTFT.

III. EXPERIMENT CONFIGURATION

The regular NIST/OP 1MChip/s TWSTFT measurements are made in a 2-minute session during even hours of Coordinated Universal Time (UTC), 12 times a day. The experiment measurements were made in three 2-minute sessions during odd hours of UTC. Table 1 shows the measurement schedule. The measurement of odd-hour 1 MChip/s TWSTFT without SAW filters was added in March 2011 to study the effect of SAW filters on the 1 MChip/s TWSTFT. No other station, but NIST and OP, transmitted signals during the three odd-hour sessions.

TABLE 1 TWSTFT measurement schedule

Measurements	Measurement sessions	Measurement period
1 MChip/s without SAW filter (regular transatlantic link)	2-minute in even hours, 12 times a day	August 18, 2010 to April 11, 2011 (MJDs 55426 to 55662)
1 MChip/s without SAW filter	2-minute in odd hours, 12 times a day	March 25, 2011 to April 11, 2011 (MJDs 55645 to 55662)
1 MChip/s with SAW filter	2-minute in odd hours, 12 times a day	August 18, 2010 to April 11, 2011 (MJDs 55426 to 55662)
2.5 MChip/s with SAW filters	2-minute in odd hours, 12 times a day	August 18, 2010 to April 11, 2011 (MJDs 55426 to 55662)

Figure 4 shows a block diagram of the experiment configurations at NIST and OP. Both NIST and OP used

SATRE modems for all the TWSTFT measurements. A TimeTech SAW filter unit contains two SAW filters, one for the TX path and one for the RX path. The signals for the TWSTFT with SAW filters went through the TX and RX SAW filters at both NIST and OP.

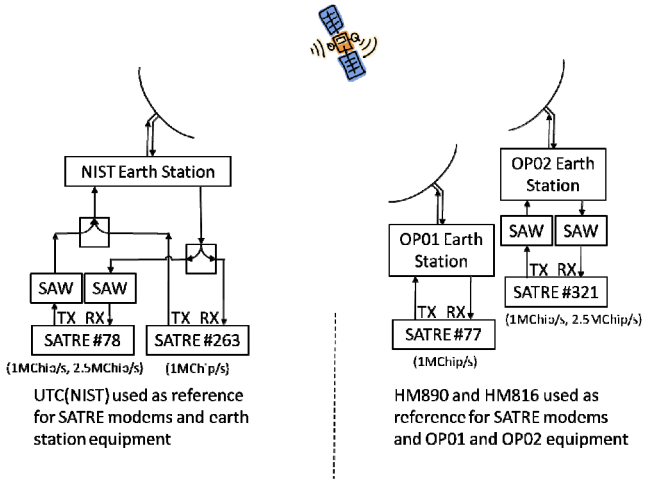


Figure 4. Block Diagram of NIST/OP TWSTFT with and without SAW filters.

At NIST, the same earth station and same radio-frequency (RF) equipment were used for both TWSTFT with and without SAW filters. A pair of power combiner and power splitter was used to channel the TWSTFT signals between the earth station equipment and the two modems. At OP, earth station OP01 was used for TWSTFT without SAW filter, and earth station OP02 was used for the TWSTFT with SAW filters.

Reference signals for the NIST TWSTFT were derived from the NIST time scale, UTC(NIST). During the experiment, reference signals for the OP TWSTFT were derived from two hydrogen MASERs, HM890 and HM816. The HM816 was used for MJDs 55496 to 55593 (October 27, 2010 to February 1, 2011) during the maintenance of HM890. The corrections for converting the OP reference signals to UTC(OP) are available, but were not used in our analysis.

IV. RESULTS OF THE 2.5 MCHIP/S TWSTFT USING SAW FILTERS

We compared the 2.5 MChip/s TWSTFT using SAW filters to the 1 MChip/s TWSTFT with and without SAW filters. We also investigated the effect of SAW filters on the 1 MChip/s PRN codes by comparing the 1 MChip/s TWSTFT with and without SAW filters. Because the NIST/OP link with SAW filters is not calibrated with respect to the NIST/OP link without SAW filters, we are not able to estimate the delay change from use of SAW filters. We will focus our analysis on the time transfer instability of the TWSTFT using SAW filters. The experiment results are shown in Figures 5 through 7.

Figure 5 shows the three TWSTFT differences between UTC(NIST) and OPH (hydrogen MASER) for MJDs from 55433 to 55662. The three differences are for the even-hour 1 MChip/s TWSTFT without SAW filters (in red), the odd-hour 1 MChip/s TWSTFT with SAW filters (in blue) and the odd-hour 2.5 MChip/s TWSTFT with SAW filters (in green). The frequency offset, drift and mean time difference are removed from the direct NIST/OPH differences. The even-hour 1 MChip/s TWSTFT without SAW filter contains larger short-term variations than the odd-hour 1 MChip/s and 2.5 MChip/s TWSTFT with SAW filters until around MJD 55610. The larger short-term variation might be introduced by interference from other TWSTFT signals transmitted at the same time with about the same frequency during even hours [4]. The reduction of short-term variation after MJD 55610 is partially attributed to a frequency offset in the NIST TX signal introduced to mitigate the multi-signal interference. Figure 5 also shows a big difference fluctuation when the HM816 was used as the reference clock at OP.

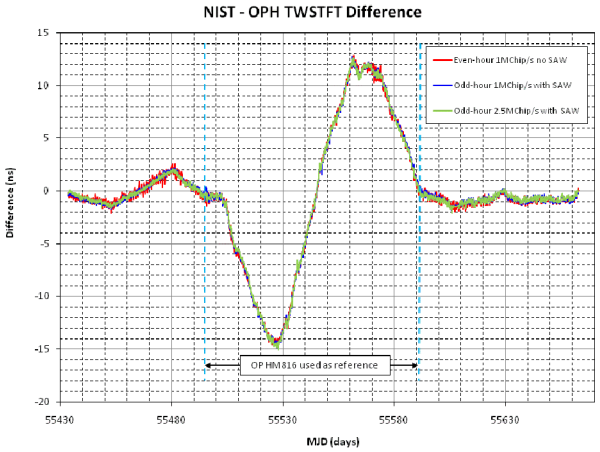


Figure 5. NIST/OPH TWSTFT differences for even-hour 1 MChip/s without SAW filters (in red), odd-hour 1 MChip/s with SAW filters (in blue), and 2.5 MChip/s with SAW filters (in green). Frequency offset, drift and mean time difference are removed from the differences.

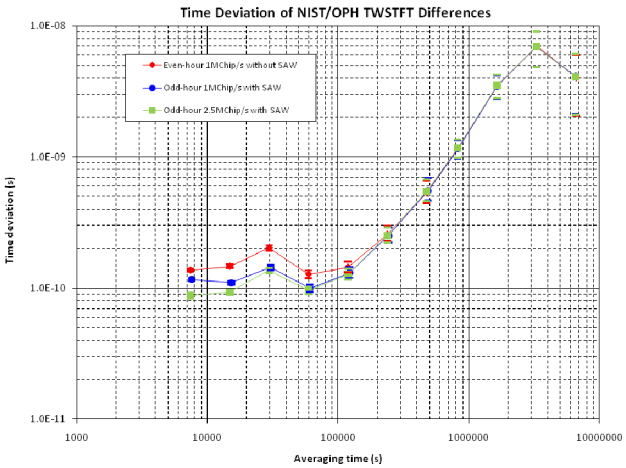


Figure 6. Time transfer instability of the NIST/OPH TWSTFT for MJDs 55433 to 55662.

TDEVs for the three differences are shown in Figure 6. All of the TDEVs are 200 ps or less for averaging times less

than 32 hours. The TDEVs are dominated by clock noise after averaging times longer than 32 hours. For averaging times of 32 hours and less, the TDEVs of the odd-hour TWSTFT with SAW filters contains less transfer noise than the even-hour TWSTFT without SAW filter which reflects the larger short-term variations in Figure 5. For averaging times from 2 hours to about 16 hours, the time transfer instability of the filtered 2.5 MChip/s TWSTFT shows a small improvement when compared to the filtered 1 MChip/s TWSTFT. This means the change of measurement noise due to chip-rate change between 2.5 MChip/s and 1 MChip/s is only a small contributor to the short-term instability. The dominant short-term instability for the transatlantic TWSTFT comes from the sources described in a 2005 study [5].

The frequency instability, as measured by the Allan Deviation (ADEV) for the three differences are shown in Figure 7. At averaging times of 1 day, the ADEVs are about 6×10^{-15} for the even-hour 1 MChip/s TWSTFT, and about 4×10^{-15} for the odd-hour 1 Mchip/s and 2.5 Mchip/s TWSTFT with SAW filters. The ADEVs are continuously averaged down to about 2×10^{-15} at averaging times of about 2.5 days. After that, the ADEVs demonstrate frequency instability mainly from the OP HM816 clock, visible in Figure 5.

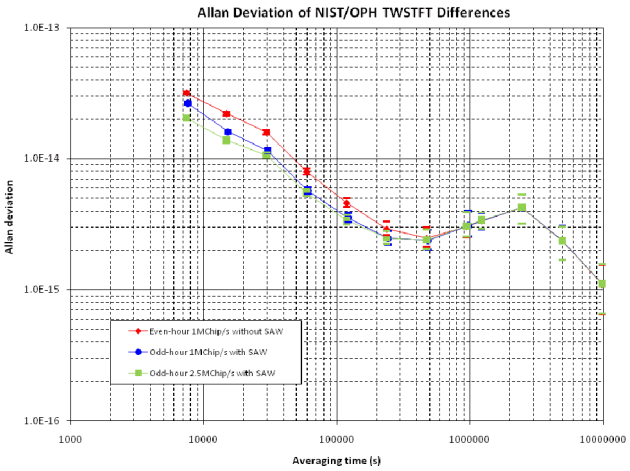


Figure 7 Frequency instability of the NIST/OPH TWSTFT for MJDs 55433 to 55662.

The comparison of time transfer instability for the 1 MChip/s TWSTFT with and without SAW filters is shown in Figure 8. Data used in computing the TDEVs were obtained from the two 2-minute measurement sessions in odd hours from MJD 55645 to MJD 55662. The measurements were made in odd-hours to minimize the interference from other TWSTFT signals transmitted at the same time with about the same frequency. The measurements of 1 MChip/s TWSTFT without SAW filters were obtained from the NIST_SATRE #263/OP01_SATRE #77 link. The measurements of 1 MChip/s TWSTFT with SAW filters were obtained from the NIST_SATRE #78/OP02_SATRE #321 link. The two measurement sessions were seven minutes apart. The TDEV for the 1 MChip/s TWSTFT with SAW filters is about 100 ps, and it is higher than that of the 1 MChip/s TWSTFT without SAW filters for averaging times of two hours and four hours. Based on our loop test results, the high level of TDEV at averaging times of two hours and four hours for the 1 MChip/s TWSTFT with SAW filters might come from the noise contributed by SAW filters. The TDEV for the 1 MChip/s TWSTFT without SAW filter increases after averaging longer than one day. That could come from the instability of the NIST_SATRE #263/OP01_SATRE #77 link.

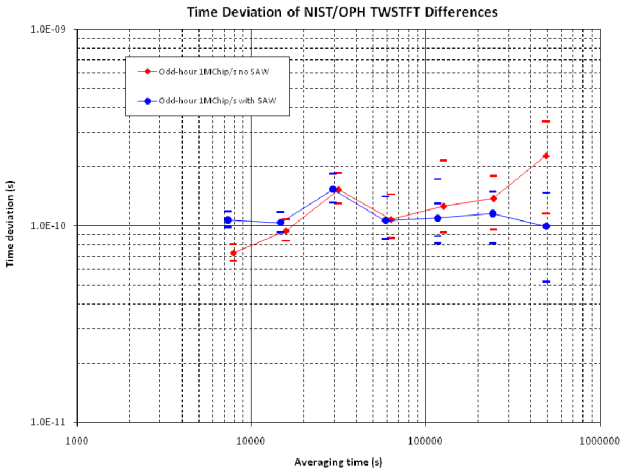


Figure 8 Comparison of 1 MChip/s TWSTFT with and without SAW filters. The TDEVs are computed with TWSTFT difference for MJDs 55645 to 55662.

All the TDEVs in Figures 6 and 8 show a diurnal at and below 200 ps. The diurnal is not from the use of SAW filters because it is also visible in the TWSTFT without SAW filters.

V. CONCLUSIONS

Our study shows the SAW filters can be used for 2.5 MChip/s TWSTFT with only 2.5 MHz bandwidth. The noise from SAW filters contributes to the short-term time transfer instability. The SAW filter's frequency dependent delay can increase the uncertainty of the TWSTFT. The 2.5 MChip/s TWSTFT with SAW filters provides a small improvement, about 10 to 20 ps in TDEV, over the 1 MChip/s TWSTFT without SAW filter for averaging times from two hours to 16 hours. The TDEV and ADEV for the filtered 2.5 MChip/s TWSTFT with SAW filters are about 100 ps and 4×10^{-15} at averaging times of one day. The TDEV and ADEV could be a little better if the TWSTFT differences had no diurnal. Because of the noise introduced by SAW filters, the 1 MChip/s TWSTFT works better without the SAW filters.

ACKNOWLEDGMENT

We thank TimeTech GmbH of Germany for loaning us the SAW filters used in our experiment. We also thank Dr. Wolfgang Schaefer for his support and advice to our study.

REFERENCES

[1] D. Kirchner, "Two-Way Satellite Time and Frequency Transfer (TWSTFT): Principle, Implementation, and Current Performance," Review of Radio Science 1996-1999, Oxford University Press, New York, NY USA, pp. 27-44, 1999.

[2] L. B. Veenstra, "International Two-Way Satellite Time Transfers Using INTELSAT Space Segments and Small Earth Stations," in Proceedings of the 22nd Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 4-6 December, 1990, Vienna, Virginia, USA, pp. 383-398.

[3] V. Zhang, T. E. Parker, J. Achkar, A. Bauch, L. Lorini, D. Matsakis, D. Piester, and D. G. Rovera, "Two-way Satellite Time and Frequency Transfer Using 1 MChip/s Codes," in Proceedings of the 41st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 16-19 November, 2009, Santa Ana Pueblo, New Mexico, USA, pp. 371-382.

[4] V. Zhang, T. E. Parker, J. Lowe, NIST Report to the 18th CCTF TWSTFT Working Group Meeting, 16-17 September, 2010, Beijing, China.

[5] T. E. Parker and Victor Zhang, "Sources of Instability in Two-Way Satellite Time Transfer," in Proceedings of 2005 Joint IEEE International Frequency Symposium and the 37th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 29-31 August 2005, Vancouver, Canada, pp. 745-751.

Dual-stage Quad-mixer Satellite Simulator for a TWSTFT Station

Faisal Mubarak[1], Erik Dierikx[2]
[1]fmubarak@vsl.nl, [2]edierikx@vsl.nl
VSL, Dutch Metrology Institute
Thijsseweg 11, 2629 JA Delft, The Netherlands

Abstract: *The new station delay calibration system, to be implemented in the VSL TWSTFT station, will perform a complete calibration of the ground station delays. The new dual-stage quad-mixer satellite simulator promises improved repeatability and less susceptibility for environmental influences. The true power of the system is the multidirectional transceiver operation within the satellite simulator. This allows including additional measurement paths, enabling absolute measurement of the path delays. Hence, the new station delay calibration system is completely stand-alone and no additional, time consuming, periodic VNA calibrations are required.*

Introduction: Soon after the introduction of the first primary time standards, the need of accurate time comparison techniques for long distances was felt. The TWSTFT has been a widely used technique for long distance time comparison since the early 90's [1]. Continuous developments and enhancements in time generation is enforcing more stringent performance from such comparison techniques as well. This text first studies the sources of instability in the TWSTFT station. In order to provide a compound solution for the problems, the TWSTFT system at VSL is completely redesigned. Resulting in a more stable and accurate delay measuring system. In figure 1 the time scale comparison principle is shown.

TA	: Time scale at station
TI	: Time Interval Counter reading
ζ_{tx}	: Transmission path delay
ζ_{rx}	: Reception path delay
ζ_{up}	: Uplink propagation delay
ζ_{dw}	: Downlink propagation delay
ζ_{trp}	: Satellite transponder delay
TC_{up}	: Sagnac correction in the uplink
TC_{dw}	: Sagnac correction in the downlink

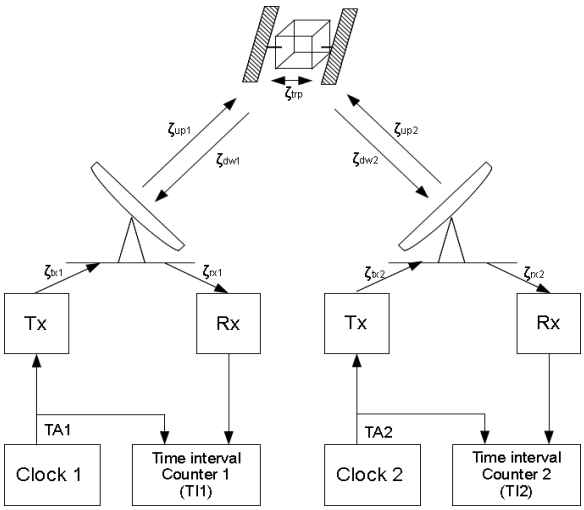


Figure 1 Principle of two way satellite time and frequency transfer system using PRN codes

$$TI_1 = (TA_1 - TA_2) + (\zeta_{tx1} + \zeta_{rx2}) + (\zeta_{up1} + \zeta_{dw2}) + \zeta_{trp} \quad (1)$$

$$TI_2 = (TA_2 - TA_1) + (\zeta_{tx2} + \zeta_{rx1}) + (\zeta_{up2} + \zeta_{dw1}) + \zeta_{trp} \quad (2)$$

The term $(TA_1 - TA_2)$ indicates the actual difference between the two compared time scales. However, to calculate the difference other terms are to be known. In [1] additional sagnac correction terms were included in equations 1 and 2, resulting in the following equations:

$$TI_1 = (TA_1 - TA_2) + (TC_{up2} + TC_{dw1}) + (\zeta_{tx1} + \zeta_{rx2}) + (\zeta_{up1} + \zeta_{dw2}) + \zeta_{trp} \quad (3)$$

$$TI_2 = (TA_2 - TA_1) + (TC_{up1} + TC_{dw2}) + (\zeta_{tx2} + \zeta_{rx1}) + (\zeta_{up2} + \zeta_{dw1}) + \zeta_{trp} \quad (4)$$

The accuracy of the compared time scales is directly linked to the accuracy of the delay components of the above equations. This text focuses on measurement of the transmitter and receiver path delays within the ground station. Detailed description of the TWSTFT principle can be found in [6],[1],[2].

Satellite Simulator Principle: The primary function of a satellite transponder is to provide a stable frequency transformation between the respective receive and transmit frequencies. A common additional function a transponder provides is the gain to compensate for the high signal path loss.

To measure receive and transmit delay components of a TWSTFT station a satellite simulator concept is used [2]. In order to simulate a satellite transponder, it is essential to provide a same frequency transformation between the respective receive and transmit frequencies of the satellite. This enables a measurement of the delay of transmit and receive path as a single component. As the satellite simulator is placed directly in front of the transmitter and the receiver antenna, the reference point of the delay component is a fixed point just in front of the TWSTFT antenna. The signal path loss is low and no additional gain is required from the satellite simulator. To determine the transmit and receive delay component separately additional measurements are needed. The additional measurements require the satellite simulator to receive and transmit PRN signals independent of the receiver and the transmitter of a TWSTFT station [2].

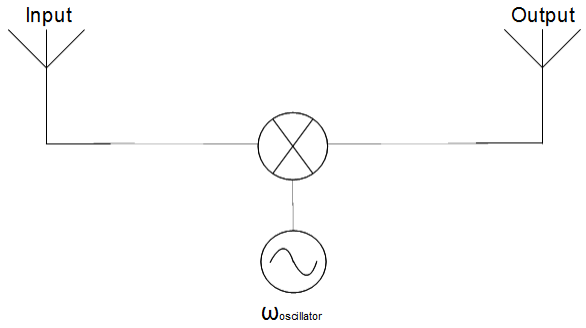


Figure 2 Single-mixer satellite simulator concept

Single-mixer architecture: In figure 2 the most simple implementation of a satellite transponder, known as the single-mixer satellite simulator , is shown [2]. Using this scheme it is only possible to measure the transmit and receive path delay as one component and it is not possible to separate the respective delay components. The main advantage of this architecture is its simplicity. This concept is very useful to monitor relative delay variation in a TWSTFT station.

Dual-mixer architecture: A satellite simulator scheme using two mixers is shown in figure 3. The two mixers perform the similar frequency transformation as the satellite transponder. Using two separate mixers driven by 70 MHz (oscillator 1) and (f_{diff} - 70 MHz) (oscillator 2) signals, the transmit and receive delay components can be measured separately. The 70 MHz oscillator port is also used to transmit and receive the PRN coded signals. Due to non-symmetrical delay path for the 70 MHz signal to the simulator's transmit and receive antennas, a systematic offset is introduced [2]. In figure 4 the spectrum transformations within the dual-mixer satellite simulator are shown.

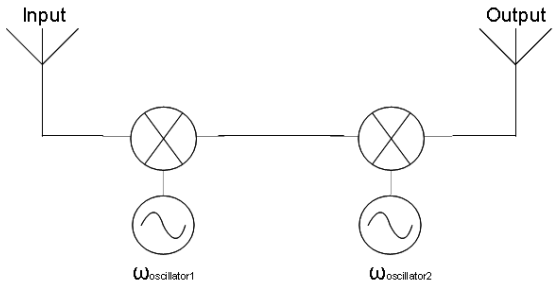


Figure 3 Dual-mixer satellite simulator concept

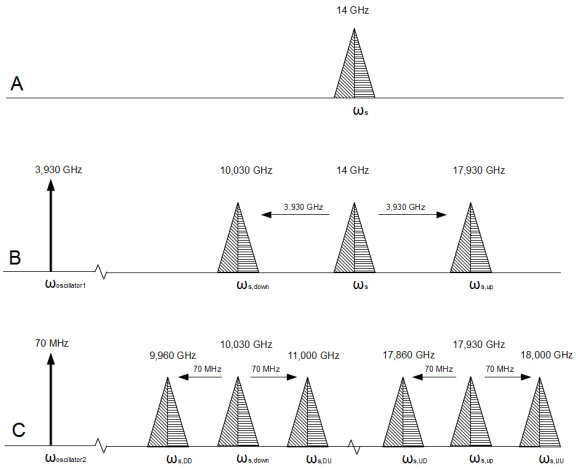


Figure 4 Spectrum transformations of the dual-mixer satellite simulator.

Modified single-mixer architecture: A modified single mixer satsim scheme is shown in figure 5 [2][4]. Using a power combiner the two oscillator signals, as shown in the dual-mixer concept, are combined and used to drive the mixer at the same time. The advantage of this architecture is the symmetrical delay path between the receive and transmit antennas. The frequency transformation is realized by utilizing the non-linear voltage-current characteristic of a diode.

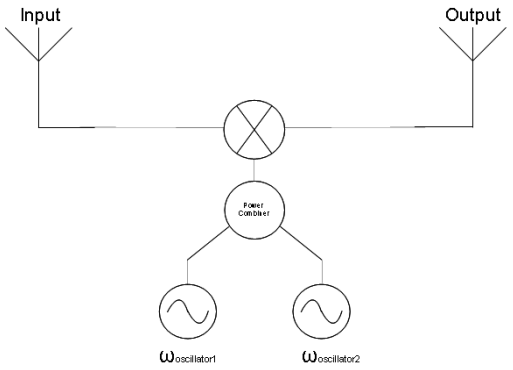


Figure 5 Modified single-mixer concept

Dual-stage quad-mixer architecture: The new proposed concept of satellite simulator is shown in figure 7. To make the satellite simulator delay measurements less sensitive for environmental influences, it is partly moved from the rooftop to the TWSTFT up- & down- conversion room. The temperature and humidity are well controlled in this room. Only a coaxial transfer switch 'M' remains mounted in front of the TWSTFT antenna, as it serves to set the reference point for transmit and receive delay components. The dual-stage quad-mixer architecture relies on four mixers for its multidirectional transceiver capability. Mixer *1A* provides down conversion from 14 GHz to 1.5 GHz. Mixer *1B* provides either up conversion from 1.5 GHz to 11 GHz or down conversion from 14 GHz to 1.5 GHz. Both mixers use the same oscillator signal and no frequency shifting is required. Similarly mixer *2A* provides down conversion from 1.5 GHz to 70 MHz and mixer *2B* provides either up conversion or down conversion between 1.5 GHz and 70 MHz. Mixer 2A and 2B also use the same oscillator signal.

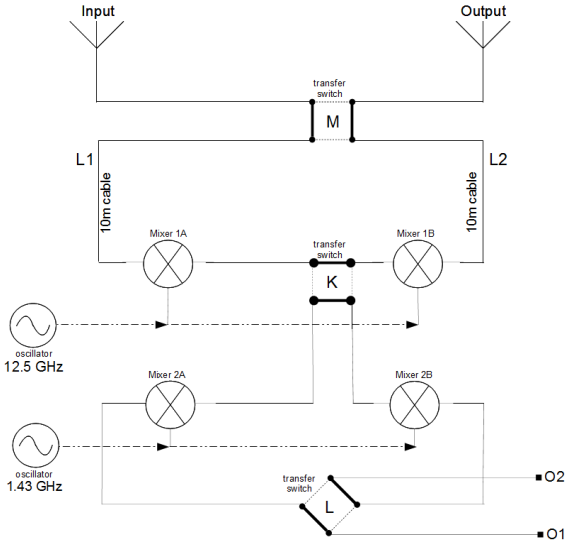


Figure 6 Dual-stage quad-mixer satellite simulator

Environmental sensitivities and biasing : To achieve a sub nanosecond accuracy for TWSTFT station measurements, sensitivity of the group delay towards environmental influences for the equipment used in TWSTFT station should be addressed adequately. However, almost no information is available for the equipment used in TWSTFT station and only a few studies have addressed this issue with significance.

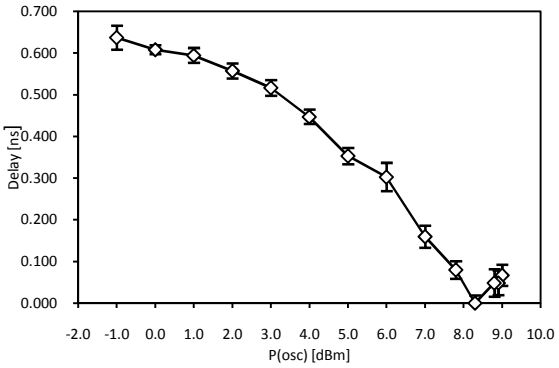


Figure 7 Mixer delay for varying oscillator power level at room temperature [22.0 ± 1.0 C]. Each measurement was performed for at least one day, with 12 measurement cycles per day.

The temperature coefficient for devices used in a TWSTFT station is typically less than 20 ps [7]. Diode-ring based mixers tend to be one of the most sensitive devices used in TWSTFT station. Another import aspect when utilizing mixers is the stability of the oscillator signal. In figure 7 the power level of the oscillator signal was varied for a double balanced diode-ring mixer. The delay decreases for higher power levels. This is caused by the larger voltage generated across the diodes. In this

region of biasing the mixer also exhibits higher group delay sensitivity towards oscillator power level variations. For lower power levels the magnitude of the group delay component is somewhat larger, but less sensitive to power level variations. Biasing the mixer at lower power level also increases noise level, as the conversion loss would be somewhat larger. Hence a tradeoff is to be made between system noise figure and group delay sensitivity to oscillator signal power level. Depending on the location of the mixer, stability of the environmental parameters, a proper setting of the biasing power level can be selected.

Automated delay calibration in a TWSTFT station: The TWSTFT station at VSL is located at three separate locations. Time scale generation, signal modulation and up conversion to 70 MHz intermediate frequency (IF) is performed in the Time & Frequency Laboratory. The time scale is generated using high performance cesium standards. For modulation and IF up conversion a SATRE modem is used. The 70 MHz signal is then sent to the TWSTFT up- & down- conversion room through 100 meter coaxial cables. In this room the required up- and down- conversion is performed to the transmit and receive microwave frequencies and sent to the rooftop. Finally the signals are amplified to the appropriate power level using a transmission amplifier and receiver LNA (Low Noise Amplifier) mounted on the antenna. See figure 8 for a layout of the present TWSTFT station at VSL.

As the TWSTFT station is located at three separate locations, three sets of delay measurements are performed. The first set of measurements is performed to determine the group delay components up to the TWSTFT up- & down- conversion room. In order to measure the delay component of the A and B cable, an additional third cable, *C*, is introduced in the system, shown in figure 8. This set of three cables allows three different loop delay measurements. From these measurements, the delay of the individual cables can be derived using linear algebra. As no frequency transformation takes place within these paths and all devices involved are located within the building, very stable delay measurements are achieved.

The second set of measurements is performed to determine the delay components of the quad-mixer satellite simulator itself, including delay paths up to the transfer switch at the TWSTFT antenna. In the case, the O1 and O2 ports in figure 6 are connected to the Rx and Tx paths of the TWSTFT modem.

The third set of measurements is performed to determine the delay of the TWSTFT transmit and receive paths respectively. The measurements are performed with the satellite simulator receive and transmit ports, hence the results from second set of measurement are required.

Meas.	Components of the measured delays
AB	$T_{switch-1} + T_{switch-2} + T_{cable-A} + T_{switch-A} + T_{pc} + T_{switch-B} T_{cable-B}$
CA	$T_{cable-A} + T_{switch-C} + T_{pc} + T_{switch-B} + T_{cable-B}$
CB	$T_{cable-C} + T_{amplifier} + T_{switch-C} + T_{pc} + T_{switch-A} + T_{cable-A} + T_{switch-2}$
M1	$T_{switch-1} + T_{switch-2} + T_{cable-A} + T_{switch-A} + T_{up-con.} + T_{cable-3} + T_{Tx-amplifier} + T_{cable-1} + T_{Satsim,mix1A} + T_{switch-K} + T_{Satsim,mix1B} + T_{cable-2} + T_{Rx-LNA} + T_{cable-4} + T_{down-con.} + T_{switch-B} + T_{cable-B} + T_{switch-2}$
M2	$T_{switch-1} + T_{switch-2} + T_{cable-A} + T_{switch-A} + T_{up-con.} + T_{cable-3} + T_{Tx-amplifier} + T_{cable-1} + T_{Satsim,mix1A} + T_{switch-K} + T_{Satsim,mix2A} + T_{switch-L} + T_{switch-D} + T_{switch-C} + T_{pc} + T_{switch-B} + T_{cable-B} + T_{switch-2}$
M3	$T_{switch-1} + T_{switch-2} + T_{cable-A} + T_{switch-A} + T_{up-con.} + T_{cable-3} + T_{Tx-amplifier} + T_{cable-1} + T_{Satsim,mix1B} + T_{switch-K} + T_{Satsim,mix2B} + T_{switch-L} + T_{switch-D} + T_{switch-C} + T_{pc} + T_{switch-B} + T_{cable-B} + T_{switch-2}$
M4	$T_{switch-1} + T_{cable-C} + T_{Amplifier} + T_{switch-C} + T_{switch-L} + T_{Satsim,mix2B} + T_{switch-K} + T_{Satsim,mix1B} + T_{cable-2} + T_{Rx-LNA} + T_{cable-4} + T_{down-con.} + T_{switch-B} + T_{cable-B} + T_{switch-2}$
M5	$T_{switch-1} + T_{cable-C} + T_{Amplifier} + T_{switch-C} + T_{switch-L} + T_{Satsim,mix2B} + T_{switch-K} + T_{Satsim,mix2A} + T_{switch-L} + T_{switch-D} + T_{switch-C} + T_{pc} + T_{switch-B} + T_{cable-B} + T_{switch-2}$
M6	$T_{switch-1} + T_{cable-C} + T_{Amplifier} + T_{switch-C} + T_{switch-L} + T_{Satsim,mix2B} + T_{switch-K} + T_{Satsim,mix1A} + T_{cable-1} + T_{cable-2} + T_{Satsim,mix1B} + T_{switch-K} + T_{Satsim,mix2A} + T_{switch-L} + T_{switch-D} + T_{switch-C} + T_{pc} + T_{switch-B} + T_{cable-B} + T_{switch-2}$

Table 1 Components of the measured delay paths

In table 1 components of the all measurements are shown. The final receive and transmit delay components

are calculated in equations 5 and 6. For simplification A, B and C cable delays are excluded from both equations.

$$Rx = (M2 + M3 - M6)/2 \quad \textbf{(5)}$$

$$Tx = M1 - M6 - Rx \quad \textbf{(6)}$$

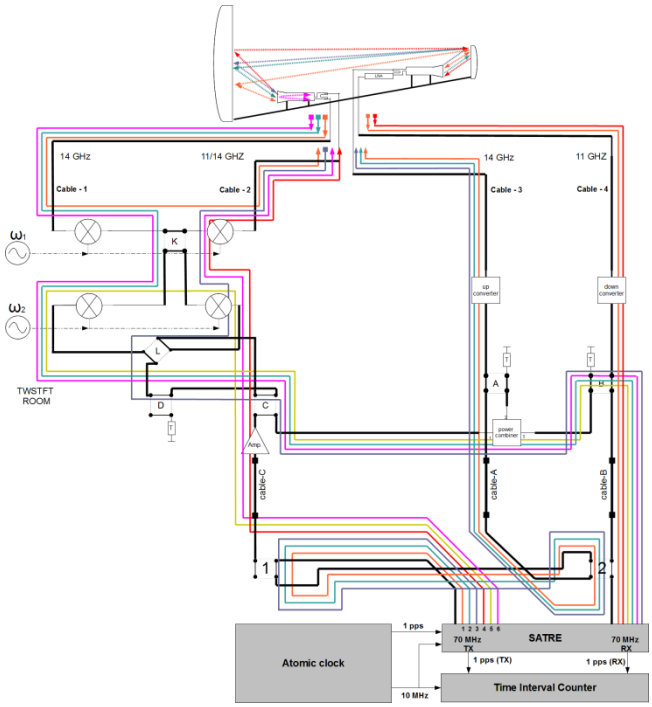


Figure 8 Diagram of the TWSTFT station with the automated delay calibration system included. The AB, CB and CA delay paths are not shown.

Meas.	Measurement path			
	Cable-3	Cable-4	Cable-2	Cable-1
M1	14 GHz	11 GHz	11 GHz	14 GHz
M2	14 GHz	-	-	14 GHz
M3	14 GHz	-	14 GHz	-
M4	-	11 GHz	11 GHz	-
M5	-	-	-	-
M6	-	-	14 GHz	14 GHz

Table 2 New TWSTFT delay measurements

Design & Measurement results: In figure 9 the final prototype of the satellite simulator is shown. Besides up- & down-convertors all hardware required to perform the TWSFTF station delay measurements is integrated into a single unit. Only transmission amplifier, the LNA and switch-M remain at the rooftop. External signal generator are required to bias the satsim mixers. The transfer switches are through RS485 interface.

Figure 9 Final prototype of the new satellite simulator switch box designed and assembled at VSL.

The TWSTFT station delay measurements are scheduled for twelve measurement sessions per day (once every two hours). In figure 10 the measurement results for AB, CB, and CA delay paths are shown.

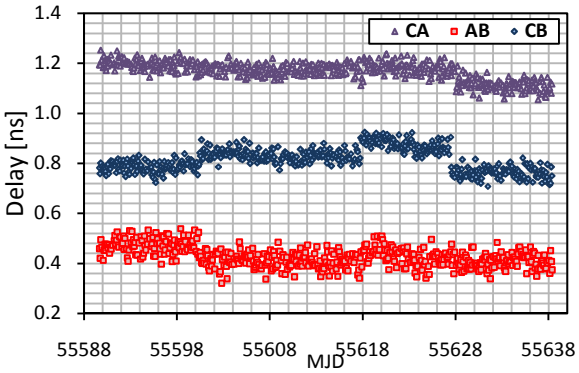


Figure 10 Measurement results of AB, AC, and BC paths. The delay values of the measurements are adjusted to be able to fit the data on a single graph. Using three different path measurements, delay values of A, B and C cables can be calculated using linear algebra.

The total fluctuation in time for the AB, CB, and CA delay paths does not exceed 200 ps over the whole measurement period (after correction of the offsets). In figure 11 the time stability of the AB delay path is shown. Time stability of 10 ps/s is achieved for averaging time of 1d. The 10 ps limitation is set by the noise floor of the modem.

After installing the new satsim unit a series of measurements were performed to test the functionality of the satsim unit. The measurements of the delay component M1-M3 are shown in figure 12.

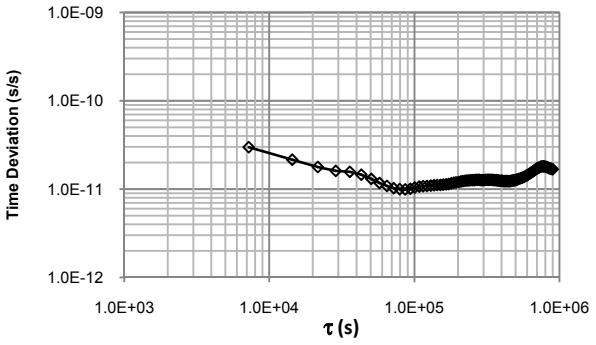


Figure 11 Time stability of AB path measurements

These measurement results are not to be takes as final values achievable by the delay measurement system, as every delay path is yet to be optimized.

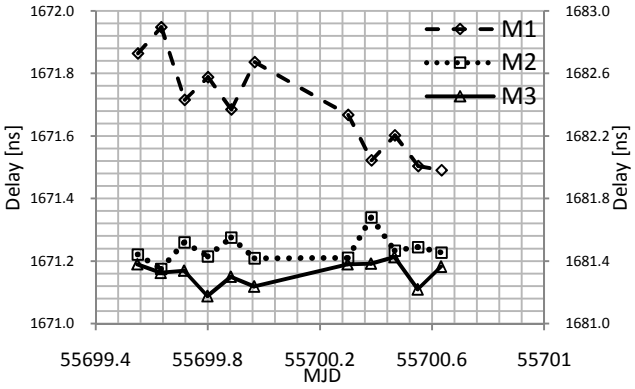


Figure 12 Measurement results of M1-M3, scheduled for twelve sessions per day (once every two hours). M1 results are displayed on the left vertical axis and M2-M3 results are displayed on the right vertical axis.

As the delay paths of M2 and M3 are almost identical, the delay components were expected to be equivalent as well. This is noticeable in figure 12 for M2 and M3. Measurement results of M1 seem to be drifting. However, the measurement period is rather short to draw concrete conclusions regarding the stability. In figure 13 the influence of C/N-ratio on the stability of delay measurement is shown.

Conclusions: A new dual-stage quad-mixer satellite simulator has been designed and assembled. With this satellite simulator all delays in the TWSTFT station can be determined in absolute terms. The implementation of the satellite simulator in the TWSTFT station is such that the delay calibration measurements are completely automated.

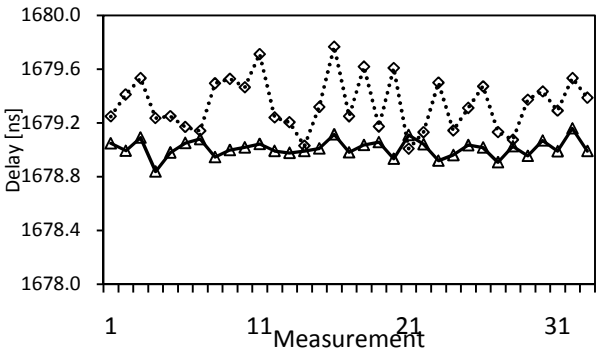


Figure 13 Measurement results for M1 for C/N-ratio of 56 dB/Hz [dotted line] and C/N-ratio of 62 dB/Hz [solid line].

Furthermore, the new satellite simulator was designed in a way that the sensitive components can be placed in-doors in a controlled environment.

Initial test measurements have started and are running 12 times per day (every two hours). The first results are quite promising, but it is too early to draw conclusions on the long term behavior. In future, for each measurement path an optimum transmission power level will be determined to achieve an optimal signal-to-noise (S/N) ratio.

Acknowledgements: Authors are greatly thankful to Henk van Ulden for helping with assembling the satellite simulator unit. Authors are also thankful to fellow colleagues for many useful discussions.

References

[1] Gerrit de Jong, et. All., "Prospects of Two-Way International Time Scale Comparison", EFTF, 1992, ESA-ESTEC, The Netherlands

[2] Gerrit de Jong, et. All., "Automated Delay Measurement System For An Earth Station for two-Way Satellite Time and Frequency Transfer", PTTI , 1994, Reston VA, USA

[3] Gerrit de Jong, et.all, *"Delay Stability Of The TWSTFT Earth Station At VSL,"* in Proceedings of the 29th Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 2-4 December 1997, Long Beach, California,

[4] Gerrit de Jong, et. All., "Evaluation And Improvements Of The Calibration Of A TWSTFT Station Using Satsim", 34th annual PTTI meeting, 2003.

[5] Joseph Achkar, "Design, Realization and Application of a Satellite Simulator in a TWSTFT Station", 40th European Microwave Conference, 2010

[6] ITU- 7/67-E, "The operational use of a two-way satellite time and frequency transfer employing PRN codes", 14th September 2009

[7] T.E.Parker, V.Zhang, "Sources of Instabilities in Two-Way Satellite Time Transfer", IEEE, 2005

Carrier-envelope phase dependent photoemission from a nanometric metal tip

Markus Schenk, Michael Krüger, and Peter Hommelhoff
Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany
Email: peter.hommelhoff@mpq.mpg.de

***Abstract*— We tightly focus $\sim 6\,\mathrm{fs}$ Ti:sa pulses onto a nanometric metal tip and record photoelectron spectra at laser intensities of about $\sim 4\times10^{11}\,\mathrm{W/cm^2}$. Here we present a strong carrier-envelope phase (CEP) dependent current modulation. In the high energy region of the spectra the CEP modulation rises up to 100%, meaning that the CEP can switch the photocurrent fully on and off. The metal tip resembles a record sensitive miniaturized CEP sensor with a precision exceeding about $15\,\mathrm{mrad}$ in $20\,\mathrm{s}$ measurement time. Even CEP-stabilizing a Ti:sa-oscillator with the signal of the emitted photoelectrons of a tip seems feasible.**

I. Introduction

The versatile use of femtosecond laser frequency combs relies on the possibility of CEP-stabilizing the pulses. Typically this is done via f-2f self-referencing the comb [1], [2]. Experimental effort, comparably high costs and a complicated setup are needed, especially if the CEP is kept konstant from pulse to pulse. A potential alternative approach is discussed in t paper. As already proposed and discussed in the IEEE Proceedings of the same conference series by one of the authors and former colleagues in 2006 [3], we report here on the experimental observation of a strong CEP modulation in the photocurrent. The pulse energy used is about $240\,\mathrm{pJ}$. Such a low pulse energy suffices because of optical field enhancement: boundary effects (small tip radius) and collective electronic excitations in the metal lead here to a up to six-fold increase in laser electric field strength (mimicking a factor of 36 higher laser intensity) [4].

The paper is organized as follows: In Section II we mention results previously achieved in photoemission experiments from sharp metal tips and on CEP effects in photoemission from metals. Section III describes the experimental setup, which is followed by Section IV describing the main results of this paper. Section V concludes the article by a discussion of tip-based carrier-envelope (CE) frequency ($\mathrm{f_{CEO}}$) stabilization of the oscillator.

II. Photoemission Studies - Related Work

Nonlinear photoemission from planar metals has been studied in detail in various papers (for a review see [6]). CEP effects in nonlinear photoemission are theoretically predicted [7], [8] but experimentally hard to observe [9], [10]: In Ref. [9] record-short $4\,\mathrm{fs}$ IR laser pulses were focused on a gold cathode leading to a minuscule CEP modulation in the total current. The difficulty in observing these effects is presumably caused by two effects. First, the emission area in photoemission from a planar surface is typically on the order of the focal spot

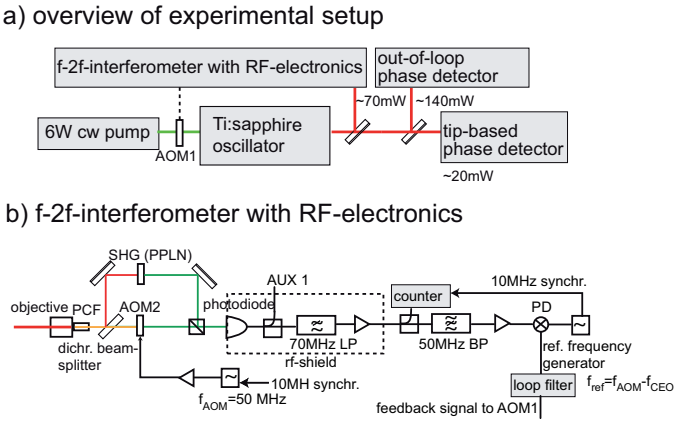


Fig. 1. Schematic experimental setup and CEP-stabilization scheme. a) A large fraction of the power of the Ti:sa laser is needed for phase-stabilization (70 mW) and out-of-loop monitoring of the phase (140 mW). For the photoemission only a small fraction of light is needed ($\sim 20\,\mathrm{mW}$). b) Detailed sketch of the CEP-stabilization setup. In the optical setup (left) the pulses are spectrally broadened to span a full octave in a photonic crystal fiber (PCF). Long and short wavelength regions are seperated by a dichroic beamsplitter. In the long-wavelength arm the second harmonic generation (SHG) is achieved with a periodically poled lithium niobate (PPLN) crystal. Both green beams are superimposed onto a photodiode. The photodiode signal is then RF-filtered and amplified in a RF-shielded box. This signal and a reference frequency are superimposed in a phase detector (PD) yielding the error signal. The feedback signal is generated and fed into the acousto-optic modulator (AOM1) in a). The offset-frequency and phase are set via this reference frequency generator. A 10 MHz signal is distributed amongst all components to provide a common clock. LP and BP stands for low pass and band pass filter. AOM2 shifts the frequency in one arm of the interferometer. This way $\mathrm{f_{CEO}}$ can be stabilized to 0 without being sensitive to DC drifts.

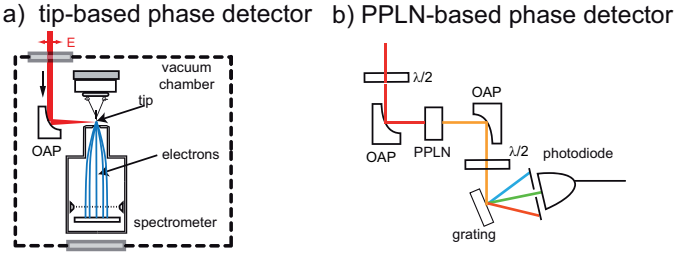


Fig. 2. The two out-of-loop CEP detectors. a) The pulses are focused onto the tip with an off-axis-parabolic mirror (OAP). The photoemitted electrons are then spectrally analyzed and detected. b) PPLN-based phase-detector [5]. The light is spectrally broadened and frequency doubled inside the PPLN. The f-2f-offset beat is then detected in the green spectral region with a photodiode.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

size. The observed signal is therefore usually integrated over the spatial intensity profile. This intensity averaging may lead to blurring of the electron spectra and hence washing out of CEP effects. Second, on a surface there are typically several hot spots with varying field enhancement that contribute to electron emission and may also obscure CEP effects. One way of circumventing those difficulties is to confine the emission area to a single well-controlled hot spot. This is the basic idea of using sharp metal tips, where the very localized optical field enhancement leads easily to an emission site area down to several $100\,\mathrm{nm}^2$.

III. Experimental Setup

A. Laser and phase stabilization

The setup consists of the laser with CEP stabilization and a tip inside a vacuum chamber (see Fig. 1 a)). The laser is a Kerr-lens mode-locked Ti:sapphire laser with $80\,\mathrm{MHz}$ repetition rate at a center wavelength of $\sim 800\,\mathrm{nm}$ and pulse duration of ∼6-fs. The CEP stabilization is achieved with a homebuilt f-2f-interferometer. The feedback mechanism is pump power modulation of the cw pump laser via an acousto-optic modulator (AOM1) [1]. Note that with another acousto-optic modulator in the f-interferometer arm (AOM2), it is possible to lock $\mathrm{f_{CEO}}$ to zero, meaning that CEP is constant. We change the phase in discrete steps by changing the phase of the reference frequency generator. We can monitor the CEP change in the PPLN-based out-of-loop interferometer (see Fig. 2b) [5]).

B. Tip environment

The nanometric metal tip is an electrochemically etched tungsten W(310) field emitter with a radius of curvature of about 10-20 nm. It is placed inside an ultrahigh vacuum chamber (base pressure of $\sim 3\times10^{-8}\,\mathrm{Pa}$)(see Fig. 2a)). The laser beam is tightly focused on the apex of the tip with a spot size of $\sim 2.4\,\mu\mathrm{m}$ ($1/e^2$ intensity radius). A small negative voltage is applied to the tip resulting in a maximum d.c. static electric field of $\sim 0.4\,\mathrm{GV/m}$ at the tip's apex. The emitted electrons are spectrally analyzed with a retarding field spectrometer (for details see [11]). Typically not more than one electron per pulse is emitted to avoid space charge effects.

IV. Carrier-envelope phase dependent photoemission

We record electron energy spectra as shown in Fig 3a) (blue solid curve) by focusing light pulses onto the tip. This spectrum is averaged over CE-phases. Two main observations can be made here. First, peaks are visible and separated roughly by the photon energy ($\sim 1.6\,\mathrm{eV}$). These represent above-threshold photoemission peaks [11]. Second, next to the dominant main contribution a plateau-like high-energetic region is found between 5 and 15eV. It is roughly two orders of magnitude smaller in count rate than the low-energetic part. It ends with a cut-off, i.e. a rather sharp decrease in count rate, at about $\sim 15\,\mathrm{eV}$. The green data points in Fig. 3a) indicate the CEP-dependent current modulation depth at the respective

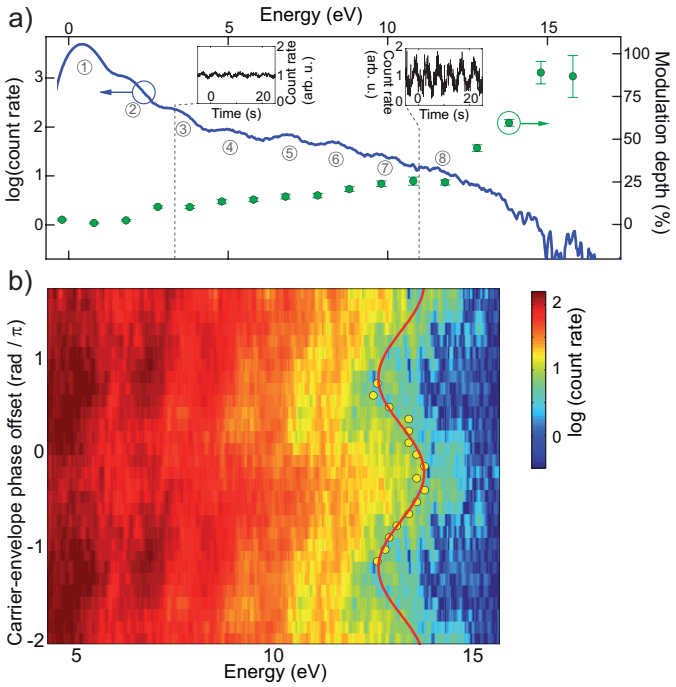


Fig. 3. Electron spectra and observed CEP effects. a) Phase-averaged measured electron spectra (blue line) in a logarithmic plot. A plateau-like structure is visible from about $5-15\,\mathrm{eV}$. The insets show the C-E phase modulation in the photocurrent with the spectrometer acting as energy high-pass filter at 3 eV and 11 eV (with the C-E offset frequency set to $f_{\mathrm{CEO}} = \dot{\phi}_{\mathrm{CE}}/2\pi = 0.2\,\mathrm{Hz}$). The CEP-modulation depth (green points) rises from few percent in total current up to near 100% for the high-energetic electrons (see Ref. [12] for details). b) Color map plot of the CEP dependent spectra in the high energy region. Clearly a modulation in the cut-off position is visible (red line). Measured data range spans 2π and is extended to $\sim 4\pi$ for better visibility.

energy. In this contribution we focus our attention on the high-energy electrons, since the CEP effects are most pronounced in that energy region. Fig. 3b) shows electron spectra in this high energy region for varying CE-phase. The position of the aforementioned cut-off is strongly modulated with the CEP (red line is sinusoidal fit in Fig. 3.). The amplitude of the cut-off modulation is $\sim 1.3\,\mathrm{eV}$. The CE-phase for the highest cut-off energies is at -0.22π.

We modeled these observation theoretically using two approaches. In the first model we numerically integrated the time-dependent Schrödinger equation. The basis for this is a rather simple 1-D model of the potential as described in Ref. [12]. We chose to calibrate the experimental CE phase offset according to the behaviour of the cut-off in the TDSE model. CE-phase of 0 means that the pulse is cosine-like, with the electric field pointing away from the tip surface. It is noteworthy that the phase of the driving field is expected to be phaseshifted with the enhanced near-field at the tip surface due to the collective excitation response of the metal [8], [13]. The second model employs the Simple-Man's model. Both models give good agreement with the experimental parameters as pulse length and laser intensity [12]. We attribute the occurence of the observed high-energetic electrons to rescattering taking

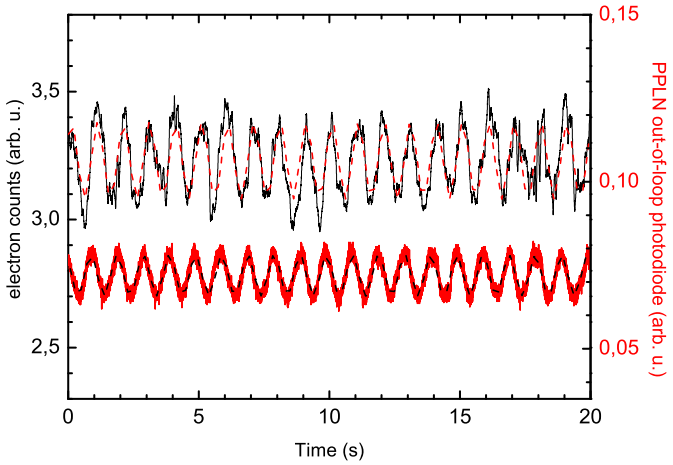


Fig. 4. Comparison between tip-based CEP-signal and PPLN-based signal for an offset-frequency set to $f_{\mathrm{CEO}} = 1\,\mathrm{Hz}$. The lower curve (red) is the PPLN based photodiode signal, the upper (black) the energy-filtered photoemission current ($E >\sim 6\,eV$). The dashed curves are sinusoidal fits. From the fit error one can extract a phase error of about $15\,\mathrm{mrad}$ for the photoemission signal (for $20\,\mathrm{s}$ measurement time).

place at the metal surface.

Next, in a crude first study we study the sensitivity of the tip-based CEP detector. For this we set $f_{\mathrm{CEO}} = 1\,\mathrm{Hz}$ and record signal of the high-energetic electrons ($E >\sim 6\,\mathrm{eV}$, see Fig. 4). From sinusoidal fits we can extract the precision in CEP determination of about $15\,\mathrm{mrad}$ for the photoemission signal in $20\,\mathrm{s}$ measurement time.

V. Outlook

The tip already represents a solid-state based CEP detector [14]. This contribution represents therefore the realisation of the idea proposed in the proceedings of the same conference [3]; however there are more possible applications: As an outlook one might ask whether it is feasible to stabilize the offset-frequency of the Ti:sa oscillator using the CEP-modulation in the electron current as feedback. Under the assumption that a signal-to-noise power ratio of $30\,\mathrm{dB}$ at $10\,\mathrm{kHz}$ resolution bandwidth is needed and assuming that the detection may be shot-noise limited we arrive, after a short calculation, at a minimum electron count rate of about $10^7\mathrm{s}^{-1}$. The maximum observed CEP-modulation amplitude is still 3 orders of magnitude smaller.

Several means to bridge this apparently large gap might be taken. Ideas include improving the collection efficiency of the emitted electrons, or using higher repetition rate Ti:sa oscillators [15], trying different tip materials or many emitters in an array.

References

[1] D. J. Jones, S. A. Diddams, J. K. Ranka, A. Stentz, R. S. Windeler, J. L. Hall, and S. T. Cundiff, "Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis," *Science*, vol. 288, no. 5466, pp. 635–639, Apr 28 2000.

[2] R. Holzwarth, T. Udem, T. W. Hänsch, J. C. Knight, W. J. Wadsworth, and P. S. J. Russell, "Optical frequency synthesizer for precision spectroscopy," *Phys. Rev. Lett.*, vol. 85, pp. 2264–2267, 2000.

[3] P. Hommelhoff, C. Kealhofer, and M. A. Kasevich, "Femtosecond laser meets field emission tip – a sensor for the carrier envelope phase?" in *Proceedings of the 2006 IEEE International Frequency Control Symposium and Expositions, Vols 1 and 2*, 2006, pp. 470–474.

[4] L. Novotny, R. X. Bian, and X. S. Xie, "Theory of nanometric optical tweezers," *Phys. Rev. Lett.*, vol. 79, p. 645, 1997.

[5] T. Fuji, J. Rauschenberger, A. Apolonski, V. S. Yakovlev, G. Tempea, T. Udem, C. Gohle, T. W. Hänsch, W. Lehnert, M. Scherer, and F. Krausz, "Monolithic carrier-envelope phase-stabilization scheme," *Optics Lett.*, vol. 30, pp. 332–334, 2005.

[6] G. Ferrini, F. Banfi, C. Giannetti, and F. Parmigiani, "Non-linear electron photoemission from metals with ultrashort pulses," *Nuclear Instruments & Methods in Physics Research A*, vol. 601, no. 1–2, pp. 123–131, 2009.

[7] C. Lemell, X.-M. Tong, F. Krausz, and J. Burgdörfer, "Electron emission from metal surfaces by ultrashort pulses: Determination of the carrier-envelope phase," *Phys. Rev. Lett.*, vol. 90, no. 7, p. 076403, Feb 2003.

[8] M. I. Stockman and P. Hewageegana, "Absolute phase effect in ultrafast optical responses of metal nanostructures," *Appl. Phys. A*, vol. 89, pp. 247–250, 2007.

[9] A. Apolonski, P. Dombi, G. G. Paulus, M. Kakehata, R. Holzwarth, T. Udem, C. Lemell, K. Torizuka, J. Burgdörfer, T. W. Hänsch, and F. Krausz, "Observation of light-phase-sensitive photoemission from a metal," *Phys. Rev. Lett.*, vol. 92, no. 7, p. 073902, Feb 2004.

[10] P. Racz, S. Irvine, M. Lenner, A. Mitrofanov, A. Baltuska, A. Elezzabi, and P. Dombi, "Strong-field plasmonic electron acceleration with few-cycle, phase-stabilized laser pulses," *Applied Physics Letters*, vol. 98, no. 11, 2011.

[11] M. Schenk, M. Krüger, and P. Hommelhoff, "Strong-field above-threshold photoemission from sharp metal tips," *Phys. Rev. Lett.*, vol. 105, no. 25, p. 257601, Dec 2010.

[12] M. Krüger, M. Schenk, and P. Hommelhoff, "Attosecond control of electrons emitted from a nanoscale metal tip," *to appear in Nature*, 2011.

[13] S. Kim, J. Jin, Y. Kim, I. Park, Y. Kim, and S. Kim, "High-harmonic generation by resonant plasmon field enhancement," *Nature*, vol. 453, no. 7196, pp. 757–760, 2008.

[14] P. Hommelhoff, M. Schenk, and M. Krüger, "Patent pending," 2010.

[15] A. Bartels, D. Heinecke, and S. A. Diddams, "10-GHz self-referenced optical frequency comb," *Science*, vol. 326, no. 5953, pp. 681–681, 2009.

Surface Functionalization and Monolayer Formation on Silicon Resonant Nanoballances

Babak Tousifar
Department of Mechanical and Material Engineering
University of Denver
Denver, USA

Siavash Pourkamali
Department of Electrical and Computer Engineering
University of Denver
Denver, USA

Miroslav kvasnica, Byron Purse
Department of Chemistry and Biochemistry
University of Denver
Denver, USA

Abstract— This work presents formation of molecular monolayers on silicon surfaces of micromechanical resonant devices and measurement of the mass of the deposited monolayer using such. Thermally actuated single crystalline silicon micromechanical resonators with resonant frequencies in the 20-30 MHz range were used as highly sensitive mass sensors. Monolayers of long chain amine molecules were formed by reacting the amine groups with epoxide groups covalently bonded to the silicon surfaces. As a result a consistent negative frequency shift of ~100ppm was measured for the resonators. Mass calculations based on the measured frequency shift show a 6.9% surface coverage for the monolayer. XPS surface analysis shows existence of nitrogen atoms on the resonator surface confirming that the measured frequency shifts are due to the addition of the amine molecules.

I. INTRODUCTION

Highly sensitive biosensors capable of detecting and measuring smallest quantities of biochemicals in biological samples are needed in various biotechnology and biomedical applications. Most of the commercially available biosensing and bioanalysis platforms, such as the biomedical diagnostics micro-arrays are based on optical methods requiring costly and sophisticated readout instrumentation with a very complex and time consuming procedure including fluorescent labeling, incubation and hybridization steps. Furthermore, such techniques only show the end results and do not provide any real-time information, e.g. about the reaction rates and effect of different parameters on it. For example measuring the affinity binding constant of the protein-protein reaction is desirable to study the disease biochemical pathways that can involve several millions of assays performed over several months [1]. Therefore, label free biomolecular sensors with fully electronic readout capable of real time monitoring of molecular absorption events can be highly beneficial in state of the art biomedical science and engineering.

Electromechanical resonant microbalances such as Quartz Crystal Microbalances (QCM) [2,3] or surface acoustic wave (SAW) sensors [4] are capable of label-free molecular detection by direct real-time monitoring of the added mass to their surfaces. However, such devices have relatively large sizes and cannot compete with fluorescent microarrays in terms of size, cost and throughput. In many cases their sensitivity is not adequate either. Micro-electromechanical resonators are smaller integrated versions of the same devices and can potentially lead to highly sensitive label-free sensor arrays with electronic readout capable of eliminating the need for amplification process such as PCR [5].

In order to use such devices as a biosensor capable of detection and concentration measurement of specific biomolecules of interest, the resonators surface is modified by attaching receptor molecules that can selectively absorb the target molecules. Upon exposure to the target molecules and absorption of such on the resonator surface, the added mass leads to an electronically measurable change in the resonance frequency of the resonator.

II. THERMAL PIEZO-RESISTIVE RESONATORS

Thermally actuated single crystalline silicon resonators with piezoresistive readout [6] have been used in this work as highly sensitive mass sensors. The simplicity and monolithic structure of such devices makes them very robust and durable and therefore suitable for sensory applications where the device needs to be in direct contact with the surroundings. Operation in liquid is generally a major challenge for

This work was supported by National Science Foundation (NSF) under grants MRI-RAPID #1061489 and MRI #0923518.

microscale resonators due to large surface to volume ratio and excessive viscous damping from the liquid. However, disk-shaped thermal-piezoresistive resonators can be designed for operation in liquid with relatively high quality factors [7]. Although all the measurements in this work have been performed in air after drying the resonators, the eventual goal is to perform direct analysis with resonators while immersed inside a biological solution.

Fig. 1 shows the COMSOL finite element thermal analysis of the resonator structures used in this work referred to as I-shaped resonators. Thermal actuation occurs by passing a current between the two pads on the two sides of the structure resulting in an ohmic power loss in the structure. The fluctuating power loss by passing a combination of AC and DC currents, results in a fluctuating temperature gradient and therefore fluctuating thermal expansion of the structure. Due to their higher electrical resistance, most of the heat gradient is generated in the narrow beams in the middle of the structure. The AC extensional force resulting from the fluctuating temperature in the beams can actuate the resonator in its in-plane extensional resonance mode. As the resonator vibrates, the alternating tensile and compressive stress in the pillars results in fluctuations in their electrical resistance (due to the piezo-resistive effect). This results in fluctuations in the DC current passing through the resonator that represents the vibration amplitude (output signal) of the resonator. (Fig.1.)[8].

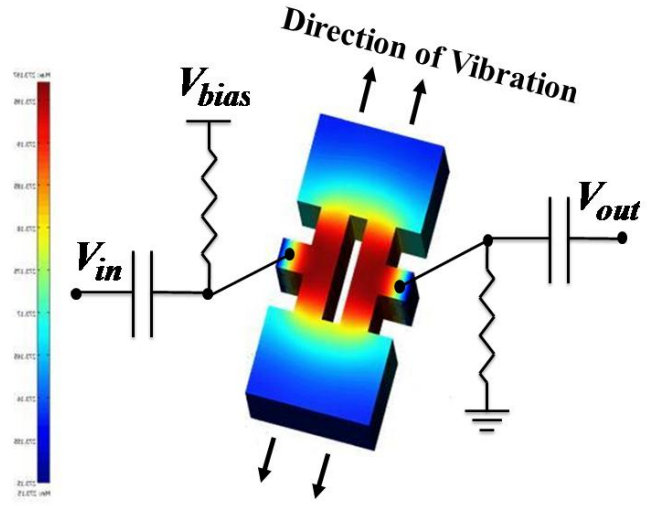


Figure 1. COMSOL finite element analysis of a thermally actuated I-shaped resonator showing the vibration direction and the qualitative distribution of AC temperature fluctuation amplitude (red being the maximum and blue minimum).

The resonators were fabricated on low resistivity N-type Silicon-On-Insulator (SOI) substrates. This was done using a single mask micromachining process by patterning the silicon device layer and undercutting the structures in hydrofluoric acid (Fig. 2a). Fig. 2b shows the SEM view of a fabricated array of resonators as well as zoomed in view of a single resonator.

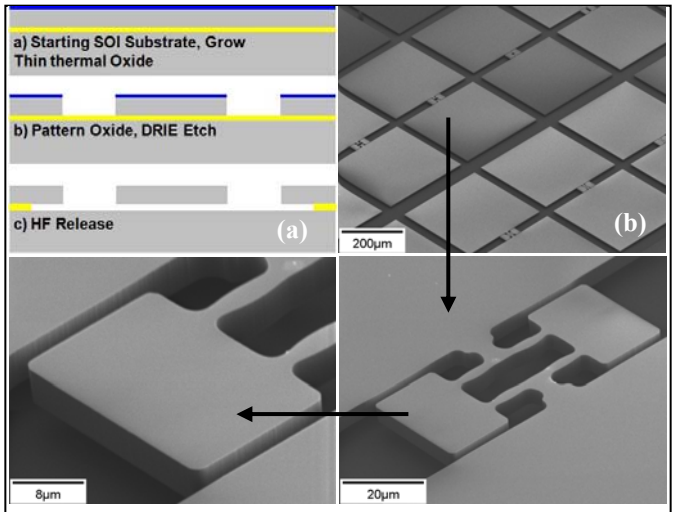


Figure 2. (a) Single-mask resonator microfabrication process. (b) SEM view of the micro fabricated resonators.

III. Experiments and Measurements

A. Surface Linkage Synthesys and Frequency Measurements

The goal is to form a stable chemical linkage, via covalent bonds, between the silicon surface and the amino groups present in octadecylamine molecules (an amine molecule with a straight chain of 18 carbon atoms), which is the target molecule in this work. For this purpose intermediate molecules containing epoxide groups, which are reactive towards amino groups, are to be attached to the silicon surface first. The previously fabricated device surfaces were activated by oxidation in 1N HNO_3 for 30-60 min followed by a rinse in DI water, acetone, and ethanol for 5 min each. In order to attach the epoxide groups to the silicon surface, the chips were soaked for 5 h at 80~ in a solution consisting of 24 parts xylene, 8 parts 3-glycidoxypropyltrimethoxysilane and 1 part N,N-diisopropylethylamine (Fig. 3a). Excess reagents were removed by three washes in acetone and ether. The first frequency measurement was performed in this step after drying the resonators.

After measuring and recording the resonator resonance frequencies with different bias currents, octadecylamine attachment was performed by immersing the chips in a 5mM solution of octadecylamine in ethanol overnight (Fig. 3b). The excess reagents on the chips were rinsed off in methanol and ether. Device resonance frequencies were measured again after drying the sample. Comparing the measured frequencies before and after octadecylamine attachment shows a clear and consistent frequency shift of ~100ppm for 3 out of the 4 tested resonators (Table 1, Fig 4). The 4rd resonator shows a larger frequency shift of 140ppm, which is suspected be due to addition of contaminants during the several chemical synthesis and rinsing steps, or absorption of airborne particles during exposure of the sample to air (e.g. during testing or transferring from one place to another).

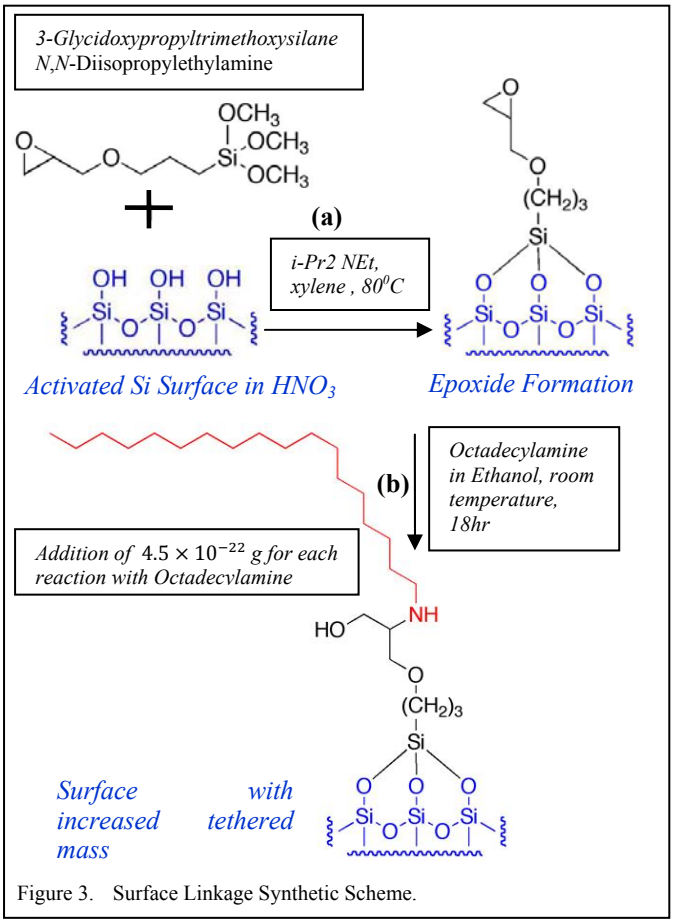


Figure 3. Surface Linkage Synthetic Scheme.

TABLE I. SUMMARY OF MEASUREMENT RESULTS FOR DIFFERENT FUNCTIONALIZED RESONATORS

L_1, L_2,W_1, W_2,W_3 (μm)	f_1(MHz) Functionalized	f_2(MHz) Octadecylamine	Δf (Hz)	Δf (ppm)
23,36,33,4,2	30.437427	30.434413	3014	99
23,36,33,4,2	29.891373	29.888322	3051	100
23,36,33,--, 2	23.809419	23.805919	3400	140
23,36,33,--, 2	27.001593	26.998914	2679	99

Figure 5 shows the measured frequencies for a 30.4MHz resonator before and after octadecylamine attachment at different bias currents showing a consistent shift at each bias current due to the added mass.

B. Surface Coverage Analysis

During the surface fictionalization, each crystal with a lattice constant of $5.43A^0$ consists of eight open bonds on 100 silicon surface exposed to reaction reagents (Fig. 6). Attachment of each epoxide group requires linking to three silicon atoms on the surface (fig. 8). This means that if all the available bonds are used for linking with 100% packing, 8.8 epoxide groups will be available in every nm^2 of silicon surface.

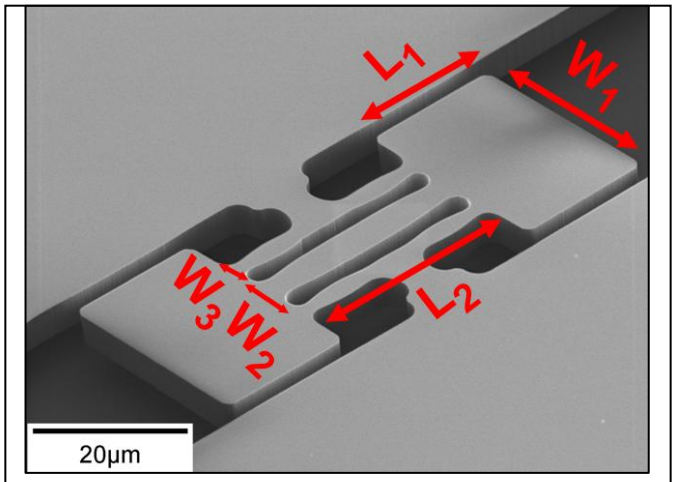


Figure 4. SEM view of a 3-bar I-shaped resonator with different resonator dimensions labeled on it.

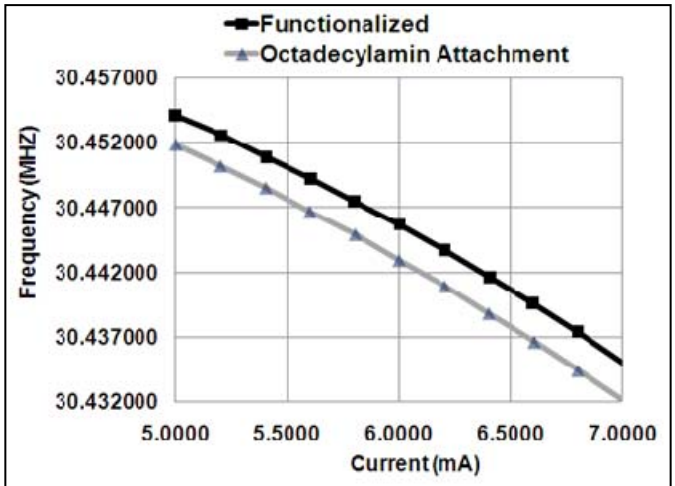


Figure 5. Consistent frequency shift for different resonator bias currents before and after octadecylamine attachment.

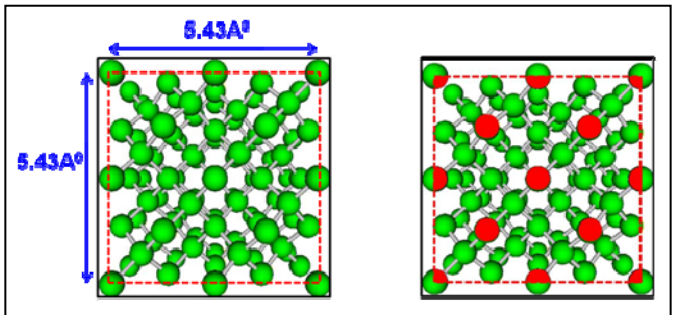


Figure 6. (a) Model of a silicon crystal as seen on the 100 surfaces, The area of every single silicon crystal is equal to $0.3nm^2$; (b) 8 silicon atoms on the surface with open bonds in every single silicon crystal unit available for surface linking .

On the other hand, according to the device mass, reference frequency value and frequency shift after octadecylamine attachment, the added mass for the 1st measured resonator with the resonance frequency of 30.4 MHz is: $\frac{\Delta f}{f} = -\frac{\Delta m}{m} \rightarrow \Delta m = 1.0pg$.

Therefore, considering octadecylamine molecular weight, the added 1.0pg of octadecylamine is equal to 2.2×10^{9} added molecules on the on the surface. Based on the resonator surface area of $3.6 \times 10^{-9} m^2$, we will end up having 0.61 molecules in every nm^2 which is equivalent to 6.93% of the calculated theoretical limit of the added molecules in $1nm^2$.

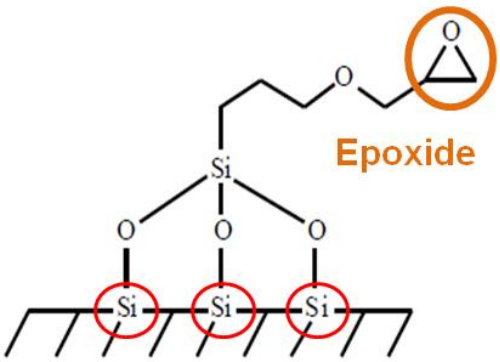
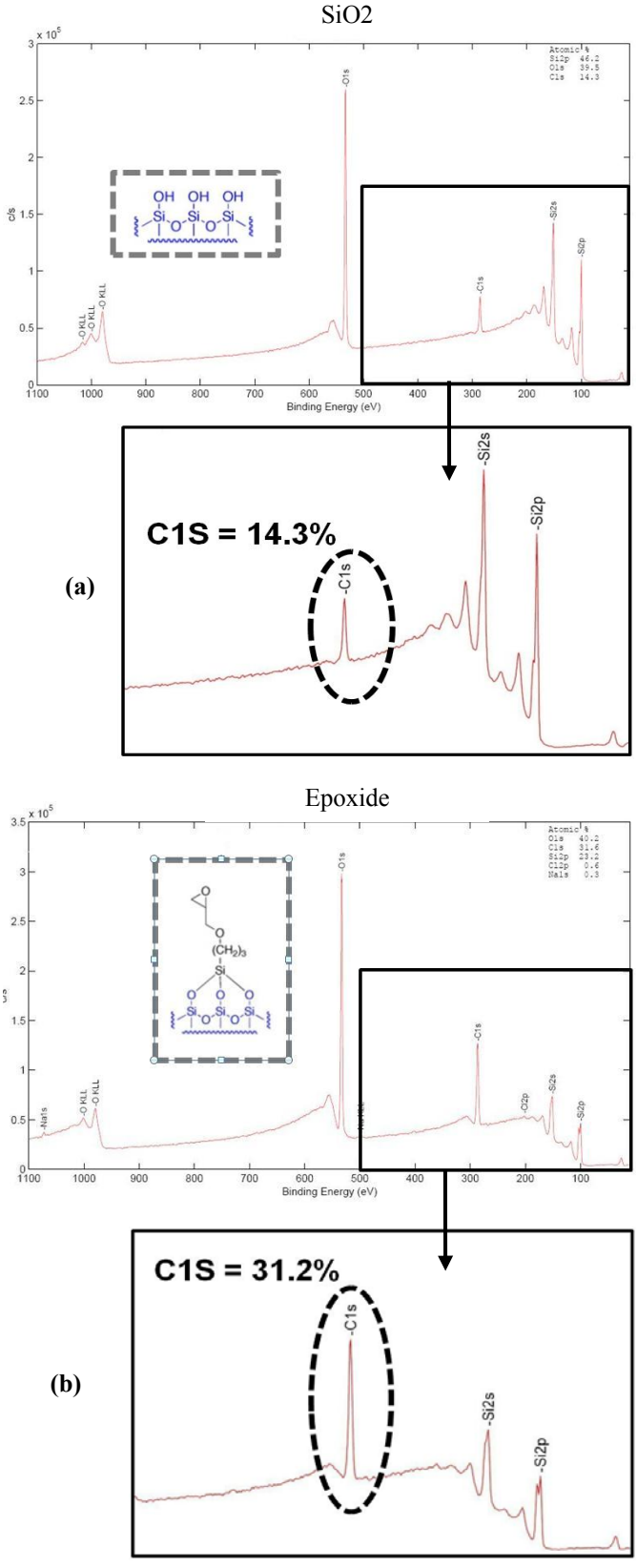


Figure 8. Each epoxide group links to the surface by bonding to three silicon open bonds on the surface

C. *Surface XPS Analysis*

To make sure that the surface linking has resulted in the attachment of the expected molecules and groups to the silicon surface and that the measured resonator frequency shifts are not a result of measurement errors or contamination, X-ray Photoelectron Spectroscopy (XPS) was performed on samples at different steps of the process. The elemental compositions on the surface support the addition of the desired molecules after each reaction. In silicon dioxide surface analysis a considerable amount of carbon was detected which is most likely due to the organic contaminants or residues (e.g. solvents used to clean the surface or airborne particles) (Fig. 9a). The carbon concentration on the surface increased significantly after the surface modification by linking the epoxide groups (Fig. 9b). After octadecylamine attachment a small amount of nitrogen (from the amine groups) could be detected on the surface while the carbon concentration further increased (Fig. 9c).



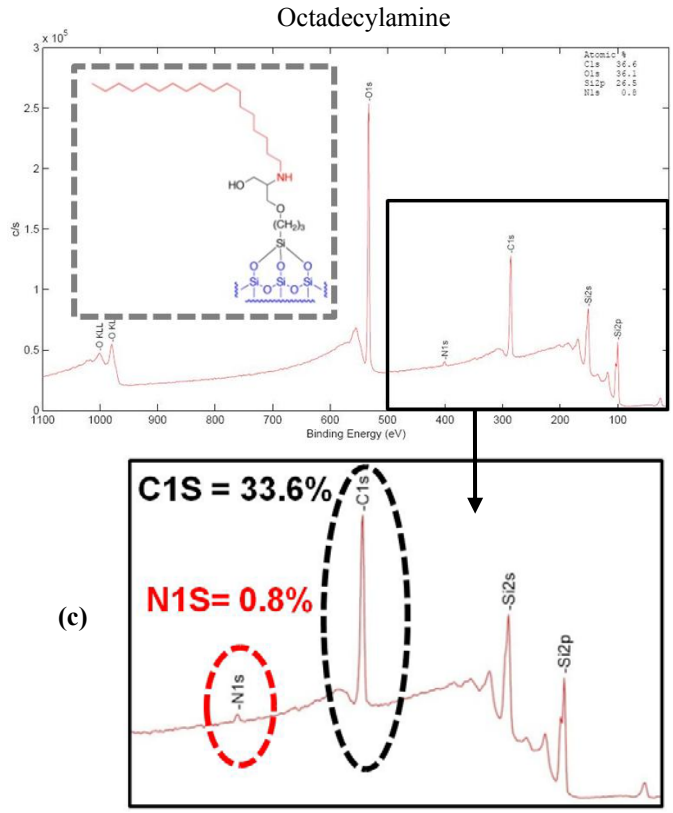


Figure 9. XPS analysis results of the resonators surfaces confirms addition of the carbon and nitrogen containing amine groups to the surface in different reaction steps. (a) SiO2, (b)Epoxide, (c) Octadecylamine

IV. Conclusions And Future Work

Thermally actuated micro-electromechanical resonators were functionalized successfully to form a monolayer of octadecylamine. Using the resonators as ultra-sensitive balances, the mass of the deposited molecular monolayer was measured by measuring resonant frequency of the resonators before and after addition of the amino-terminated carbon chains. A consistent negative frequency shift of ~100ppm after amine attachment shows a 6.9% surface coverage. Given that the attached carbon chains are relatively long (18 carbon atoms, ~2.5nm long), the calculated coverage is close the maximum achievable molecular density for such molecules.

The same surface linking chemistry can be utilized to attach amino terminated DNA single strands or other bio-receptors to the silicon surface to be used as molecular probes for detection of complementary target molecules.

Acknowledgment

Authors would like to thank Mr. Amir Rahafrooz for device fabrication and finite element analysis, staff at Georgia Tech nanotechnology research center, Dr. Daniel Armentrout for his supports, and Colorado State University for providing XPS facility.

V. References

[1] E. Jacoby, R. Bouhelal, M. Gerspacher, K. Seuwen, ChemMedChem 1, 761 (2006).

[2] Avramov, I.D., "A 0-phase circuit for QCM-based measurements in highly viscous liquid environments, IEEE Sensors Journal, Vol. 5, Issue 3, June 2005, pp.425-432.

[3] Sang-Hun Lee; Stubbs, D.D.; Cairney, J.; Hunt, W.D., "Real-time detection of bacteria spores using a QCM based immunosensor," Proceedings of IEEE Sensors, 2003, Vol. 2, pp.1194–1198.

[4] Mannelli, I.; Minunni, M.; Tombelli, S.; Mascini, M., "Bulk acoustic wave affinity biosensor for genetically modified organisms detection," IEEE Sensors Journal, Vol. 3, Issue 4, Aug. 2003, pp.369-375.

[5] Wabuyele, M.B., et al., Approaching real-time molecular diagnostics: Singlepair fluorescence resonance energy transfer (spFRET) detection for the analysis of low abundant point mutations in K-ras oncogenes. Journal of the American Chemical Society,2003, 125(23): 6937–6945.

[6] A. Rahafrooz, and S. Pourkamali, "High frequency thermally actuated electromechanical resonators with piezoresistive readout," IEEE Transactions on Electron Devices, April 2011

[7] A. Rahafrooz, and S. Pourkamali, "Characterization of Rotational Mode Disk Resonator Quality Factors in Liquid," 2011 IEEE International Frequency Control Symposium (IFCS 2011)

[8] A. Rahafrooz, A. Hajjam, B. Tousifar, S. Pourkamali., "Thermal actuation, a suitable mechanism for high-frequency electromechanical resonators," Proceedings, IEEE MEMS conference, 2010, pp. 200-203.

Damping and Mass Sensitivity of Laterally Vibrating Resonant Microcantilevers in Viscous Liquid Media

Russell Cox[1], Jinjin Zhang[1]
Fabien Josse[1], Stephen M. Heinrich[2]
[1] Electrical and Computer Engineering
[2]Civil and Environmental Engineering
Marquette University
Milwaukee, USA
Fabien.Josse@marquette.edu

Isabelle Dufour
Université de Bordeaux
CNRS, IMS Laboratory
Talence, France

Luke A. Beardslee, Oliver Brand
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, USA

Abstract— **The effect of liquid viscosity and density on the characteristics of laterally excited microcantilevers is investigated and compared to transversely excited microcantilevers. When immersed into a viscous liquid medium such as water from air, the resonant frequency of laterally (in-plane) vibrating microcantilevers is shown to decrease by only 5-10% as compared to ~50% reduction for transversely (out-of-plane) vibrating microcantilevers. Furthermore, as the viscosity of the medium increases the resonant frequency of a laterally vibrating beam is shown to decrease at a slower rate than that of a transversely vibrating beam. The decreased viscous damping also leads to increases in the quality factor of the system by a factor of 4-5 compared to beams vibrating transversely. The mass sensitivities of laterally vibrating beams are also theoretically predicted to be roughly two orders of magnitude larger in water for some cantilever geometries. The increase in the quality factor and mass sensitivity indicate that operating in the in-plane flexural mode (lateral vibration) will decrease the limit of detection compared to operating in the more common out-of-plane flexural mode (transverse vibration). These improvements in device characteristics indicate that microcantilevers excited laterally are more suited for operating in media of high viscosities.**

I. INTRODUCTION

Microcantilevers are microelectromechanical devices that can be used as highly sensitive chemical sensing platforms. When coated with a chemically selective layer, shifts in the resonant frequency of these beams can be related to changes in the ambient concentration of specific analytes in the medium of operation. A change in mass in the range of picograms has been detected using these devices, with projected detection limits on the order of femtograms or attograms [1,2]. However, when these beams are operating in a viscous liquid medium as opposed to air, the increased fluid damping and displaced fluid mass cause the resonant frequency and the mass sensitivity of the device to decrease. The increased fluid damping also causes the quality factor to decrease drastically, which increases the frequency noise and the limit of detection.

It was previously proposed that exciting microcantilevers in the in-plane flexural mode (instead of the more common out-of-plane flexural mode) would decrease the amount of viscous damping and displaced fluid mass, making the device a more appropriate platform for liquid phase sensing applications [3]. In a previous work, a semi-analytical expression for the hydrodynamic forces acting on laterally vibrating beams was derived to analyze these in-plane flexural mode cantilevers [3]. The expression predicts values for the hydrodynamic forces that compare well with other results presented in the literature [4].

In the present study, this semi-analytical expression for the hydrodynamic forces will be used to predict the characteristics of laterally vibrating microcantilevers operating in media having higher viscosities than that of water (e.g., up to 20% aqueous glycerol solution). The characteristics of interest are the resonant frequency, the quality factor, and the mass sensitivity of the device. The predicted values for the resonant frequency and quality factor will be compared to experimentally obtained values, and the trends in these characteristics investigated. The predicted mass sensitivity for these devices will also be discussed, both as a function of the beam's geometry and as a function of the medium of operation. Finally, as the expressions for the characteristics of microcantilevers vibrating transversely are well-known, analytical expressions for the ratio of the characteristics of laterally and transversely vibrating beams of similar geometries will be determined. Such expressions will help to indicate the geometries and characteristics of interest for operation in the in-plane flexural mode.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

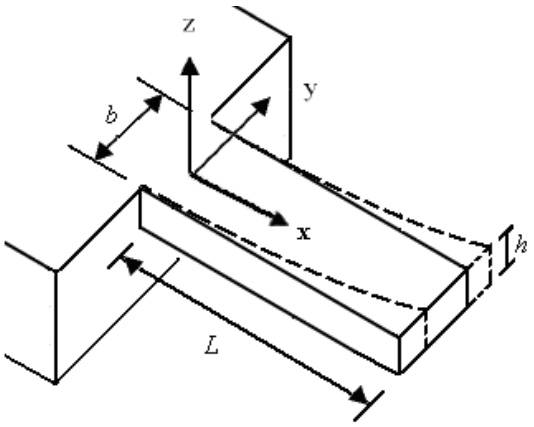


Figure 1: Geometry of a microcantilever of length *L,* width *b,* and thickness *h* deflecting laterally

II. Theoretical Analysis

A. Resonant Frequency

The geometry of a laterally excited microcantilever is shown in Fig. 1. The resonant frequency of a laterally vibrating beam in a viscous liquid medium was previously shown to be [3]

$$f_{res,lat,i} = \frac{\alpha_i^2}{2\pi}\sqrt{\frac{k_{lat}}{M_{lat}}} \tag{1}$$

where α_i is a constant dependent on the mode number ($\alpha_1 \cong 1.875$ for the fundamental in-plane flexural mode) and the effective spring constant, k_{lat}, and effective mass, M_{lat}, are given by

$$k_{lat} = \frac{Eb^3h}{12L^3}, \tag{1a}$$

$$M_{lat} = \left(\rho_B bhL + Lg_{2,lat}\right) + L\frac{\left(\left(g_{1,lat}/\omega_{res,lat}\right) + \left(\frac{\omega_{lat}}{2}\right)\frac{d}{d\omega}\left(g_{1,lat}/\omega\right)\right)}{\left(\rho_B bh + g_{2,lat} + \left(\frac{\omega_{lat}}{2}\right)\frac{d}{d\omega}\left(g_{2,lat}\right)\right)}\left(g_{1,lat}/\omega_{lat}\right) \tag{1b}$$

Here, ρ_B denotes the beam density and the variables $g_{1,lat}$ and $g_{2,lat}$ are related to the amount of viscous damping and the amount of displaced fluid mass, respectively, and are given by

$$g_{1,lat} = \frac{\pi}{4}\rho_L b^2 \Gamma_{lat,I}(\mathrm{Re}, h/b)\omega_{lat}, \tag{2}$$

$$g_{2,lat} = \frac{\pi}{4}\rho_L b^2 \Gamma_{lat,R}(\mathrm{Re}, h/b). \tag{3}$$

The terms $\Gamma_{lat,R}$ and $\Gamma_{lat,I}$ are the real and imaginary parts, respectively, of the hydrodynamic function, a normalized version of the hydrodynamic force defined in [3] that depends only on the aspect ratio *h/b* and the modified Reynolds number (Re) of the medium. The modified Reynolds number is a measure of the ratio of inertial forces to the viscous forces acting on the beam, and is defined as [5]

$$\mathrm{Re} = \frac{\rho_L \omega_{lat} b^2}{4\eta}, \tag{4}$$

where ρ_L and η are the mass density and dynamic viscosity of the fluid, respectively. It is assumed that the medium of operation is infinitely large and incompressible. It is also assumed that the velocity is continuous at the fluid-beam interface and the deflections of the beam are small compared to the dimensions of the beam. The beam's length is also assumed to be much greater than its width and thickness. Since the microcantilevers investigated in this study are composite beams made up of several layers, the Young's modulus, *E*, used in the theoretical predictions will be the effective Young's modulus

$$E_{effective} = 12\rho_B\left(\frac{2\pi L^2 f_0}{\alpha_i^2 b}\right)^2, \tag{5}$$

where f_0 is the resonant frequency in vacuum. In this work, we used the experimentally determined resonant frequency of the beam in air instead, which has approximately the same value as the resonant frequency in a vacuum. An effective beam density, ρ_B, of 2330 kg/m^3 will be assumed as the density of each layer is approximately that of silicon.

Immersed in a fluid, the resonant frequency of a laterally vibrating beam decreases from its value in vaccum, f_0, [6] by

$$\Delta f = f_0\left(1 - \sqrt{\frac{\rho_B Lbh}{M_{lat}}}\right). \tag{6}$$

As the viscous damping or displaced fluid mass increases, the total effective mass will increase, decreasing the resonant frequency. Since the amount of viscous damping is smaller when operating in the in-plane flexural mode compared to the out-of-plane flexural mode, the resonant frequency is expected to decrease by a smaller amount. The resonant frequency of the fundamental flexural mode of a laterally vibrating beam in a viscous liquid medium, $f_{res,lat}$, can be compared to the resonant frequency, $f_{res,trans}$, of the same beam vibrating transversely. Dividing (1) by the expression for the transverse resonant frequency in [7], the ratio of these two values is predicted to be

$$\frac{f_{res,lat}}{f_{res,trans}} = \frac{b}{h}\sqrt{\frac{M_{trans}}{M_{lat}}} \tag{7}$$

where M_{trans} is

$$M_{trans} = \left(\rho_B bhL + Lg_{2,trans}\right) + L\frac{\left(\left(g_{1,trans}/\omega_{res,trans}\right) + \left(\frac{\omega_{trans}}{2}\right)\frac{d}{d\omega}\left(g_{1,trans}/\omega\right)\right)}{\left(\rho_B bh + g_{2,trans} + \left(\frac{\omega_{lat}}{2}\right)\frac{d}{d\omega}\left(g_{2,trans}\right)\right)}\left(g_{1,trans}/\omega_{trans}\right) \tag{7b}$$

and $g_{1,trans}$ and $g_{2,trans}$ are defined by [7]. Note that using thinner beams (or beams with larger values for *b/h*) will result in larger differences in their resonant frequency when

operating in the in-plane flexural mode. Also note that as the dynamic viscosity and density of the medium increases, the ratio of effective masses increases.

B. Quality Factor

Microcantilevers operating in the liquid-phase suffer from a drastic decrease in their quality factors. This decrease in the quality factor increases the frequency noise (which is proportional to $f_{res,lat}/Q_{lat}$ when operating in an oscillator configuration [8]), thus increasing the limit of detection in biochemical sensing applications. The quality factor is defined as 2π times the ratio of the maximum energy stored in a resonating system to the amount of energy dissipated in one cycle [9]. However, when a resonating device is used in an oscillator configuration, the quality factor is commonly obtained using the 3-dB bandwidth Δf_{3dB} of the device and is given by

$$Q_{lat,3dB} = f_{res,lat} \Big/ \Delta f_{3dB} \,. \tag{8}$$

When the damping is small, (8) represents a good approximation to the exact definition [9]. The quality factor of a laterally vibrating beam was previously derived as [3]

$$Q_{lat} = \left(2\left(1 - \sqrt{1 - \frac{g_{1,lat}/\omega_{lat}}{\rho_B bh + g_{2,lat}}} \right) \right)^{-1} \tag{9}$$

and, when Re>>1, the quality factor can be approximated as

$$Q_{lat,approx} = 2\pi f_{res,lat} \frac{\rho_B bh + g_{2,lat}}{g_{1,lat}} \,. \tag{10}$$

The quality factor of laterally vibrating beams is predicted to be higher than that of transversely vibrating beams due to the decreased damping. The ratio of the quality factor of a laterally vibrating beam to that of a transversely vibrating beam of similar geometry is approximately given by

$$\frac{Q_{lat,approx}}{Q_{trans,approx}} = \frac{\rho_B bh + g_{2,lat}}{\rho_B bh + g_{2,trans}} \frac{\Gamma_{I,trans}}{\Gamma_{I,lat}} \,. \tag{11}$$

Note that the first fraction on the right hand side of (11) is smaller than one (when $b>h$), as transversely vibrating beams displace more fluid compared to laterally vibrating beams. The second fraction is the ratio of the amounts of viscous damping which is normally much greater than one for the cases considered.

C. Mass sensitivity

The sensitivity of the resonant frequency to changes in the beam mass, or the mass sensitivity, can be defined as [8]

$$S_m = \left| \frac{\Delta f}{\Delta m} \right| . \tag{12}$$

An analytical expression for the mass sensitivity can be obtained using the same procedures described in [7]. The expression for the mass sensitivity of a laterally vibrating beam is then the same as the expression for the transversely vibrating beam [7], but with different values for the hydrodynamic function, Γ_{lat}, and the moment of inertia I_{lat}. If the thickness of the sensing layer is considered small enough, then the mass sensitivity can be given as

$$S_m = \lambda_{m,lat} f_{res,lat} \tag{13}$$

where

$$\lambda_{m,lat} = \frac{\left(g_{1,lat} \Big/ \omega_{lat} \right)\left(g_{1,lat} \Big/ \omega_{lat} + \left(\frac{\omega_{lat}}{2} \right) \frac{d}{d\omega} \left(g_{1,lat} \Big/ \omega \right) \right) L^2}{2M_{lat} \left(\rho_B bhL + Lg_{2,lat} + L\left(\frac{\omega_{lat}}{2} \right) \frac{d}{d\omega} \left(g_{2,lat} \right) \right)^2} - \frac{1}{2M_{lat}} \tag{13a}$$

The second term in (13a) dominates the first term when Re>>1. The mass sensitivity, S_m, can then be approximated as

$$S_{m,approx} = \left| \frac{\alpha_i^{\,2}}{2\pi} \frac{\sqrt{Eb^3 hL^{-3}}}{4\sqrt{3}\left(M_{lat} \right)^{3/2}} \right| . \tag{14}$$

When the beam is operating in air or low viscosity media, the effective mass can be approximated as the beam mass, $M_{lat} \cong \rho_B Lbh$. In this case, the mass sensitivity has a geometric dependence of $h^{-1}L^{-3}$. As the decrease in the resonant frequency from air to water is small for laterally vibrating beams, this geometric dependency mostly holds when operating in a viscous liquid medium. Since the chemical sensitivity is proportional to the mass sensitivity multiplied by the volume of the sensing layer (e.g., Lbh_2, where h_2 is the thickness of the sensing layer) the chemical sensitivity is predicted to be proportional to $(h_2/h)*(b/L^2)$. As the resonant frequency is known to be proportional to b/L^2 (e.g., [3]), beams with higher resonant frequencies are then predicted to have higher chemical sensitivities.

The ratio of the mass sensitivity of a laterally vibrating beam and that of the same beam vibrating transversely is also dependent on the beams' geometry. Using (14), this ratio can be approximated as

$$\frac{S_{m,lat,approx}}{S_{m,trans,approx}} = \frac{f_{res,lat}}{f_{res,trans}} \frac{M_{trans}}{M_{lat}} = \frac{b}{h} \left(\frac{M_{trans}}{M_{lat}} \right)^{3/2} . \tag{15}$$

The effective mass of transversely vibrating beams is larger than that of laterally vibrating beams. Thus, the increase in the mass sensitivity when using lateral excitation compared with transverse excitation will be equal to or greater than b/h. Note again that thinner beams will benefit more when operating in the in-plane flexural mode than thicker beams. Also note that this ratio will be larger than the ratio of the resonant frequencies by a factor of M_{trans}/M_{lat}.

III. Results and Discussion

Experimental results for the resonant frequency and the quality factor of laterally vibrating beams both in water and aqueous mixtures of glycerol were obtained. The set-up used to extract the experimental data is described in [10]. The microcantilevers were laterally excited using thermal excitation. The beam's deflection as a function of excitation frequency was obtained using a Wheatstone bridge fabricated on the beam with the sensing resistors oriented so that the bridge voltage could be related to the lateral deflection of the beam. Solutions of varying aqueous percent glycerol (from pure water to 20% w/w) were placed in a flow cell containing the microcantilevers so that the beams were completely immersed in the solution. The spectrum of each device was then obtained and the resonant frequency and quality factor of the first in-plane flexural mode were extracted. Several microcantilevers were tested and sample results will be discussed for selected geometries. The silicon-based microcantilevers used in this work have a nominal silicon thickness of either 4.8 or 9.0 µm and are covered with a 1.9 µm passivation layer on the top made up of alternating layers of silicon nitride and silicon oxide (for fabrication details, see [10]).

The decrease in the resonant frequency of a 200x60x6.7 μm^3 laterally vibrating beam due to the increase in the viscous damping and displaced fluid mass as a function of the percent aqueous glycerol is shown in Fig. 2. The theoretical results calculated using (1) show similar trends; however, the decrease in the experimentally determined resonant frequency is smaller than the decrease in the theoretically predicted resonant frequency. The predicted resonant frequency in water is slightly different from the experimentally determined value (1390.7 kHz vs. 1381.3 kHz, respectively). The predicted decrease in the resonant frequency also has better agreement with the experimental results for longer beams, i.e. 1000x90x10.9 μm^3 beam. This suggests that the difference could be due to effects that are not accounted for when the assumption of a long beam is made. This assumption neglects the effects of the shear and rotational inertia of the beam, which cause the resonant frequency to be lower. These effects tend to be larger for shorter beams.

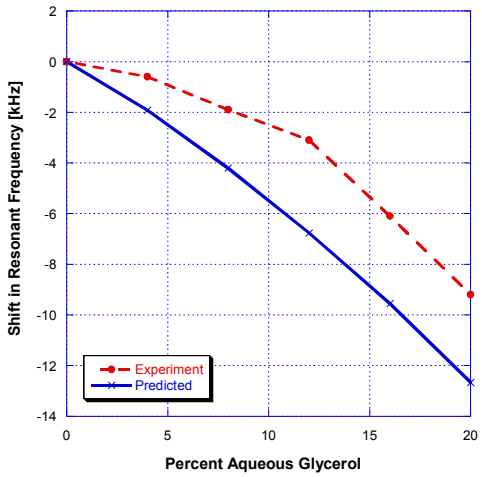


Figure 2: Predicted and experimentally determined shift in the resonant frequency of a 200x60x6.7 µm laterally vibrating microcantilever from water to a solution of aqueous glycerol.

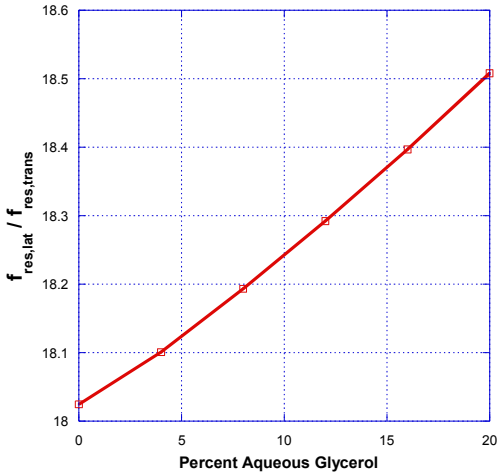


Figure 3: The predicted ratios of the first lateral and transverse resonant frequencies as a function of the medium's percent aqueous glycerol for a 200x60x6.7 µm beam

Due to the increased stiffness of operating in the in-plane flexural mode, the resonant frequencies of laterally vibrating beams will be higher than those of transversely vibrating beams of similar geometry. When the effective mass can be approximated by the beam mass (such as when operating in vacuum or air), the predicted ratio of resonant frequencies of laterally and transversely vibrating microcantilevers is approximately equal to the ratio of their width and thickness. When the beam is operating in a more viscous liquid medium, this ratio of frequencies increases. The predicted ratio of the resonant frequencies of a laterally and transversely vibrating 200x60x6.7 μm^3 microcantilever is given in Fig. 3 as a function of the percent aqueous glycerol in the operational medium. This ratio is larger than the width to thickness ratio of 8.95 due to the differences in the effective mass. As noted before, the resonant frequency of a laterally vibrating beam decreases with increasing percent aqueous glycerol. This means that the increase in the ratio is due to the transverse resonant frequency dropping more drastically than its lateral counterpart as the viscosity is increased. Because a high

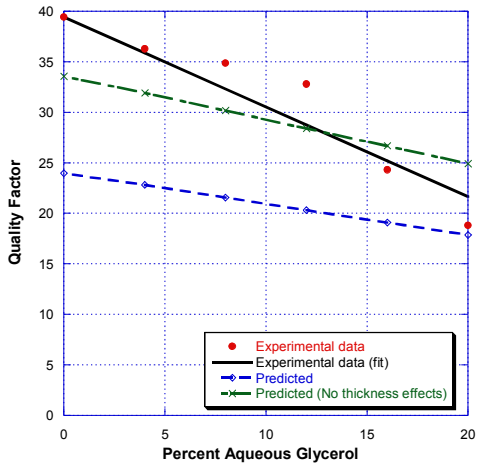


Figure 4: Predicted and experimentally determined quality factors of a 200x60x6.7 µm laterally vibrating microcantilever as a function of percent aqueous glycerol.

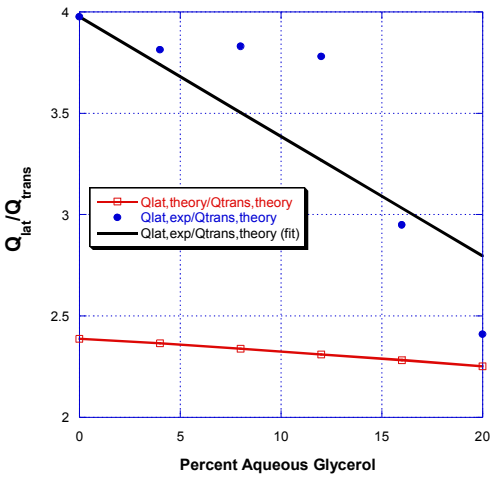


Figure 5: The ratios of the quality factors of laterally (both predicted and experimentally determined) and transversely vibrating beams (200x60x6.7 µm) as a function of percent aqueous glycerol (medium of operation).

resonant frequency is desired for sensing applications, it is thus more advantageous to operate in the in-plane flexural mode when operating in high-viscosity media.

The quality factor as a function of percent aqueous glycerol for the same beam (200x60x6.7 µm^3) is given in Fig. 4. The experimentally determined quality factors are larger than the commonly obtained values for transversely vibrating beams, which are usually on the order of 10 [11]. While the experimental results show similar trends as the theoretically predicted results, the experimentally determined quality factors are larger than the theoretically predicted quality factors. Calculated values for the quality factor using a simpler expression given by [12] are also shown in Fig. 4. However, this expression neglects the pressure effects of the fluid on the smaller faces (of dimension *h*), i.e., it is based on assuming only a Stokes-type fluid resistance on the top and bottom faces of the beam. The theoretically predicted quality factors would be smaller if the shear deformation and support compliance effects were taken into account. Thus, these effects are not the source of the discrepancy. Another assumption made when the hydrodynamic forces were estimated in [3] was that the beam was infinitely long. This assumption was made to neglect the edge effects near the end of the beam in the lengthwise direction, and might not be appropriate for the given geometry. It is noted that the quality factors of longer beams (e.g. 1000x90x10.9 µm^3) more closely match the characteristics predicted by theory.

Regardless of the length used, the quality factor tends to be higher when operating in the in-plane flexural mode compared to the out-of-plane flexural mode. It is predicted, when operating in air, that the quality factor of a 200x60x6.7 µm beam is a factor of 6.16 times larger for the in-plane flexural mode than for the out-of-plane flexural mode. This increase, unlike the resonant frequency, is expected to be smaller when operating in a viscous liquid medium. Fig. 5 shows the predicted ratio of the quality factor of a 200x60x6.7 µm^3 laterally vibrating silicon microcantilever and that of a transversely vibrating microcantilever of the same geometry as a function of percent aqueous glycerol. As the experimental results show higher quality factor values than the predicted theoretical results, this improvement is expected to be higher when using the experimental results for comparison. The improvement in the quality factor is also a function of the beam's geometry. For example, longer and narrower beams will have a smaller increase in quality factor when compared to shorter and wider beams. This is because shorter and wider beams have higher resonant frequencies, Reynolds numbers, and thus larger damping ratios.

The mass sensitivity was not experimentally determined in this study. However, the mass sensitivity can still be estimated for the 200x60x6.7 µm^3 microcantilever as a function of percent aqueous glycerol, and is given in Fig. 6. While the predicted mass sensitivity is expected to decrease as the percent aqueous glycerol increases, the rate of decrease is much smaller than that of similar beams vibrating transversely. For example, a 200x60x6.7 µm^3 laterally vibrating microcantilever in water is predicted to have a 3.26 Hz/pg mass sensitivity compared to a 0.0443 Hz/pg mass sensitivity of the same beam vibrating transversely. When the viscosity of the operational environment is increased to 1.734 cP (20% aqueous glycerol), the mass sensitivity of the transversely vibrating beam decreases by 10.4% to 0.0397

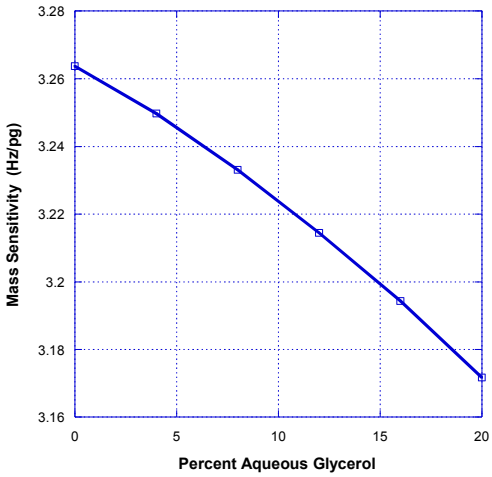


Figure 6: Predicted mass sensitivity of a 200x60x6.7 µm laterally vibrating microcantilever as a function of percent aqueous glycerol.

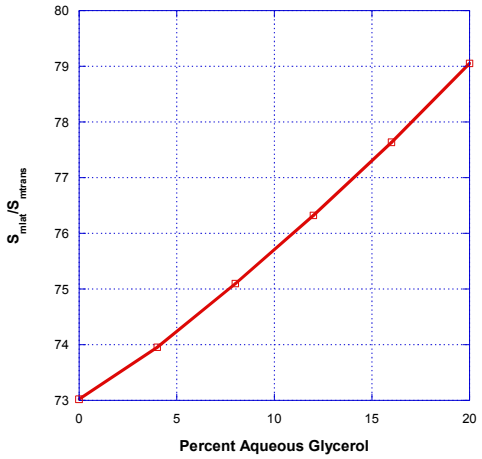


Figure 7: The ratio of the mass sensitivities of a laterally and a transversely vibrating beam for a 200x60x6.7 µm microcantilever as a function of percent aqueous glycerol in the operational medium.

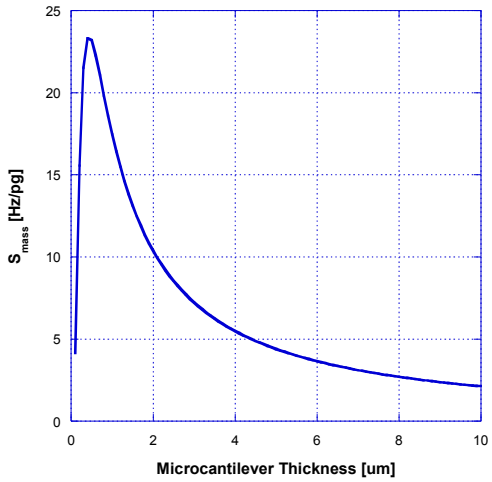


Figure 8. The mass sensitivity of a laterally vibrating 200x60 µm microcantilever in water as a function of beam thickness.

Hz/pg, while the same beam laterally vibrating only decreases by 2.76% to 3.17 Hz/pg. This indicates a further improvement in more viscous fluid. The ratio of the mass sensitivity as a function of percent aqueous glycerol is given in Fig. 7.

The ratio of the chemical sensitivities, assuming the same geometry and sensing layer, is the same as the ratio of the mass sensitivities. If the thickness of the beam is decreased, this ratio would be even larger, as the mass sensitivity is also a function of the beam's geometry. Shorter and thinner beams are expected to have higher mass sensitivities. The predicted mass sensitivity of a 200x60 μm^2 silicon microcantilever in water as a function of thickness is given in Fig. 8. Note that the mass sensitivity is predicted to initially increase with increasing thickness. This increase occurs because, for small thicknesses, the fluid mass dominates the beam mass. It's noted that the stiffness is linearly dependent on the thickness. When the beam mass dominates the fluid mass as the thickness increases, the mass sensitivity decreases. However, as discussed in [3], the quality factor will also increase with increasing thickness. As the limit of detection is a function of both the chemical sensitivity (which depends on the mass sensitivity) and the quality factor, both must be considered when choosing the appropriate thickness of a microcantilever for a given application. However, increasing the width will increase the quality factor while not greatly affecting the mass sensitivity of laterally vibrating beams; and decreasing the length will increase both the quality factor and the mass sensitivity of laterally vibrating beams.

IV. CONCLUSIONS

It is shown that laterally vibrating microcantilevers can be excited and have relatively high (20 to 60) quality factors, even when operating in media with higher viscosity than water. The resonant frequency of laterally vibrating beams in a viscous liquid medium is also observed to experience relatively insignificant decreases (compared to its resonant frequency in air). When compared to transversely vibrating beams, the predicted characteristics (i.e. the resonant frequency, quality factor, and mass sensitivity) are shown to be higher. Considering the resonant frequency and mass sensitivity, this increase is larger for media with higher dynamic viscosities. These improvements result in an almost two orders of magnitude higher chemical sensitivity in viscous solutions. Besides for chemical sensing, this is ideal for biosensor applications where biofluids may have higher viscosities than water. It is also shown that shorter and wider beams will have higher quality factors and larger improvements in their mass sensitivity compared to transversely vibrating microcantilevers. This indicates that the limit of detection of microcantilever bio-chemical sensors operating in the in-plane flexural mode will be lower than that of microcantilevers operating in the out-of-plane flexural mode in liquid-phase sensing applications.

ACKNOWLEDGMENTS

This work has been funded in part by the National Science Foundation under award ECCS-0824017

REFERENCES

[1] L. A. Pinnaduwage, V. Boiadjiev, J. E. Hawk, and T. Thundat, "Sensitive Detection of Plastic Explosives with Self-Assembled Monolayer-coated Microcantilevers," Applied Physics Letters, vol. 83, no. 7, p. 1471, 2003.

[2] J. Davis, G. Abadal, O. Kuhn, O. Hansen, F. Grey, and A. Boisen, "Fabrication and Characterization of Nanoresonating Devices for Mass Detection," Journal of Vacuum Science & Technology B, vol. 18, no. 2, pp. 612-616, 2000.

[3] R. Cox, F. Josse, S. Heinrich, I. Dufour, O. Brand, "Resonant Microcantilevers Vibrating Laterally in Viscous Liquid Media," IEEE International Frequency Control Symposium (FCS), pp. 85-90, 2010.

[4] D. Brumley, M. Willcox, and, J. Sader, "Oscillation of Cylinders of Rectangular Cross section Immersed in Fluid," Physics of Fluids, vol. 22, no. 052001, 2010.

[5] J. Sader, "Frequency Response of Cantilever Beams Immersed in Viscous Fluids with Applications to the Atomic Force Microscope," Journal of Applied Physics, vol. 84, no. 1, pp. 64-76, July 1998.

[6] I. Dufour, S. Heinrich, and F. Josse, "Theoretical Analysis of Strong-Axis Bending Mode Vibrations for Resonant Microcantilever (Bio)Chemical Sensors in Gas or Liquid Phase," Journal of Microelectromechanical Systems, vol.16, iss.1, pp. 44-49, Feb. 2007.

[7] R. Cox, F. Josse, M. J. Wenzel, S. Heinrich, and I. Dufour, "Generalized Model of Resonant Polymer-Coated Microcantilevers in Viscous Liquid Media," Analytical Chemistry, vol. 80, no. 15, pp. 5760-5767, 2008.

[8] I. Dufour, F. Lochon, S. Heinrich, F. Josse, D. Rebière, "Effect of Coating Viscoelasticity on Quality Factor and Limit of Detection of Microcantilever Chemical Sensors ", IEEE Sensors Journal, vol. 7, 2007, pp. 230-236

[9] B. Razavi, "A Study of Phase Noise in CMOS Oscillators," IEEE Journal of Solid-State Circuits, vol. 31, no. 3, pp. 331-343, 1996.

[10] L. Beardslee, A. M. Addous, S. Heinrich, F. Josse, I. Dufour, and O. Brand, "Thermal Excitation and Piezoresistive Detection of Cantilever In-Plane Resonance Modes for Sensing Applications," Journal of Microelectromechanical Systems, vol. 19, no. 4, pp. 1015-1017, 2010.

[11] J. Chon, P. Mulvaney, and J. Sader, "Experimental Validation of Theoretical Models for the Frequency Response of Atomic Force Microscope Cantilever Beams Immersed in Fluids," Journal of Applied Physics, vol. 87, no. 8, pp. 3978-3988, April 2000.

[12] S. M. Heinrich, R. Maharjan, I. Dufour, F. Josse, L. Beardslee and O. Brand, "An Analytical Model of a Thermally Excited Microcantilever Vibrating Laterally in a Viscous Fluid," Proceedings IEEE Sensors 2010 Conference, Waikoloa, Hawaii, pp. 1399-1404., 2010

Liquid Sensor Utilizing a Regular Phononic Crystal with Normal Incidence of Sound

Ralf Lucklum*, Mikhail Zubtsov*
* Otto-von-Guericke-University
Institute of Micro and Sensor Systems
Magdeburg, Germany
ralf.lucklum@ovgu.de

Manzhu Ke*,§
§ Wuhan University
Department of Physics,
Wuhan, China

***Abstract*— A phononic crystal device is investigated as a sensor platform combining bandgap engineering with resonant transmission. Here we apply a regular phononic crystal with a solid matrix and liquid filled holes, fully immersed into the liquid with incidence direction of sound normal to the plate.**

I. INTRODUCTION

Phononic crystals are periodic composite materials with spatial modulation of acoustically relevant parameters like elasticity, mass density and longitudinal and transverse velocities of elastic waves. When applied as sensor, the material of interest constitutes one component of the phononic crystal, e.g., a fluid in the holes of a regular phononic crystal having a solid matrix. If the value of interest, here the concentration of one component in a liquid binary mixture, changes acoustic properties of this mixture, the acoustic properties of the phononic crystal will also change. Transmission or reflection coefficients are appropriate parameters for measurement and used to localize a characteristic feature of the phononic crystal. For a sensor application, a transmission peak within the band is the most favorable feature since the respective frequency of maximum/minimum transmission is easy to determine. The sensor scheme therefore relies on the determination of the frequency dependence of maximum transmission on the physical or chemical value of interest.

The sensitivity of the sensor, S_f can be defined as the ratio of frequency shift, Δf, and change of the input parameter, Δx:

$$S_f = \frac{\Delta f}{\Delta x} \tag{1}$$

Similar to acoustic microsensors the sensitivity has been found to be dependent on the probing frequency, f_0. We furthermore consider a certain frequency measurement resolution by taking the peak half band width, f_{HBW}, into account. We define therefore a reduced sensitivity, S_{fr}, which gives much better insights to the sensor capabilities:

$$S_f = \frac{\Delta f}{\Delta x \; f_0 \; \Delta f_{HBW}} \tag{2}$$

We recently have applied the one-dimensional concept for the first time with an arrangement of parallel solid plates and a liquid-filled gap in-between to determine the concentration of propanol in water [1,2]. This kind of realization is a reduction of a phononic crystal to one dimension in its most simple form. The difference of the acoustic impedance between aluminum, the material the solid plates are made of, and the liquid, the water-propanol mixture in the gap and pure water surrounding the phononic crystal is large enough to create a band gap sufficiently large for sensing purposes. We have designed the phononic band gap that a broad transmission window appears at the resonance frequency of the ultrasonic transducers used for generation and detection of ultrasonic waves. The narrow transmission peak sensitive to the composition of the liquid mixture in the gap emerges in the fist band gap below. The sensitivity $\Delta f/\Delta x_2$ in the concentration range between x_2 = 0–0.035 of 1-propanol (0–10%) is about 1.086 MHz, the reduced sensitivity is about 0.19 kHz^{-1}. One can estimate a limit of detection of propanol of about 0.1%, which has good practical relevance.

Although the one-dimensional approach can be reduced by size, one always has to realize a certain aspect ratio between radius of the plates/gaps and thickness of the plate/gaps. Moreover, the sensor only probes bulk properties since only longitudinal waves are useful.

We therefore go the next step and apply a two-dimensional regular phononic crystal keeping the incidence of sound normal to the plate. Still, a certain diameter of the plate is required, however, more opportunities exist in miniaturization. In addition, this realization features rich physics which can be applied in different sensing schemes based on different modes which are excited in the plate by the incoming longitudinal wave. Here we concentrate on the most apparent feature, the so-called extraordinary acoustic transmission phenomenon [3].

978-1-61284-110-6/11/$26.00 ©2011 IEEE

II. THEORY

Extraordinary acoustic transmission through plates has been found for several geometries, e.g., through metal plates with narrow gratings [4], sub-wavelength hole arrays [5-10] and through plates coupled by pillars [11,12] or a plate having linear gratings on both sides [13]. There are no generally recognized theory or conceptual framework concerning EAT, though the behavior of acoustic waves propagating through these structures has been a subject of intensive research. Extraordinary transmission has been attributed to the coupling between the compositions of diffractive waves excited on the surfaces of the grating and the Fabry-Perot resonant modes inside the apertures [4]. Evanescent waves are assumed altering Fabry-Perot resonance conditions and resonance can occur in apertures less than half wavelength of the incident wave. Resonance is smoothly linked to the surface-wave-like modes induced by the periodic structure factor [5,6]. Several transmission mechanisms are involved which produce a complex interplay phenomena, specifically interaction among holes arranged in an array and direct transmission through the bulk material [7]. When varying the thickness of the plate, the transition from the Fabry-Perot resonances (thick plate limit) to the structure-factor-induced surface modes resonance (thin plate limit) will take place. We have chosen a thin plate whose thickness is smaller than the minimum half wavelength of the scanned frequency range. The structure we have used consists of a steel plate with liquid-filled holes in square arrays. The lattice distance is 1.5 mm, the thickness of the plate is 0.5 mm and the diameter of the hole is 0.5 mm.

The three-dimensional finite-difference time-domain (FDTD) method has been applied to calculate the transmission spectrum through the plate. The elastic wave equation in homogeneous solids is given by equation (3) and (4):

$$\rho\frac{\partial^2\mu_x}{\partial t^2}=\frac{\partial T_{xx}}{\partial x}+\frac{\partial T_{xy}}{\partial y}+\frac{\partial T_{zx}}{\partial z} \tag{3a}$$

$$\rho\frac{\partial^2\mu_y}{\partial t^2}=\frac{\partial T_{xy}}{\partial x}+\frac{\partial T_{yy}}{\partial y}+\frac{\partial T_{yz}}{\partial z} \tag{3b}$$

$$\rho\frac{\partial^2\mu_z}{\partial t^2}=\frac{\partial T_{zx}}{\partial x}+\frac{\partial T_{yz}}{\partial y}+\frac{\partial T_{zz}}{\partial z} \tag{3c}$$

$$T_{xx}=(\lambda+2\mu)\frac{\partial\mu_x}{\partial x}+\lambda(\frac{\partial\mu_y}{\partial y}+\frac{\partial\mu_z}{\partial z}) \tag{4a}$$

$$T_{xy}=\mu(\frac{\partial\mu_x}{\partial y}+\frac{\partial\mu_y}{\partial x}) \tag{4b}$$

$$T_{yy}=(\lambda+2\mu)\frac{\partial\mu_y}{\partial y}+\lambda(\frac{\partial\mu_x}{\partial x}+\frac{\partial\mu_z}{\partial z}) \tag{4c}$$

$$T_{yz}=\mu(\frac{\partial\mu_y}{\partial z}+\frac{\partial\mu_z}{\partial y}) \tag{4d}$$

$$T_{zz}=(\lambda+2\mu)\frac{\partial\mu_z}{\partial z}+\lambda(\frac{\partial\mu_y}{\partial y}+\frac{\partial\mu_x}{\partial x}) \tag{4e}$$

$$T_{zx}=\mu(\frac{\partial\mu_z}{\partial x}+\frac{\partial\mu_x}{\partial z}) \tag{4f}$$

$\vec{u}(\vec{r})$ is the displacement vector, $\lambda(\vec{r})$ and $\mu(\vec{r})$ are the Lame coefficients, and $\rho(\vec{r})$ is the mass density. One obtains the relation equations of stress (T) and the velocity (v) by using discretization of the equations in both space and time domains. In the FDTD calculations, we have chosen a computational cell with dimension 30×30×300 grid points, corresponding to 30 grid points per lattice constant. A spatial grid with mesh size $\Delta x = \Delta y = \Delta z = 0.05$ mm has been used. The faces of the computational cell normal to x and y axes are chosen to have Bloch boundary conditions, while the faces normal to z axis (the top and bottom ones) have Mur's absorbing boundary conditions. For transmission spectrum calculations, a broad band wave packet is launched to illuminate the plate with a normal incidence. The transmitted signal is recorded as a function of time. The frequency dependence of the field at a particular point can be obtained by fast Fourier transforming the time results. The transmission coefficient is calculated by dividing the transmitted energy flux by the energy flux of the incident wave.

COMSOL™ simulations have been performed revealing further insights. We especially have been concentrating on coupled elements: liquid confined in the holes, regular set of portions of the plate surrounded by the holes and regular set of holes with adjacent portions of the plate. Each of these elements has their own set of modes and each of these modes can support either transmission or reflection. Eigenfrequencies of each of these elements depend on the boundary conditions including rheological properties of the liquid, which may lead to sensing schemes not considered so far.

III. EXPERIMENTAL

Different liquid properties are obtained by gradually changing the liquid in the holes and the surrounding from pure DI-water to a series of liquid mixture with different molar concentration of 1-propanol similar to [1,2]. Density and sound velocity have been taken from [3]. Whereas density decreases monotone with molar ratio of 1-propanol from 998 kg m^{-3} for pure DI-water to 804 kg m^{-3} (about 20 %) speed of sound increases from 1483 m s^{-1} to a maximum of 1588 m s^{-1} at $x_2 = 0.056$ (7 %) and decreases to 1220 m s^{-1} (for pure 1-propanol (-25 %; with respect to water).

The transmission curves have been obtained using a network analyzer (Agilent 4395A) together with an S-parameter test set (Agilent 87511A). Transmission S21 has been measured with and without sample in place respectively. The transmission amplitude is obtained by normalizing with the amplitude of the equivalent setup without the sample.

Several stages have been used to fix and align the sample parallel between two piezoelectric transducers with a central frequency of 1 MHz having a half bandwidth of 45 kHz, Fig. 1. Although the signal amplitude reduces considerably off the resonance, the amplitude is still sufficiently above the noise floor. The distance between the transducers and metal plate is large enough to avoid the near field effect and multiple reflections. A small tank contains the liquid. Porous plastic discs have been placed in the container to reduce unwanted reflection of the acoustic waves at the container walls.

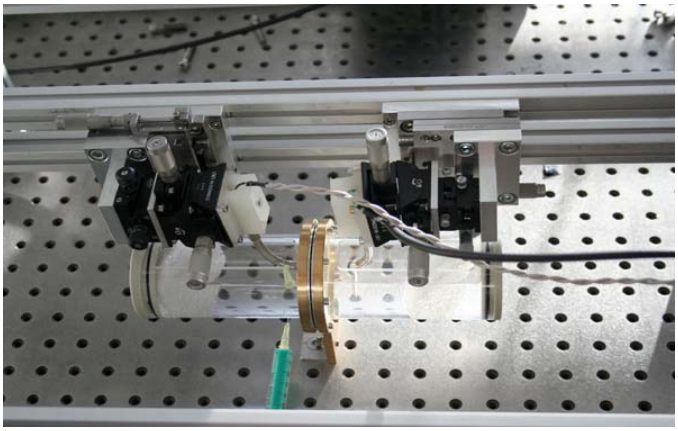


Figure 1. Photo of the experimental setup

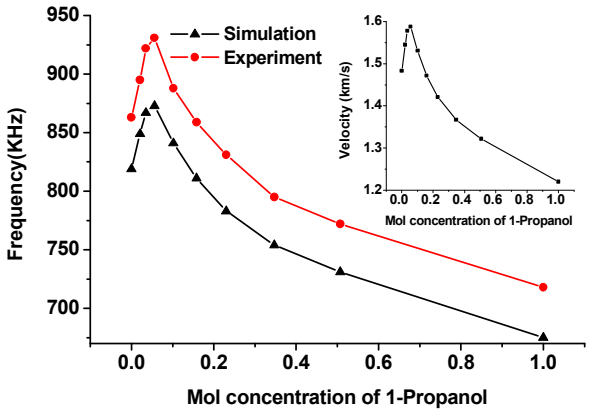


Figure 2. Results of resonant frequency *vs.* molar concentration of 1-propanol in the liquid mixture. The black curve with triangular symbols reflects the FDTD simulation and the red curve with circular symbols summarizes the experimental results. The maximum transmission peak frequency corresponds to the extreme in speed of sound at molar ratio $x_2 = 0.056$. The inset reflects speed of sound of the liquid mixture.

IV. Results

We observed both in FDTD simulation and in experiment that the frequency of resonant transmission peak significantly changes with the molar ratio of 1-propanol in the liquid mixture, Fig. 2. This finding resembles the results of measurement and simulation of the one-dimensional arrangement [1,2]. Specifically, the shift of the maximum transmission peak frequency again reflects the difference in speed of sound. The highest resonance frequency appears at a molar ratio $x_2 = 0.056$, corresponding to the maximum acoustic velocity of the mixture. The sensitivity $\Delta f/\Delta x_2$ in the concentration range between $x_2 = 0$–0.035 is about 3 MHz and three times larger, the reduced sensitivity is about 0.065 kHz^{-1} and hence only 3 times smaller than in the one-dimensional case. The major contribution for this decline in the reduced sensitivity is the 8 times larger half band width of the extraordinary transmission peak. Figs. 4 a-d (next page) reveal another effect not yet understood, the decline of the peak maximum. Viscosity changes of the liquid mixture are not very likely, since this value has also passes an extreme whereas the peak amplitude decreases monotone. Finally, Fig. 4 exhibits also a systematic frequency difference between simulation and experiment. Further analyzes reveal the reason for the difference. We have found that only a slight difference in the lattice constant between the fabricated sample and the design can cause this shift whereas the peak frequency is much less sensitive to similar deviations in the hole diameter as well as in the acoustic velocity of steel.

Fig. 3 exemplarily shows acoustic wave propagation through the device at the frequency of extraordinary transmission. First of all, the calculations confirm earlier findings, the sound pressure before and after the phononic crystal is the same. The value is much higher in the hole, an effect which does not appear having only one hole in the plate. Fig. 3 also demonstrates the excitation of several modes in the plate by the incoming longitudinal wave which very much depend on the frequency. The extraordinary acoustic transmission peak must therefore be attributed to 'collective' resonances of the liquid inside the holes.

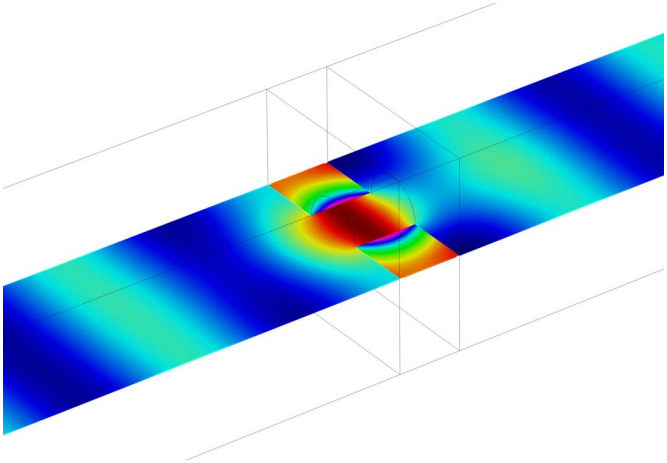


Figure 3. Comsol Simulation of the extraordinary transmission of sound through an array of sub-wavelenth holes

V. Conclusions

Regular 2-dimensional phononic crystals in an arrangement with normal incidence of sound can be used as sensor to determine acoustic properties of liquids like speed of sound concentration of a component in a mixture with sensitivities comparable to ultrasonic sensors. The sensing effect is based on collective resonances inside the holes, hence the principle can be used for very small volumes.

Acknowledgment

The work has been supported by a grant of the German Research Foundation (Lu 606/12-1) and the European Commission Seventh Framework Program (233883, TAILPHOX) which is gratefully acknowledged. Furthermore, the authors wish to acknowledge financial support by a project funded by the Ministry of Education of Saxonia-Anhalt (FKZ 5159AD/0308T).

REFERENCES

[1] R. Lucklum, I. Hauer, "Transmission properties of a 1D resonant cavity, " 2009 EFTF-IEEE IFCS Joint Conf. (Besancon, France), Proc. pp. 248–253

[2] R. Lucklum J. Li, "Phononic crystals for liquid sensor applications", Meas. Sci. Technol., 20, pp. 124014, 2009.

[3] R. Kuhnkies, W. Schaaffs, Acustica 13, 407, 196.

[4] M.H. Lu, X.K. Liu, L. Feng, J. Li, C.P. Huang, Y.F. Chen, Y.Y. Zhu, S. N. Zhu, and N.B. Ming, "Extraordinary Acoustic Transmission through a 1D Grating with Very Narrow Apertures, " Phys. Rev. Lett. 99, pp. 174301, 2007.

[5] B. Hou, J. Mei, M. Ke, W. Wen, Z. Liu, J. Shi and P. Sheng, "Tuning Fabry-Perot resonances via diffraction evanescent waves," Phys. Rev. B., 76, pp. 054303, 2007.

[6] B. Hou, J. Mei, M. Ke, Z. Liu, J. Shi, W. Wen, "Experimental determination for resonance-induced transmission of acoustic waves through subwavelength hole arrays," J. Appl. Phys., 104, pp. 014909, 2008.

[7] H. Estrada, P. Candelas, A. Uris, F. Belmar, F.J. García de Abajo, F. Meseguer, "Extraordinary Sound Screening in Perforated Plates," Phys. Rev. Lett. 101, pp. 084302, 2008.

[8] J. Christessen, L. Martin-Moreno, F.J. Garcia-Vidal, "Theory of Resonant Acoustic Transmission through Subwavelength Apertures," Phys. Rev. Lett. 101, pp. 014301, 2008.

[9] X. Wang, "Theory of resonant sound transmission through small apertures on periodically perforated slabs," J. Appl. Phys. 108, pp. 064903, 2010.

[10] P. Peng, C.Qiu, R. Hao, J. Lu, Z. Liu, "Acoustic transmission enhancement through a stiff plate drilled with subwavelength side openings," EPL 93 pp. 34004, 2011.

[11] B. Djafari-Rouhani, Y. Pennec, H. Larabi, "Band structure and wave guiding in a phononic crystal constituted by a periodic array of dots deposited on a homogeneous plate," Proc. SPIE, 7223, 72230F, 2009.

[12] B. Djafari-Rouhani, Y. Pennec, H. Larabi, " Band Structure and phonon transport in a phononic crystal made of a periodic array of dots on a membrane," IUTAM Symposium on Recent Advances of Acoustic Waves in Solids (Taipei), Springer IUTAM Book Series 26, pp. 127-138, 2010.

[13] Z. He, H. Jia, C. Qiu, S. Peng, X. Mei, F. Cai, P. Peng, M. Ke, Z. Liu, "Acoustic Transmission Enhancement through a Periodically Structured Stiff Plate without Any Opening," Phys. Rev. Lett. 105, pp. 074301, 2010.

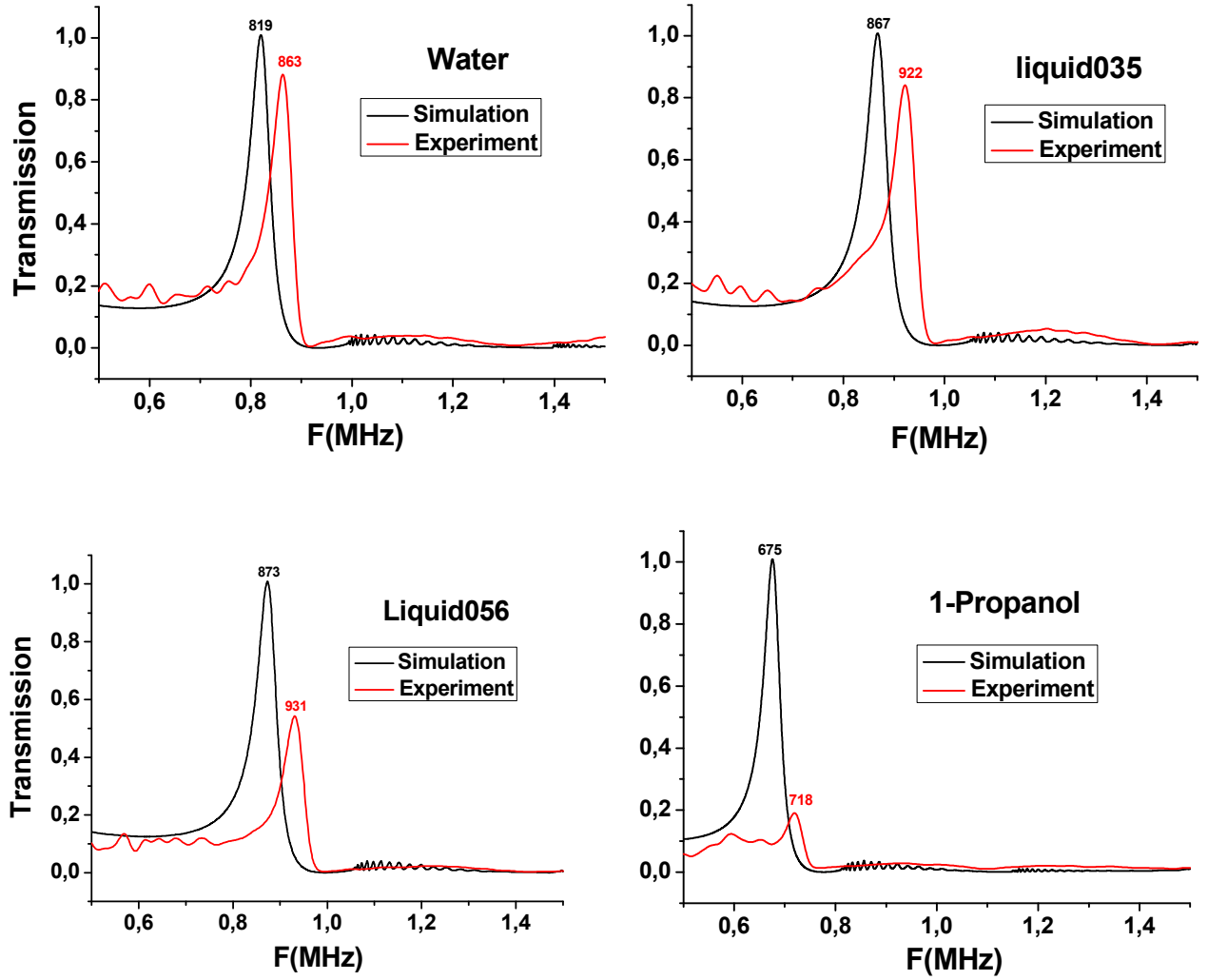


Figure 4. Comparison between simulated and experimentally determined transmission spectra for different molar ratios of 1-propanol (x_2 = 0, 0.035, 0.056, 1)

Influence of Ambient Parameters on the Response of Polymer-Coated SH-Surface Acoustic Wave Sensors to Aromatic Analytes in Liquid-Phase Detection

Florian Bender and Fabien Josse
Department of EECE
Marquette University
Milwaukee, WI 53201-1881
florian.bender@mu.edu
fabien.josse@mu.edu

Antonio J. Ricco
Department of Electrical Engineering
Center for Integrated Systems
Stanford University
Stanford, CA 94305-4075

***Abstract*—This work addresses the importance of ambient conditions in the design of a sensor system for real-time monitoring of contaminated water sites and groundwater *in situ* for fuel and oil spills. The sensor system is based on shear-horizontal surface acoustic wave devices with chemically sensitive polymer coatings for analyte sorption. In addition to the use of a sensor array, evaluation of response time (absorption time constant) is discussed as a tool for analyte identification. The influence of relevant ambient parameters (temperature, pH, salinity) on the sensor response is described, and resulting requirements for the design of the sensor system are discussed.**

I. INTRODUCTION

Various types of chemical compounds are known to pose a demonstrable or potential health hazard through contamination of groundwater. This has resulted in the introduction of legal limits for a broad range of these compounds both for industrial effluent and for drinking water [1]. Current methods for monitoring contaminants include gas/liquid chromatography and mass spectroscopy. Although these methods can be very accurate, they often require samples to be taken to laboratories for analysis, are very time consuming, and relatively expensive. Moreover, vital information can be lost during sample collection, transportation, and storage. Therefore, there is a need to develop *in-situ* monitoring systems for rapid analysis and characterization of samples. An ideal solution would be a compact, portable sensor system that permits real-time monitoring of critical sites for contamination. Various types of sensor systems have been proposed for this application, including acoustic wave sensor devices [1]. It has been demonstrated that shear-horizontal surface acoustic wave (SH-SAW) sensor devices can be used to detect small concentrations of chemical and biochemical analytes *in situ* in liquid phase [2]–[4]. In addition, SH-SAW sensor systems can be made small and portable [1]. This permits real-time, distributed monitoring of groundwater or contaminated sites without the need to collect samples and analyze them off-site in a laboratory. Consequently, such sensor systems show the potential to reduce analysis time and cost. In addition, they can be used to gain a more detailed and comprehensive understanding of the distribution and dispersion pathways of chemicals in contaminated sites.

In contrast to a controlled laboratory environment, this approach requires a very thorough understanding of the ways a changing environment will affect the sensor system. For example, it has been shown that the responses of absorption-based vapor sensors can be significantly influenced by environmental parameters such as ambient temperature, which affects the partitioning of analyte vapor into polymer sensing films [5]–[7]. For *in-situ* measurements where ambient parameters cannot always be controlled, it is crucial to understand such effects and to compensate or correct for them by appropriate system design.

In this work, the investigation of a SH-SAW sensor system to detect aromatic hydrocarbons at trace levels in contaminated water *in situ* and in real time is reported. Aromatic hydrocarbons such as benzene, toluene, ethylbenzene, and xylenes (the BTEX compounds) are major components of gasoline and diesel fuels [8] and therefore are important targets in groundwater and contaminated site monitoring. In this application, temperature, pH, and salt concentration of a particular monitoring site depend significantly on various factors including geographic location. Active control of the environmental variables would require unacceptable increases in system complexity, size, and power consumption. Therefore, the influence of temperature and other ambient parameters such as pH or salt concentration on the sensor response must be determined. Results can then be used to develop a system for accurate *in situ* measurement of

978-1-61284-110-6/11/$26.00 ©2011 IEEE

trace concentrations of aromatic hydrocarbons such as BTEX in contaminated water.

II. Experimental

A. Sensor Design

The sensor platform used in this work consists of a piezoelectric substrate (36° rotated Y-cut, X-propagation $LiTaO_3$) supporting a SH-SAW, and electrodes (interdigital transducers, IDTs) with a periodicity of 40 µm. This corresponds to a frequency of operation of about 103 MHz. The entire device (including the IDTs) is coated with a chemically sensitive polymer and immersed in the liquid to be monitored. A schematic of the sensor device is shown in Fig. 1. For simplicity, Fig. 1 shows only a single SH-SAW sensing channel; the actual device consists of a dual delay line design where the second SH-SAW channel is coated with poly(methyl methacrylate) (PMMA) baked for 120 min at 180 °C and serves as a reference line. It is noted that the treatment of the PMMA, which results in a glassy polymer layer, prevents analyte absorption. The purpose of the reference channel is to compensate for signal drift due to secondary effects such as changes in temperature.

Various polymers have been investigated as chemically sensitive coatings. In this work, poly(ethyl acrylate) (PEA), poly(epichlorohydrin) (PECH), and poly(isobutylene) (PIB) have been used. The polymers were deposited from solution by spin coating and baked for 15 min at 60 °C, resulting in thicknesses ranging from 0.5 to 1.0 µm as indicated below. The design of the SH-SAW sensor platform has been described in detail in previous work [2], [4].

B. Materials and Instrumentation

Among the polymers used in this study, only PEA was obtained as a solution (20% w/v in toluene, Sigma-Aldrich). The other polymers were obtained in undissolved form from Sigma-Aldrich (PECH, PIB) and Scientific Polymer Products (PMMA), respectively. PECH and PIB were dissolved in chloroform (99.8%, Sigma-Aldrich), and PMMA was dissolved in 2-ethoxyethyl acetate (98%, Sigma-Aldrich). Spin coating of the polymers was performed using a Specialty Coating Systems 6708D spin coater. All chemicals were used as received.

The BTEX compounds were obtained from Sigma-Aldrich and had the following purities: 99.8% (benzene and toluene), 99% (ethylbenzene), and 98.5% (xylenes, mixture of isomers), respectively. Only filtered and degassed DI water (Milli-Q) was used throughout this study. For pH-dependent measurements, pH was adjusted using potassium phosphate dibasic and potassium phosphate monobasic (both 98%, Sigma-Aldrich); all pH values were verified using a pH meter (Corning 220). NaCl (99%) was obtained from Sigma-Aldrich as well.

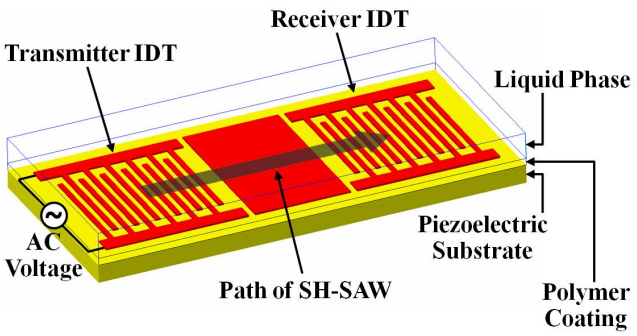


Fig. 1: Schematic of the SH-SAW sensor device.

The experimental set-up consisted of a network analyzer (Agilent 8753ES) and a switch/control system (Agilent 3499A) to switch between the two SH-SAW channels. The sensor was placed inside a flow cell made in-house, and the pump (Eppendorf EVA) was set to a sample flow rate of 0.4 mL/min. Temperature was held constant at 22.0 ± 0.1 °C; for the temperature-dependent measurements, the temperature of the flow cell and sensor device was varied using a Peltier element (Custom Thermoelectric TEC module).

III. Theoretical Considerations

Generally, the absorption of an analyte by a polymer coating on an acoustic wave sensor device affects both the velocity and the attenuation of the acoustic wave. For this study, priority is given to the change in acoustic wave velocity which is measured by a network analyzer as a proportional change in frequency. Acoustic wave loss (attenuation) has sometimes been investigated only for the purpose of interpreting the observed frequency changes.

Absorption of the analyte into the polymer coating will not only increase the mass loading on the sensor surface but will also lead to swelling of the polymer, accompanied by changes in its shear modulus. Both mechanisms will affect the characteristics of the acoustic wave. The resulting change in frequency, Δf, can be expressed as [4]:

$$\Delta f = \Delta f_C \, C_{Amb} \, K_{P\text{-}L} \, (1 / (\rho_P + C_W) + \beta). \qquad (1)$$

In this equation, Δf_C is the frequency shift due to the coating (including water absorbed by the coating), C_{Amb} is the ambient concentration of the analyte, $K_{P\text{-}L}$ is the partition coefficient of the analyte between polymer and liquid phase, ρ_P is the density of the polymer, C_W is the concentration of water in the polymer, and β is a sensitivity coefficient which is a function of device, polymer, and medium properties and contains the viscoelastic contributions to the sensor signal. Equation (1) implies that viscoelastic contributions in Δf are linear with Δf_C and C_{Amb}. Note that this applies only for small coating thicknesses and low analyte concentrations.

For some polymers, partition coefficients have been calculated for the liquid phase [9], [10]. For the BTEX compounds, liquid-phase partition coefficients for PECH and PIB are summarized in Table I.

Various groups have performed calculations to determine the effect of the viscoelastic parameters of a sensor coating on both frequency and loss of an acoustic wave [3], [11], [12]. Li *et al.* investigated the effect of the complex shear modulus of a polymer coating, $G = G' + jG''$, and determined the influence of a variation in G' and G'' on acoustic wave velocity, v, and

Table I: Calculated partition coefficients in aqueous phase, $K_{P\text{-}L}$, for two polymers, PECH and PIB [9], [10].

Analyte	PECH	PIB
Benzene	67	30
Toluene	249	140
Xylenes	700	458
Ethylbenzene	673	464

acoustic wave loss of a coated sensor device in liquid phase [3], [12]. Results of these calculations are shown in Fig. 2.

While the exact dependence of the shear moduli of the polymers used in this study as a function of temperature and analyte concentration is not known, such calculations can nevertheless serve as a guideline to the interpretation of experimental observations. As Fig. 2b indicates, the influence of shear modulus on acoustic loss is generally small for glassy or moderately rubbery polymers but can rapidly increase as the polymer gets softer. It is well known that the shear modulus of a polymer depends on temperature in a complex way [13], undergoing a rapid decrease around the glass transition temperature, T_g, and changing at a lower rate in the rubbery regime above T_g. The exact behavior of the shear modulus in the rubbery regime depends on the type of polymer [13]. The chemically sensitive polymers used in this study have glass transition temperatures below the freezing point of water (see Table II), and thus will be in the rubbery state in groundwater monitoring applications.

While exact data on the mechanical properties of our chemically sensitive polymers in the rubbery regime are not available from the literature, experimental observation shows an increase in acoustic loss with temperature for all three polymers around room temperature (data not shown). This observation implies a further softening of the polymers as the temperature is increased beyond T_g.

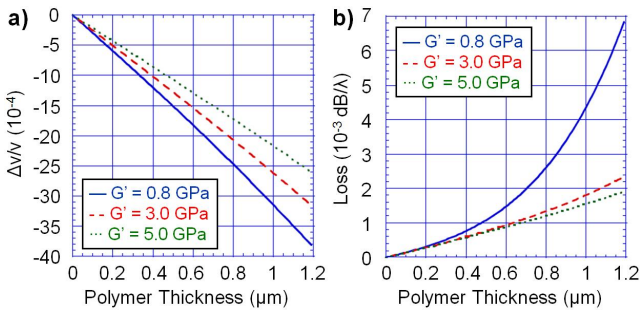


Fig. 2: Calculated change in acoustic wave velocity (a) and loss (b) as a function of polymer thickness for three different real parts, G', of the polymer shear modulus, G. The imaginary part was kept constant at $G'' = 0.8$ GPa.

Table II: Glass transition temperatures, T_g, for the chemically sensitive polymers used in this study [12], [14].

Polymer	T_g (°C)
PEA	-25…-22
PECH	-22…-15
PIB	-73

IV. Results and Discussion

Coated sensor devices were exposed to various concentrations of BTEX compounds in water to evaluate the performance of the various coating materials in the liquid phase. Sample results are shown in Fig. 3 for the detection of ethylbenzene using a 1.0-µm-thick PEA coating.

Assuming non-specific sorption of the analytes into the polymer, Henry's law should be valid at low concentrations [15], and the measured frequency shift should be linear with concentration. For the coating, type of analyte, and concentration range investigated in Fig. 3, this assumption appears to be valid. Measurements were conducted for a larger concentration range to test the limits of Henry's law. Fig. 4 shows the result for benzene and a PEA-coated device.

In Fig. 4, no systematic deviation from a linear fit can be observed (only a small and random scatter of the data due to manual pipetting in sample preparation is noticeable). For this example, as well as for other analytes and sensor coatings, Henry's law was always observed to be valid for

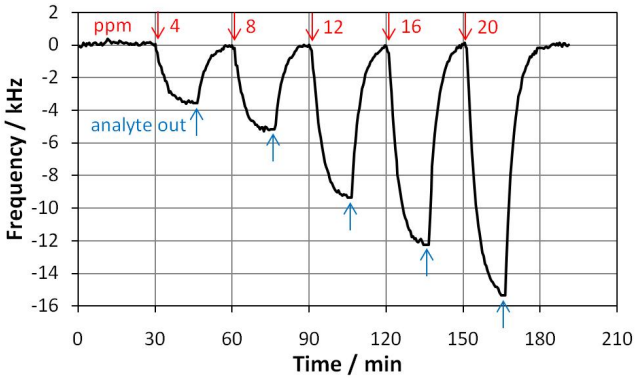


Fig. 3: Frequency response of a SH-SAW device coated with 1.0 µm PEA to various concentrations of ethylbenzene in water; concentrations are indicated on the graph (1 ppm = 1 mg/L). The analyte was flushed out with water between individual samples. The signal was corrected for baseline drift.

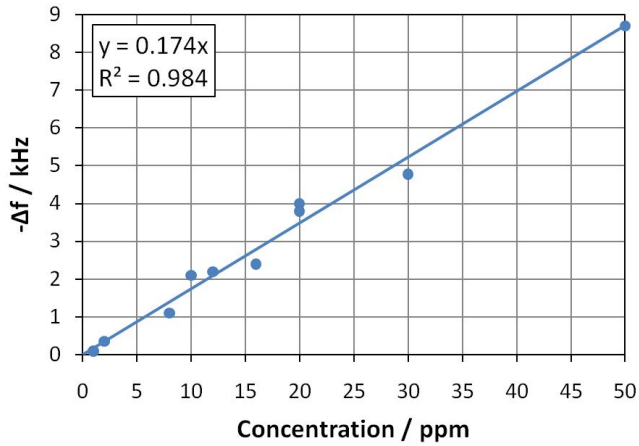


Fig. 4: Equilibrium frequency shift of a SH-SAW device coated with 1.0 µm PEA vs. benzene concentration in water (1 ppm = 1 mg/L). The blue line is a linear fit to the experimental data (circles) with slope indicated on the graph.

concentrations below 50 mg/L.

The ultimate goal of the design of the SH-SAW sensor system is to identify and quantify a given analyte in the presence of other contaminants. Therefore, the system needs a certain degree of selectivity. Non-specific analyte sorption is associated with partial selectivity, permitting analyte identification using an array of sensors with different coatings [16], [17]. An example of partial selectivity is reflected in the data presented in Table I. However, it should be noted that the partition coefficients depend on thermodynamic driving forces based on two input variables: the affinity of the coating for the analyte, and the solubility of the analyte in the liquid phase. In order to judge the partial selectivity of a given coating for an analyte, the measured frequency shift has to be divided by analyte solubility. In addition, since the sensor response is predominantly based on mass loading, the data should also be normalized with respect to analyte molar mass in order to compare data in terms of the number of molecules absorbed into a coating. If both normalization procedures are applied, a parameter is obtained which permits a thermodynamically more meaningful comparison of different coating materials:

$$\Delta f\,(\text{normalized}) \;=\; (\Delta f / M_W) \,/\, (\Delta C_{Amb} / C_S). \qquad (2)$$

In this equation, M_W is the analyte molar mass, and C_S is the analyte solubility; the other variables are defined as in Equation (1). Data for solubilities and molar masses of organic hydrocarbons have been taken from the literature [18].

Experimental results have been collected for the detection of 20 mg/L of the BTEX compounds and the three chemically sensitive polymers investigated. The data have been normalized according to Equation (2) and are presented in Fig. 5. Of the BTEX compounds, benzene is the most hazardous to human health. Therefore, coating materials have been selected with a focus on maximum partial selectivity for benzene, and the data in Fig. 5 have been normalized to the respective value for benzene for each coating material.

A useful concept to understand the interaction between two substances is the Hildebrand solubility parameter [19], which measures the cohesive energy density of a given substance, i.e. the amount of energy needed to remove a unit volume of molecules of this substance from their neighbors to infinite distance. The cohesive energy contains various contributions from polarity, polarizability, hydrogen bonding, etc. Chemical substances with similar solubility parameters will usually show good solubility in each other. Fig. 5 indicates that PEA shows the strongest partial selectivity for benzene. This is in agreement with the fact that PEA and benzene have almost identical Hildebrand solubility parameters [19].

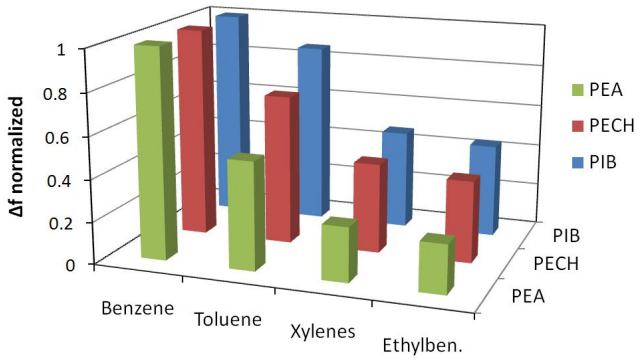


Fig. 5: Normalized frequency shifts for the BTEX compounds for sensors coated with 1.0 µm PEA, 0.6 µm PECH and 0.5 µm PIB, respectively. Data have been normalized as described in the text to extract the effect of partial selectivity of the polymers.

Fig. 5 also shows slightly different response patterns for the investigated polymers, a fact that can be exploited for pattern recognition. However, since this effect is not pronounced, this study also investigates the potential of the sensor response time to be an additional tool to identify a given analyte. To use response time as an input parameter for pattern recognition, an appropriate mathematical procedure must be identified to extract this parameter from experimental data. This has been done by modeling experimental response curves with an exponential fit using a state-space model [4]:

$$y(t) \;=\; y_0\,(1 - \exp(-t / \tau)). \qquad (3)$$

In this equation, $y(t)$ is the frequency change as a function of time, starting at the moment the analyte is added. The response time, τ, has been determined for 20 mg/L of each BTEX compound. Results are given in Table III for two polymer coatings.

Note that the coating thickness has to be known because it affects the value of τ. On the other hand, it was observed that τ is independent of analyte concentration within experimental error. Also note that benzene has a very short response time, a fact that might be useful for rapid extraction of benzene concentration from the sensor array response in the presence of other contaminants. In light of these findings, both equilibrium frequency change and response time were evaluated for their dependence on ambient parameters.

Fig. 6 shows an example of the influence of ambient temperature on the frequency response of a device coated with 0.5 µm PIB. The change in device loss is also presented, as it will be helpful in the interpretation of the results. The equilibrium frequency change is not much affected by ambient temperature; however, the latter does affect both the response time and the change in device loss. Fig. 6b shows that the change in loss increases with temperature, a result which is in agreement with the assumption that the polymer gets softer with increasing temperature as discussed in section III. The increase in softness is also expected to facilitate analyte diffusion into the polymer, thereby reducing response time.

Table III: Response times, τ, extracted from experimental data for the detection of 20 mg/L of the BTEX analytes with two polymer coatings, 1.0 µm PEA and 0.6 µm PECH, respectively.

Analyte	τ_{PEA} **(s)**	τ_{PECH} **(s)**
Benzene	63	52
Toluene	98	92
Xylenes	262	241
Ethylbenzene	265	216

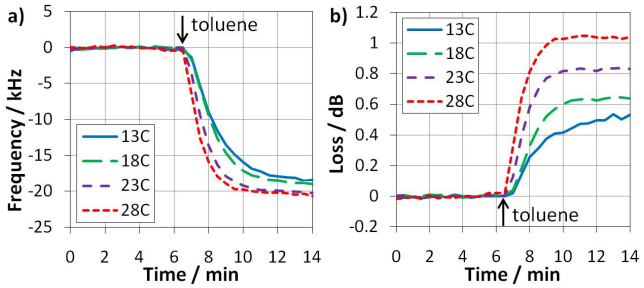


Fig. 6: Frequency shift (a) and change in device loss (b) for detection of 200 mg/L toluene for a sensor coated with 0.5 µm PIB, measured at four different temperatures as indicated on the graphs (in °C). The signal was corrected for baseline drift.

In order to quantify the effect of temperature on both frequency shift and response time, response curves measured at different temperatures were modeled with exponential fits according to Equation (3), and data for equilibrium frequency shift, y_0, and τ were extracted. Results for the detection of 10 mg/L toluene using a device coated with 0.6 µm PECH are shown in Fig. 7; the extracted fitting parameters are listed in Table IV.

Table IV shows that while y_0 is only slightly affected by temperature, τ varies by more than a factor of two in the temperature range investigated. Similar findings were made for the other investigated chemically sensitive polymers.

An additional set of experiments was conducted where the pH value of the aqueous phase was varied between 5.5 and 8, and the BTEX compounds were detected at each respective pH value. It was shown that pH had no significant impact on either frequency shift or response time (results not shown). In another set of experiments, the salinity of the aqueous phase was varied by adding NaCl in concentrations of 0 to 3%. Again, it was found that salinity has no significant influence on either frequency shift or response time in BTEX analyte detection. This is important if the device is to be used for detection in contaminated ocean or brackish water sites.

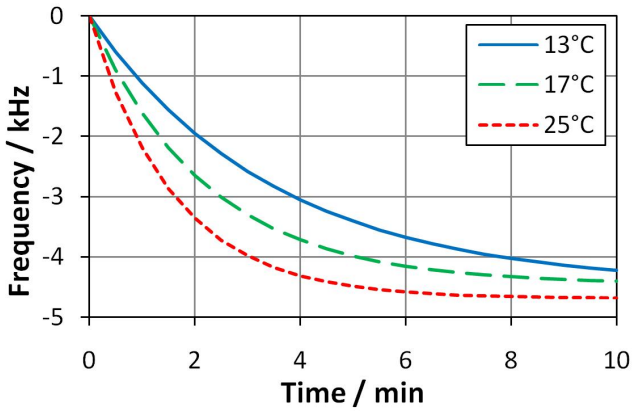


Fig. 7: Exponential fits for the frequency-vs.-time responses measured for detection of 10 mg/L toluene using a device coated with 0.6 µm PECH, measured at three different temperatures as indicated on the graph.

Table IV: Response times, τ, and equilibrium frequency shifts, y_0, extracted from the exponential fits shown in Fig. 7 for detection of 10 mg/L toluene using a device coated with 0.6 µm PECH.

T_{Amb} (°C)	τ (s)	y_0 (kHz)
13	210	-4.48
17	133	-4.45
25	95	-4.69

V. Conclusion

A system using polymer-coated SH-SAW sensor devices was presented, which is under investigation for the detection and monitoring of BTEX compounds in groundwater or contaminated sites. It was shown that in addition to using the partial selectivity of the polymer coatings, it will be useful to evaluate the response time as an additional sensor parameter for pattern recognition. Since the sensor system is going to be used *in situ*, the influence of various ambient parameters on both the frequency shift and response time of the sensors has been investigated.

It is known that the temperature of groundwater and contaminated water sites depends on geographic location. In the US (Alaska and Hawaii not included), near-surface groundwater temperatures vary between about 5 °C and 25 °C [20]. In addition, the acidity and salinity of groundwater and contaminated water can vary significantly. It was shown that the influence of all three ambient parameters investigated on the equilibrium frequency shift is small. Therefore, the latter is regarded as a robust parameter under field conditions. Within the relevant range, groundwater acidity and salinity also have little influence on response time for the polymers investigated. In contrast, it was shown that ambient temperature significantly affects the response time of polymer-coated sensors.

It was observed that both the baseline acoustic wave loss and the change in loss upon analyte sorption increase with temperature for the investigated chemically sensitive polymers, indicating a softening of these polymers with increasing temperature. Theoretical calculations [12] show that a decrease in shear modulus of a polymer can under certain conditions lead to a large change in acoustic loss but only a minor change in frequency, in agreement with experimental observations in this study. The softening of the polymer is also consistent with the faster response times observed for higher temperatures [21], [22].

The above findings suggest two alternative routes for the design of a sensor system for monitoring of groundwater or contaminated sites. A sensor array with pattern recognition based solely on using equilibrium frequency shift data would be the more straightforward approach because the influence of all the investigated ambient parameters on frequency shift is insignificant. However, such a system would have to rely on the limited partial selectivity of the polymer coatings for analyte identification. This might result in a decrease both in reliability in analyte identification and in accuracy in analyte quantification, and likely would preclude accurate analysis of analyte mixtures.

Therefore, a second approach is proposed that uses both equilibrium frequency shift and response time as sensor parameters for pattern recognition. It has been shown that response time is a powerful parameter in analyte identification (see Table III). In this case, the design of the sensor system requires that ambient temperature be monitored and used as an input parameter to compensate for the dependence of response time on temperature. This could be achieved by adding a small thermal sensor to the sensor array, and by modifying the pattern recognition software to take into account this additional parameter. It is expected that this approach will lead to better performance of the sensor system in analyte identification and ultimately will facilitate mixture analysis.

ACKNOWLEDGMENT

The authors would like to thank Karen A. Synowiec, Rachel Mohler, Urmas Kelmser and Arnold K. Mensah-Brown for helpful discussion.

REFERENCES

[1] C.K. Ho, A. Robinson, D.R. Miller and M.J. Davis, "Overview of sensors and needs for environmental monitoring", *Sensors,* vol. 5, pp. 4–37, 2005.

[2] F. Josse, F. Bender and R.W. Cernosek, "Guided shear horizontal surface acoustic wave sensors for chemical and biochemical detection in liquids", *Anal. Chem.,* vol. 73, pp. 5937–5944, 2001.

[3] Z. Li, Y. Jones, J. Hossenlopp, R. Cernosek and F. Josse, "Analysis of liquid-phase chemical detection using guided shear horizontal-surface acoustic wave sensors", *Anal. Chem.,* vol. 77, pp. 4595–4603, 2005.

[4] A.K. Mensah-Brown, M.J. Wenzel, F.J. Josse and E.E. Yaz, "Near real-time monitoring of organophosphate pesticides in the aqueous-phase using SH-SAW sensors including estimation-based signal analysis", *IEEE Sens. J.,* vol. 9, pp. 1817–1824, 2009.

[5] J.W. Grate et al., "Determination of partition coefficients from surface acoustic wave vapor sensor responses and correlation with gas–liquid chromatographic partition coefficients", *Anal. Chem.,* vol. 60, pp. 869–875, 1988.

[6] J.W. Grate et al., "The predominant role of swelling-induced modulus changes of the sorbent phase in determining the responses of polymer-coated surface acoustic wave vapor sensors", *Anal. Chem.,* vol. 64, pp. 610–624, 1992.

[7] E.T. Zellers and M. Han, "Effects of temperature and humidity on the performance of polymer-coated surface acoustic wave vapor sensor arrays", *Anal. Chem.,* vol. 68, pp. 2409–2418, 1996.

[8] J.S. Arey and P.M. Gschwend, "Estimating partition coefficients for fuel–water systems: developing linear solvation energy relationships using linear solvent strength theory to handle mixtures", *Environ. Sci. Technol.,* vol. 39, pp. 2702–2710, 2005.

[9] Y.K. Jones, Z. Li, M.M. Johnson, F. Josse and J.M. Hossenlopp, "ATR-FTIR spectroscopic analysis of sorption of aqueous analytes into polymer coatings used with guided SH-SAW sensors", *IEEE Sens. J.,* vol. 5, pp. 1175–1184, 2005.

[10] Y.K. Jones, "Spectroscopic and computational investigation of polymer coatings and analyte systems for use with guided shear horizontal surface acoustic wave (SH-SAW) sensors for liquid phase detection", dissertation, Marquette University, Milwaukee, WI, 2005.

[11] M.V. Voinova, M. Rodahl, M. Jonson and B. Kasemo, "Viscoelastic acoustic response of layered polymer films at fluid-solid interfaces: continuum mechanics approach", *Phys. Scr.,* vol. 59, pp. 391–396, 1999.

[12] Z. Li, "Guided shear-horizontal surface acoustic wave (SH-SAW) chemical sensors for detection of organic contaminants in aqueous environments", dissertation, Marquette University, Milwaukee, WI, 2005.

[13] J.J. Aklonis and W.J. MacKnight, "Time – temperature correspondence", Introduction to polymer viscoelasticity, 2nd edition, New York: John Wiley and Sons, 1983, pp. 36–56.

[14] J. Brandrup, E.H. Immergut and E.A. Grulke, "Glass transition temperatures of polymers", Polymer handbook, 4th edition, vol. I, New York: John Wiley and Sons, 1999, pp. VI/193–VI/277.

[15] A. Hierlemann, A.J. Ricco, K. Bodenhöfer and W. Göpel, "Effective use of molecular recognition in gas sensing: results from acoustic wave and in situ FT-IR measurements", *Anal. Chem.,* vol. 71, pp. 3022–3035, 1999.

[16] E.T. Zellers, S.A. Batterman, M. Han and S.J. Patrash, "Optimal coating selection for the analysis of organic vapor mixtures with polymer-coated surface acoustic wave sensor arrays", *Anal. Chem.,* vol. 67, pp. 1092–1106, 1995.

[17] A.J. Ricco, R.M. Crooks and G.C. Osbourn, "Surface acoustic wave chemical sensor arrays: new chemically sensitive interfaces combined with novel cluster analysis to detect volatile organic compounds and mixtures", *Acc. Chem. Res.,* vol. 31, pp. 289–296, 1998.

[18] D.L. Lide, "Aqueous solubility and Henry's law constants of organic compounds", Handbook of chemistry and physics, 82nd edition, Boca Raton: CRC Press, 2001–2002, pp. 8–95.

[19] W.L. Archer, Industrial solvents handbook, New York: Marcel Dekker, 1996.

[20] T.E. Gass, "Geothermal heat pumps", *Geotherm. Resour. Counc. Bull.,* vol. 11, pp. 3–8, 1982.

[21] J.W. Grate and D.A. Nelson, "Sorptive polymeric materials and photopatterned films for gas phase chemical microsensors", *Proc. IEEE,* vol. 91, pp. 881–889, 2003.

[22] A.K. Mensah-Brown, D. Mlambo, F. Josse and J. Hossenlopp, "Rapid detection of organophosphates in aqueous solution using a hybrid organic/inorganic coating on SH-SAW devices", *Proc. 2010 IEEE Int. Freq. Control Symp.,* pp. 232–237, 2010.

Surface Acoustic Wave Flow Sensor

Yizhong Wang[1], Zheng Li[2], Lifeng Qin[1], Minking K. Chyu[1], Qing-Ming Wang[1*]
[1] Department of Mechanical Engineering and Materials Science
University of Pittsburgh, Pittsburgh, PA 15261 U.S.A.
[2] Department of Mechanics, Peking University, Beijing, China
* qmwang@engr.pitt.edu

Abstract—**A novel surface acoustic wave flow sensor is studied in this paper. ST-X cut quartz SAW delay line device is chosen for its ultra-high temperature stability. Test chamber is designed and built using acrylic plate and the SAW device is mounted on the chamber wall. Test Results show that SAW flow sensor has a linear relationship between the passing flow rate and system phase delay. The system has a response resolution of 4.24Deg / kPa in terms of pressure or 1 degree phase change for flow rate change of every 11.8ml/min. The change of resonant frequency is found to be not as obvious as phase change. The system can be further integrated with wireless system for remote on site monitoring.**

I. INTRODUCTION

Fluid flow sensor has broad applications in various industrial and research operation monitoring and control. It's especially important to have accurate flow measurement in some aero-space technology applications and some of the cutting edge research projects.

There has been numerous flow sensors, some of them already commercialized. Traditional mechanical flow sensors, such as venturi flow meters [1] and pitot tube flow meters [2], do not require power supply, and have simple structures. However, the readouts of these traditional flow sensors are not compatible with electronic system, making them unsuitable for modern applications.

Modern electronic flow sensors can be divided into several categories based on operation principles. One of the most important types of flow sensor is thermal flow sensor [3], including hot-wires anemometers, calorimetric flow sensors. They usually have a heater and at least one temperature sensing element to measure, e.g. the heat transfer from the heater to the thermometer. These flow sensors can have broad measurement range, from micro fluid measurement to macroscopic petroleum transport pipes. It's also convenient to integrate these sensors into CMOS circuit same chip control. The other important type of flow sensor is ultrasonic flow sensor [4]. The sensing structure is clamped to the outside wall of the pipe, where ultrasonic wave travels through the flow and is reflected back through the flow and intercepted by the receiver.

Sponsored by NSF grant #0925716 and DOE grant #0002138.

The response of the ultrasonic wave, including resonant frequency, magnitude and phase change depends on the flow rate through the pipe. However, the power consumption of these sensors makes them hard to be used in portable devices or remote areas where power supply is a challenge.

For pressure driven laminar flow, as shown in Fig. 1, the pipe wall is subject to pressure along the flow direction. The pressure on the wall decreases linearly towards the exit.

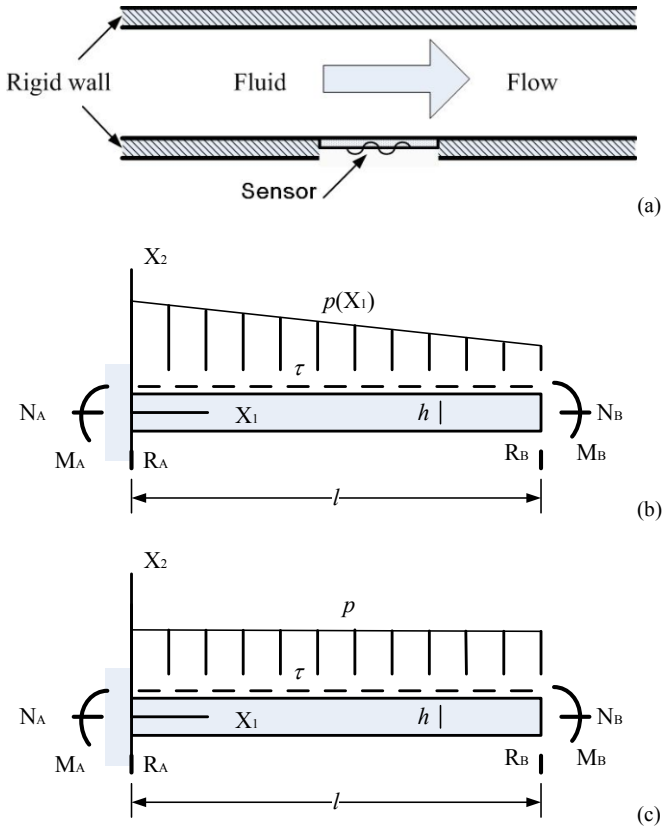


Fig. 1. Schematic view and Physics model of flow sensor. (a) schematic view of the pipe and sensor. (b) Sensor is subject to linearly decreased pressure. (c) Normal pressure can be simplified as uniform for small sensor surface area.

When a sensor is mounted on the pipe wall, the sensor will also be subject to linearly decreased pressure. For sensor with small surface area compared to pipe dimension, the pressure can be assumed to be uniform as shown in Fig. 1(c).

Previous research [5] showed that, for laminar flow, the pressure difference between specific point on pipe wall and flow exit depended linearly on flow rate. The normal pressure and shear stress on the pipe wall can be calculated as equation (1) and (2).

$$\Delta P = \int_{L/2}^{0} \frac{dp}{dx} dx + \frac{8\mu L_p}{\pi R_p^4} Q = \frac{6Q\mu L}{h^3 w \left(1 - \frac{192h}{\pi^5 w} \sum_{n=1,3,5\ldots}^{\infty} \frac{1}{n^5} \tanh\left(\frac{n\pi w}{2h}\right)\right)} + \frac{8\mu L_p}{\pi R_p^4} Q, \quad (1)$$

$$\tau_r = \frac{h}{2}\left(-\frac{dp}{dx}\right)\left[1 - \frac{8}{\pi^2} \sum_{n=1,3,5\ldots}^{\infty} \frac{1}{n^2} \frac{\cosh\left(\frac{n\pi z}{h}\right)}{\cosh\left(\frac{n\pi w}{2h}\right)}\right]_{z=0} = \frac{6\mu Q}{h^2 w} \frac{\left[1 - \frac{8}{\pi^2} \sum_{n=1,3,5\ldots}^{\infty} \frac{1}{n^2 \cosh\left(\frac{n\pi w}{2h}\right)}\right]}{\left[1 - \frac{192h}{\pi^5 w} \sum_{n=1,3,5\ldots}^{\infty} \frac{1}{n^5} \tanh\left(\frac{n\pi w}{2h}\right)\right]}. \quad (2)$$

The calculated shear stress on the pipe wall is three order magnitudes less than the normal stress thus can be ignored.

Since flow rate depends linearly on pressure, the flow meter essentially functions as a pressure sensor.

To facilitate the integration of wireless module, surface acoustic wave (SAW) is chosen as the sensing method. SAW sensor has been widely used for its low power consumption and easy operation and compatibility with wireless.

There has been some research on SAW pressure sensor design and fabrication [6]. However, the fabricated sensor is based on $LiNbO_3$ and has two-layer chamber structure. Since $LiNbO_3$ is temperature dependant, the pressure sensor is unstable without proper temperature compensation. The two-layer structure also makes it unsuitable for flow sensor application.

ST-X cut quartz is selected for flow sensor due to its excellent temperature stability. Plate structure is used for fabrication simplicity.

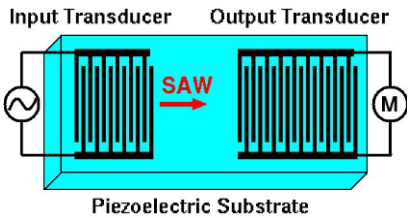
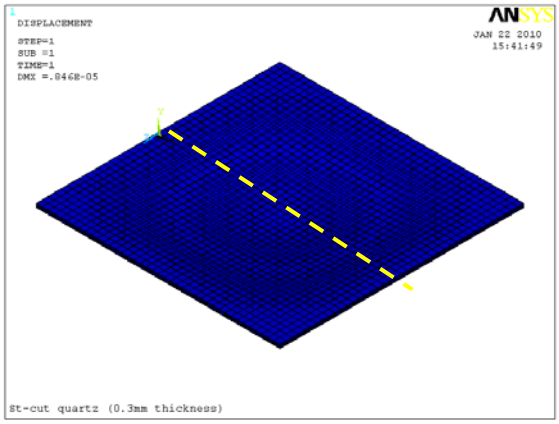
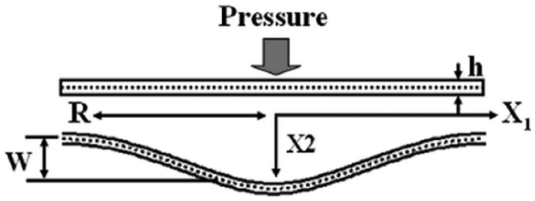
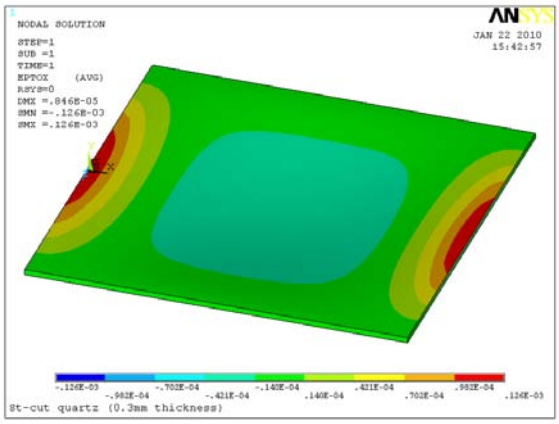


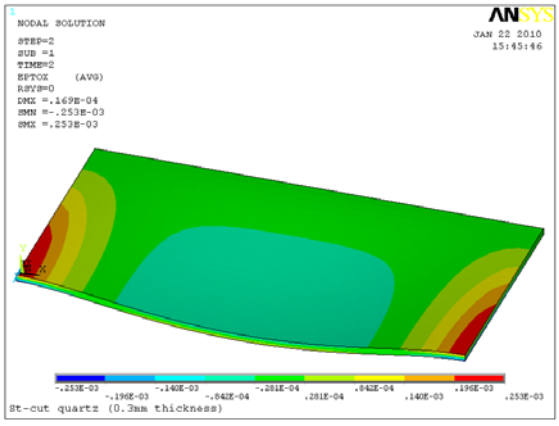
Fig. 2. Two ports SAW sensor structure





(a)





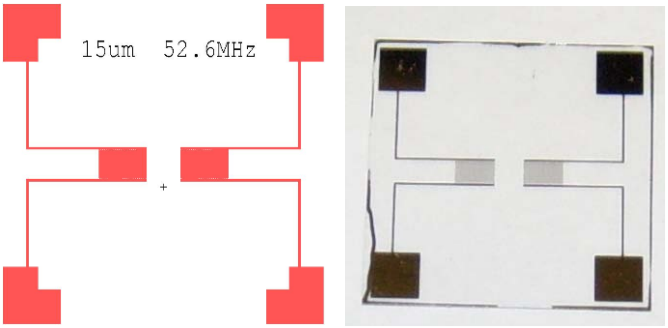
(b)

Fig. 3. Simulation of the substration deformation under pressure.

II. DEVICE DESIGN AND ANALYSIS

A. Sensor Design

Traditional two port SAW sensor structure is used in this study. As shown in Fig. 2, the excitation inter-digital-transducer (IDT) generate SAW under alternate current, the wave travels along the surface of the substrate and intercepted by the receiving IDTs. Under normal pressure, the substrate will bend and elongate. The elongation of the substrate will cause the increase of the wave travel time. The travel time change can be measured using network analyzer, and the travel time change will be further interpreted into phase change. So phase change will be shown for different normal pressure, i.e. different flow rate.



(a) (b)

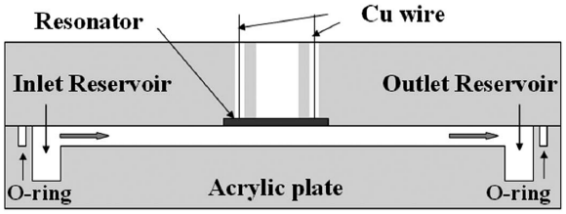
Fig. 4. Designed and fabricated Sensor top view

To determine the location of the IDTs and space between the input and output IDTs, it's necessary to simulate the deformation under pressure by finite element method (FEM). Ansys® is used to simulate the response of the substrate. As shown in Fig. 3, the plate will bend under pressure.

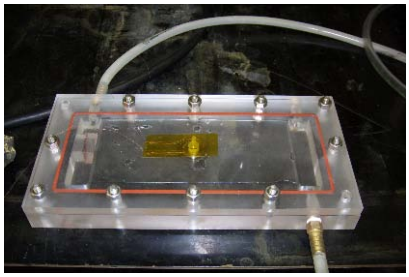
Since SAW signal will damp significantly in water and other liquids, it's natural that IDTs are located on the back of the substrate, thus SAW can propagate on the back of substrate. As shown in Fig. 3(a), along the symmetrical line of the substrate, the area close to the clamped end will be subject to compression and the area on the center will be subject to tension. The IDTs are finally selected in the manner that they will have minimal deformation under pressure and the SAW propagation path is entirely located on tension area. The designed sensor has 70 pair of input and output IDTs and 3mm apart. The resonant frequency is calculated as 52.63MHz assuming SAW velocity on ST-X cut quartz is 3158m/s. The substrate under pressure is 2cm by 2cm. Fig. 4(a) shows the sensor design over view.

B. Sensor Fabrication

A 50nm thick Cr layer was deposited on the Quartz wafer. Then a 150nm thick Au layer was sputtered on top of Cr layer. After that a 3um thick AZ4210 was coated on Au surface by spin-coating at 3000rpm for 30 seconds. After patterning, the Au and Cr on IDT area were kept and others were removed by Au etchant and Cr etchant. The remaining photoresist is then removed by acetone. Finally the wafer is cleaned using IPA and DI water and then dehydrated for 30min at 120℃. Silver paste is used to connect the IDT electrodes to the measurement instrument. Fig. 4(b) shows the fabricated sensors.



(a)



(b)

Fig. 5. Designed and completed test setup.

C. Test Setup

A rectangular flow chamber is designed with dimension of 165.10 mm × 69.85 mm × 0.51 mm (L × W × H) and is composed of 2 parallel acrylic plates. The schematic view of the chamber, which includes a rectangular flow channel and an inlet and outlet reservoir, a sensor mounting port, and an inlet port and an outlet port, is shown as Fig. 5(a). Fig. 5(b) shows the completed setup.

III. EXPERIMENT RESULT AND DISCUSSION

The SAW flow sensor is tested under different flow rates from 0ml/min up to 3500ml/min, which corresponds to normal pressure of 0Pa up to 60kPa. Using the group delay measurement function from network analyzer, the delay time and the corresponding resonant frequency of the sensor can be recorded simultaneously. When the flow rate increases from stagnant, the delay time changes accordingly. This increase of delay time is then converted into phase change after dividing by the period of the SAW. From this aspect, the phase change is linearly dependant on delay time change.

From Fig. 6, it can be seen that the phase change of the sensor is linearly depending on the fluid flow rate at 1Deg/11.8ml/min, while the resonant frequency does not show obvious change along flow rate change, probably because the choice of IDT locations minimize the deformation of IDTs.

IV. CONCLUSION

SAW flow sensor was built on ST-X cut quartz substrate. FEM simulation was done to determine the deformation of the crystal plate under flow pressure and guide the fabrication of the IDTs. Test results show linear relationship between the flow rate and delay time change, i.e. phase change.

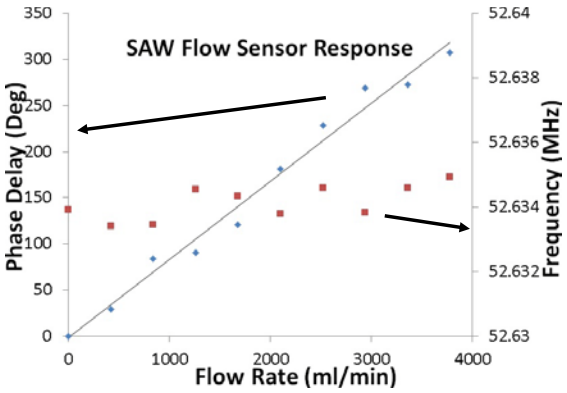


Fig. 6. Experiment Result

REFERENCES

[1] O. Wigertz, "A low-resistance flow meter for wide-range ventilatory measurement," Respiration Physiology, vol. 7, issue 2, pp. 263–270, August 1969.

[2] J. H. Preston, "The three-quarter radius pitot tube flow meter," The Engineer, vol. 190, pp. 400-403, 1950.

[3] M. Ashauer, H. Glosch, F. Hedrich, N. Hey, H. Sandmaier and W. Lang, "Thermal flow sensor for liquids and gases based on combinations of two principles," Sensors and Actuators A: Physical, vol. 73, issue 1-2, pp. 7-13, 9 March 1999.

[4] Richard S. Figliola and Donald E. Beasley, Theory and Design for Mechanical Measurements, 4th ed., John Wiley & Sons, New York, 2006, pp. 412-413.

[5] Lifeng Qin, Zijing Zeng, Hongbin Cheng and Qing-ming Wang,, "Acoustic wave flow sensor using quartz thickness shear mode resonator," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 56, issue 9, pp. 1945-1954, September 2009.

[6] Haekwan Oh, Weng Wang, Keekeun Lee, Ikmo Park, and Sang Sik Yang, "Sensitivity Improvement of Wireless Pressure Sensor By Incorporating A SAW Reflective Delay Line," Intl. J. on Smart Sensing and Intelligent Systems, vol. 1, pp. 940–954, Dec. 2008.

Solidly Mounted Resonators with Carbon Nanotube Electrodes for Biosensing Applications

García-Gancedo L.[1], Iborra E.[2], Clement M.[2], Olivares J.[2], Zhu, Z.[1], Flewitt A.J.[1], Milne W.I.[1], Ashley G.M.[3], Luo J.K.[3], Zhao X.B.[4] and Lu J.R.[4]

[1] Electrical Engineering Division, University of Cambridge, Cambridge, United Kingdom
[2] Grupo de Microsistemas y Materiales Electrónicos, Universidad Politécnica de Madrid, Madrid, Spain
[3] IMRI, University of Bolton, Bolton, United Kingdom
[4] School of Physics & Astronomy, University of Manchester, Manchester, United Kingdom

***Abstract*— The work reported here shows a direct experimental comparison of the sensitivities of AlN solidly mounted resonators (SMR)-based biosensors fabricated with standard metal electrodes and with carbon nanotube electrodes. SMRs resonating at frequencies around 1.75 GHz have been fabricated, some devices using a thin film of multi-wall carbon nanotubes (CNTs) as the top electrode material and some identical devices using a chromium/gold electrode. Protein solutions with different concentrations were loaded on the top of the resonators and their responses to mass-load from physically adsorbed coatings were investigated. Results show that resonators using CNTs as the top electrode material exhibited higher frequency change for a given load due to the higher active surface area of a thin film of interconnecting CNTs compared to that of a metal thin film electrode and hence exhibited greater mass loading sensitivity. It is therefore concluded that the use of CNT electrodes on resonators for their use as gravimetric biosensors is viable and worthwhile.**

I. INTRODUCTION

Gravimetric sensing is a well known technique utilised to detect the variation in a measurable parameter as a function of the deposited mass on an active sensing area [1-3]. Quartz crystal microbalances (QCM) resonating in the range of 5-20 MHz are one of the earliest, most successful examples of gravimetric sensors. Historically, the emergence of the Quartz crystal as a microbalance can be traced back to Sauerbrey, who realised that a resonating quartz crystal could be used for the microgravimetric measurement of sputtered thin films [4]. A few years later, King chemically sensitised a QCM to make selective chemical mass measurements of 10^{-9}g [5]. Since then, there has been an increased interest in developing gravimetric sensors with higher sensitivities for their use as biosensors. Mass loading sensitivity is proportional to the square of the operating frequency [4]; hence the emergence of bulk acoustic wave resonator (BAW) technology, a nanoscale version QCM with an operating frequency in the range of a few GHz, offered great potential to improve the sensitivity and detection limitation of gravimetric sensors. BAW resonators offer the additional advantage of lower fabrication cost due to the employment of microfabrication technologies and small dimensions, which makes them suitable for the fabrication of array of sensors for parallel or multi detection.

However, higher frequency and smaller dimensions cause all environmental and boundary conditions to affect the operation of the BAW resonators more strongly than the QCMs, resulting in a higher noise level and lower quality factors. It is for this reason that up to now, mass sensitivity of BAW resonators is not significantly better than that of QCMs [6-8].

Work has been carried out in the last few years towards the improvement of the structural and electrical properties of piezoelectric thin films for their use as the active material in these devices [9-12], with the aim of improving the electrical response of the resonators through the achievement of materials with improved piezoelectric properties. Recently, it has also been understood that the electrode material also plays an important role in the resonator's electrical response [13-16]. Electrodes not only affect the growth of the piezoelectric thin film but also provide the shape of the resonant area and, if designed properly, can confine the acoustic energy between the facets of the resonator thereby minimising losses and maximising electro-mechanical transduction (hence *Q*). Ideally, the electrode material should have low mass (in order to minimise mass-loading effects) but high acoustic impedance (in order to confine the acoustic waves within the piezoelectric layer) and high conductivity (in order to minimising the series resistance in the transmission of the excitation signal). Aluminium (Al), molybdenum (Mo)

The work reported here was supported by the EPSRC, grants number EP/F063865/1, EP/F06294X/1, EP/F062966/1 and EP/E023614/1. Iborra, Clemet and Olivares acknowledge financial support from the Ministerio de Ciencia e Innovación of Spain through project MAT2010-18933, and from the European Union through the European Regional Development Fund (FEDER).

tungsten (W), gold (Au) and platinum (Pt) are the most common metals utilised as the electrode materials. However the low acoustic impedance of Al and high weight of Mo, W, Au and Pt make them less ideal for their use as the electrode materials.

Carbon nanotubes (CNTs) have been shown to be an excellent material choice for BAW resonators [15-16]. They possess low densities in the range of 1-2 g·cm^{-3}, electrical conductivities of up to 10^6 S·m^{-1} and exceptionally high elastic moduli (hence high acoustic impedance), usually higher than 1 TPa [17]. Fig. 1 shows a comparison of densities and acoustic impedances of some of the commonly used metals for electrodes, and highlights that CNTs posses the highest acoustic impedance and lowest density allowing simultaneously low mass electrodes with high acoustic impedance and low series resistance.

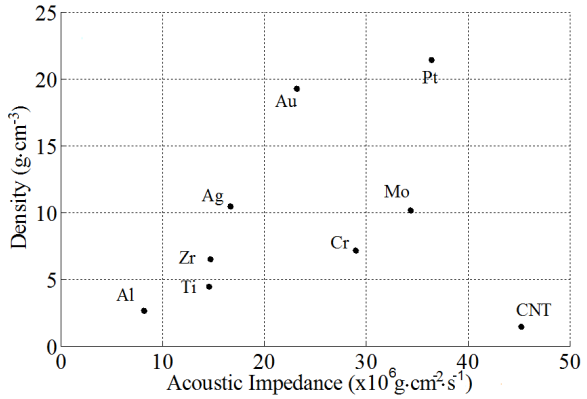


Figure 1. Density and acoustic impedance of commonly used electrode metals

Biosensors with CNT electrodes will offer an additional benefit compared to those with metal electrodes as it is possible to post-growth functionalise them for direct bonding to targeted molecules, thereby preventing non-specific binding or the need of additional bonding layers. Hence the potential exists to use chemically modified CNTs which would double both as the electrode and the sensing layer. Recent advances in chemical functionalisation of CNTs in sensing systems have been reported in [18-23]. Furthermore, the surface area of a thin film of interconnecting CNTs is significantly higher than the surface area of a metal thin film electrode; hence the binding area is increased without physically increasing the size of the devices. Consequently, for the same concentration of biological samples, a greater number of targeted molecules can be bound onto the resonator's surface compared to devices with metal electrodes, and hence devices with higher sensitivity can be realised.

The work reported here shows a direct experimental comparison of the sensitivity of AlN SMRs biosensors with standard metal electrodes and with CNT electrodes. SMRs resonating at frequencies around 1.75 GHz have been fabricated, some devices using a thin film of multi-wall CNTs as the top electrode material, and some identical devices using a top chromium/gold electrode. The CNTs were directly grown on the AlN film from an iron seed layer at 650°C in a NH_3 and C_2H_2 atmosphere, and possess very low density (<2 g cm^{-3}), high electrical conductivities (>10^5 S/m), and exhibited excellent definition and adhesion to the surface. Bovine Serum Albumin (BSA) solutions with different concentrations were loaded on the top of the resonators and their responses to mass-load from physically adsorbed protein coatings were investigated.

II. Experimental Methods

A. AlN SMR fabrication

SMRs were fabricated by growing Ir/AlN/CNT stacks on top of acoustic mirrors centred at a frequency of around 2 GHz and formed by five alternating low and high acoustic impedance layers of porous SiO_2, with an acoustic impedance of 9.5×10^6 kg·m^{-2}·s^{-1} [24], and Ir, with an acoustic impedance of 138.0×10^6 kg·m^{-2}·s^{-1} [25]. A seed layer of Ti/Mo (15-20 nm thick) was used to improve the adhesion and crystal quality of the Ir bottom electrode to the acoustic reflector. Ir slugs (99.98% pure) were electron-beam evaporated at a base pressure of less than 1×10^{-6} Torr to form bottom electrodes 120 nm thick. The AlN piezoelectric films were then reactively sputtered from a 150 mm in diameter Al target in a Leybold Z-550 sputtering system. Before AlN deposition, a 60 s soft-etch in an Ar discharge was performed to condition the Ir bottom electrode. AlN films were sputtered with a 3:7 Ar/N_2 admixture at a total pressure of 1.2 mTorr, a pulsed DC power level of 1.2 kW, and a substrate temperature of 400º C. An RF bias was applied to the substrates to tune the stress in the AlN films and to improve their crystal structure and orientation. These conditions yielded a deposition rate of 40 nm·min^{-1}. The thickness of the AlN layer was set to 1550 nm.

The crystal structure of the AlN and Ir films was assessed by X-ray diffraction (XRD) through the measurement of the θ/2θ patterns and the rocking-curves (RC) around the most intense reflections, i.e. the AlN 00·2 at 18.02º (θ) and the Ir (111) at 20.35º (θ). The surface of the different layers was inspected by atomic force microscopy (AFM), which provided the measurement of the surface roughness. Additional information on the polar orientation was obtained by wet chemical etching of the AlN layers.

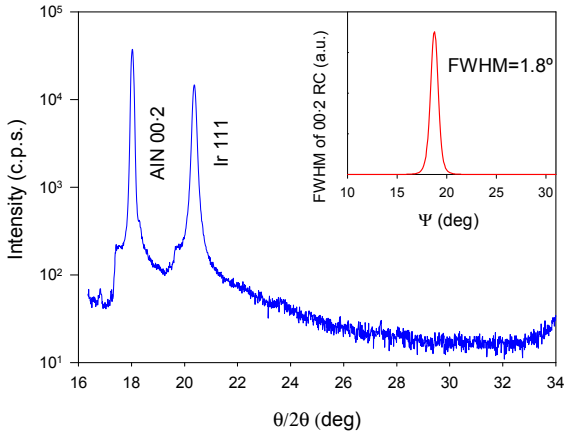
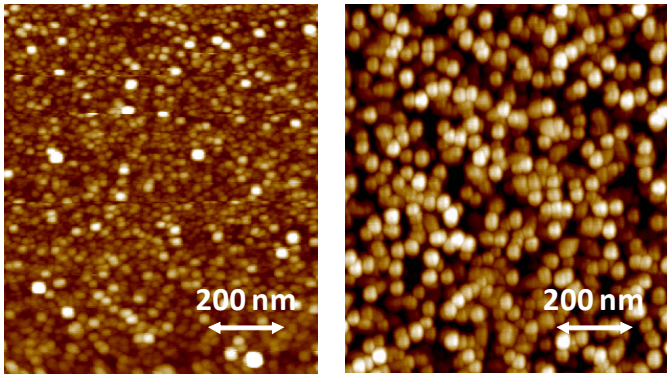


Figure 2. XRD patterns of AlN on Ir. The inset show the rocking curve (RC) of the 00·2 AlN peak.

XRD characterisation showed that highly *c*-axis-oriented AlN films were successfully grown on (111)-oriented Ir layers, provided these surfaces were etched with Ar+ ions immediately before AlN deposition (see Fig. 2). The pre-treatment of the Ir surface was essential to avoid the growth of tilted microcrystals related to the degradation of the piezoelectric response; AlN films with pure 00·2 orientation and RC ranging from 1.8º to 3º, almost independent of the texture and roughness of the Ir bottom electrode, were obtained. No degradation of the piezoelectric quality of AlN films with the texture was observed in the mentioned range.

Fig. 3 shows atomic force microscopy (AFM) images of the surface of Ir electrode (before AlN deposition) and AlN films. The roughness of the Ir electrode is around 4 nm and that of AlN films is around 7 nm. The structure of the AlN surface indicates a *c*-axis orientation of the microcrystals with a basal plane diameter of around 50 nm.



(a) (b)

Figure 3. AFM images of the surfaces of : (a) Ir bottom electrode before AlN deposition and (b) AlN film.

After AlN deposition, vias were formed through the AlN for electrical connection to the bottom electrode by wet etching the AlN in a warm (40°C) KOH solution through a Mo mask. The top electrode was then patterned with a lift-off photolithography process. At this point, a 10 nm of Cr layer followed by a 90 nm of Au layer was evaporated on one set of devices. A 7 nm of Fe layer, which acts as a catalyst layer for the CNTs growth, was evaporated on another set of devices. The CNTs were directly grown on the second set of the devices, as explained in the following section.

Fig. 4 shows the top view of one of the devices fabricated with top CNTs electrode. It can be seen that the definition of the electrodes is excellent.

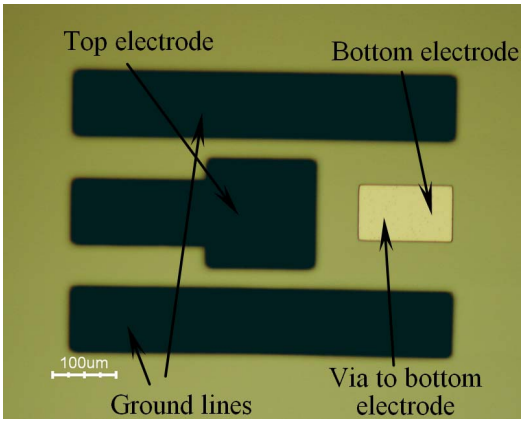


Figure 4. Image of a fabricated SMR with CNT top electrode

B. *CNT growth and characterisation*

CNT growth was carried out by thermal CVD. The substrate was loaded into the chamber and heated up to 650 °C at a rate of 4°C/min. Ammonia (NH_3) was introduced into the chamber to raise the pressure up to 5 mbar. It also helped to maintain the catalyst metal surface activity by reacting with the amorphous carbon. The system was maintained for several minutes so as to break the catalyst layer into nanoislands, after which acetylene (C_2H_2) was introduced into the chamber to commence the growth. C_2H_2 is the source of carbon necessary for CNT growth, and was introduced into the chamber through a large, flat, stainless steel plate 'showerhead' configuration, which ensures a uniform flow of gas over the whole area of growth. Once the growth is finished, the gases were evacuated and the system was left to cool down. This type of growth produces a forest of interconnecting nanotubes that should possess optimal electromechanical properties [17]. However if no particular care is taken, CVD generally produces CNTs with numerous defects. A high density of defects are favoured when the growth occurs at low temperatures, dopants such as N_2 or B are present or a residual growth process continues when the power is shutdown, but the gases are not turned off.

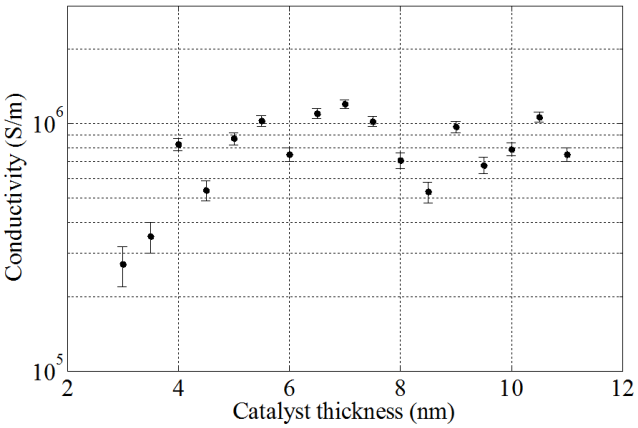


Figure 5. Conductivity of the CNTs layer *vs*. catalyst layer thickness for growth temperature of 750 ºC

The gas ratio C_2H_2 / NH_3 and growth temperature were optimised so as to avoid the presence of amorphous carbon, which would decrease significantly the conductivity. The catalyst layer thickness was varied in order to clarify its effect on the conductivity of the CNT layer. It was found that the conductivity increases with the catalyst thickness up to ~5nm, and become less thickness dependent afterwards as shown in Fig. 5.

The growth temperature was also changed in order to maximise the conductivity. A higher growth temperature produces CNT layers with a higher conductivity, as shown in Fig. 6, due to lower number of defects. However when growing the CNTs on top of an acoustic mirror, the Ir layers on the reflector stack delaminate, damaging a large number of devices. For this reason, when the growth temperature is 750 ºC, the yield is only around ~10%. Lowering the growth temperature to 650 ºC decreases slightly the conductivity but increases the yield to around ~20%. Growth does not occur at temperatures below ~600 ºC under these experimental conditions.

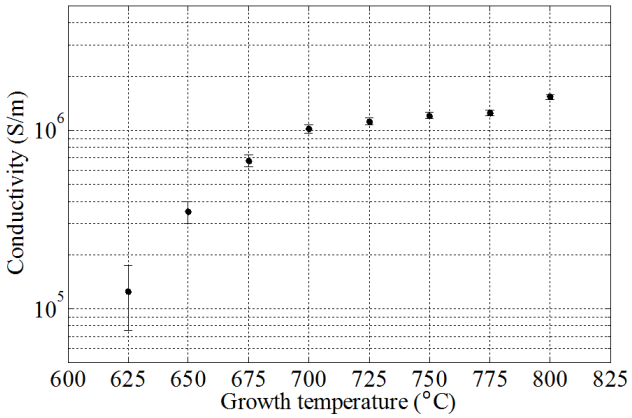


Figure 6. Conductivity of the CNTs layer vs. growth temperature

In this work CNTs were grown at 650 ºC with a catalyst thickness of 7 nm. CNT layers obtained possess an electrical conductivity of ~3.5×10^5 S/m.

III. FBAR Electrical Characterisation

Electrical characterisation of the SMR devices was carried out using an Agilent E5061B network analyser. An automation measurement routine was written with LabView to continuously monitor and record the resonance spectrum of the SMRs. Devices resonating at around 1.75 GHz, with resonance spectra that were free from significant spurious modes, were achieved and a typical response of the devices with CNT electrodes is shown in Fig. 7. It was also found that the resonant frequency of the devices covered with CNT electrodes is slightly higher due to the lower mass of the CNT electrodes compared to the Cr/Au electrodes.

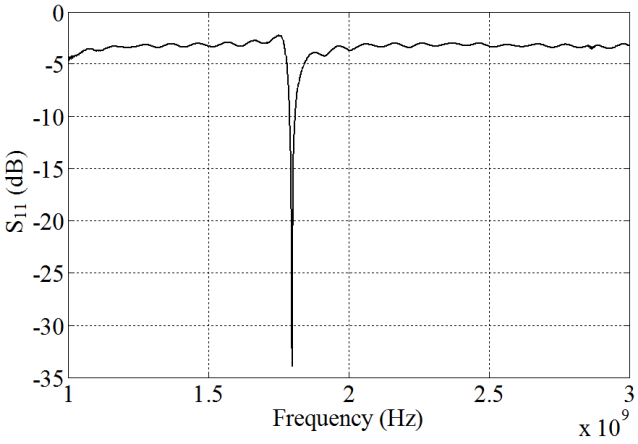


Figure 7. Typical frequency response of fabricated devices with CNT top electrode, showing the main resonance at f_0 ~ 1.75 GHz.

The quality factor of the resonance peak was calculated using the 3-dB method [26] from the equation:

$$Q_{3dB} = \frac{f_r}{|f_2 - f_1|} \tag{3}$$

where f_r is the resonant frequency taken from the S-parameter and f_1 and f_2 are the frequencies at which the insertion/return loss are +3 dB compared to the insertion/return loss at f_r.

Q was found to be higher than 2000 for the devices with CNT electrodes and around 1000 for the devices with metal electrodes. Generally devices with higher Q exhibit sharper resonance peaks that allow tracking a smaller frequency shift. Therefore, CNT electrode BAW resonators have the potential to measure smaller frequency shifts, increasing the sensitivity compared to the metal electrode devices.

IV. Preparation of Bio Targets on Devices and Biosensing Results

The resonant frequency of unloaded resonators was tracked over 30 min and was found to be stable with changes of less than 2 kHz. Bovine Serum Albumin (BSA) solutions with different concentrations (62.5, 125, 250, 500 and 1000 µg/ml) were then placed on the top of different sets of identical SMRs both with metals and CNT electrodes, and allowed them to physically adsorb onto the SMRs' surface for 15 min, after which they were rinsed by buffer and dried with N_2. Measurements were subsequently carried out on loaded devices when they are completely dried to monitor the frequency shift (compared to unloaded devices), and results are shown in Fig. 8 as a function of BSA concentration. The frequency shift was of the order of hundreds of kHz, and has a distinct linear relation with the solution concentration over the range measured. The absolute frequency changes observed on the SMRs with CNT electrodes are significantly greater than

those with metal electrodes, demonstrating its high performance and great potential for use.

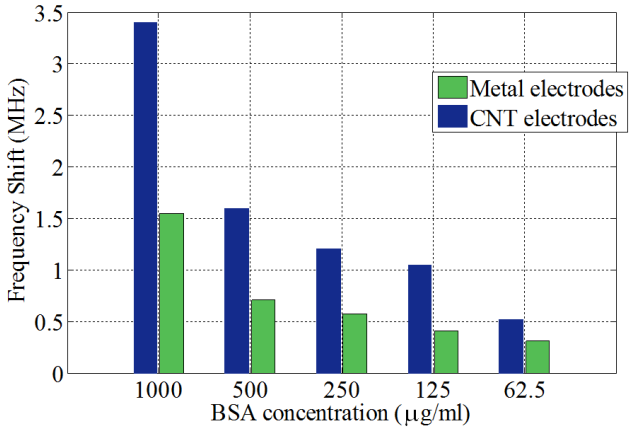


Figure 8. Frequency shift observed as a function of protein concentrations

V. CONCLUSIONS

BAW resonators with CNT electrodes have been shown to have many advantages for the development of acoustic wave based biosensors. AlN SMRs with a top CNT electrode resonating at around 1.75 GHz were microfabricated and exhibited a significantly large unloaded quality factor Q>2000, much greater than the identical devices fabricated with metal electrodes, which is a direct consequence of the low density and high acoustic impedance of the CNT electrodes. Devices with higher *Q* possess higher sensitivity due to sharper resonance peaks which facilitates tracking smaller frequency shifts. Furthermore, protein solutions with different concentrations were loaded on the top of the resonators and their responses to mass-load from physically adsorbed coatings were investigated. It was found that resonators using CNTs as the electrode material exhibited higher frequency change for a given load due to the higher surface area of a thin film of interconnecting CNTs compared to that of a metal thin film electrode.

VI. REFERENCES

[1] S.W. Wenzel and R.M. White, "*Analytic comparison of the sensitivities of bulk-wave, surface-wave and flexural plate-wave ultrasonic gravimetric sensors,*" Appl. Phys. Lett. Vol. 54 (20), pp. 1976-1978, 1989

[2] D.S. Ballantine, R.M. White *et al*., "*Acoustic wave sensors: theory, design and physico-chemical applications,*" R. Stern, M. Levy (Eds.), Applications of Modern Acoustics, Academic Press, San Diego, 1997

[3] M.D. Ward, D.A. Buttry *et al*., "*In situ interfacial mass detection with piezoelectric transducers*," Science, Vol. 249 1000–1007, 1990

[4] G. Sauerbrey, "*Erwendung von schwingqarzen zur wagung dunner schichten un zur mikrowagung*," Z. Physik, Vol. 155, pp. 206-222, 1959

[5] W.H. Jr. King, "*Piezoelectric sorption detector*," Anal. Chem. Vol. 36 (9), pp. 1735-1739, 1964

[6] J. Weber, W.M. Albers *et al*., "*Shear mode FBARs as highly sensitive liquid biosensors*," Sens. & Actuat. A: Phys., Vol. 128 (1), pp. 84-88, 2006

[7] G. Wingqvist, J. Bjurstrom *et al*., "*Immunosensor utilizing a shear mode thin film bulk acoustic sensor*," Sens. & Act. B: Chem., Vol. 123 (1), pp. 466-473, 2007

[8] G. Wingqvist, V. Yantchev *et al*., "*Shear mode AlN thin film electro-acoustic resonant sensor operation in viscous media*," Sens. & Act. A: Phys., vol. 148 (1), pp. 88-95, 2007

[9] P.B. Kirby, M.D.G. Potter *et al*., "*Thin film piezoelectric property considerations for surface acoustic wave and thin film bulk acoustic resonators*," J. Eur. Ceram. Soc., Vol. 23, pp. 2689-2692, 2003

[10] E. Iborra, M. Clement *et al*., "*Optimization of thin AlN sputtered films for X-band BAW resonators*," IEEE Ultrason. Symp. In press, 2010

[11] G. Ferblantier, F. Mailly *et al*., "*Deposition of zinc oxide thin films for application in bulk acoustic wave resonator*," Sens. & Act A, Vol. 122, pp. 184-188, 2005

[12] S-H. Choi and J-S Kim, "*Study on the c-axis preferred orientation of ZnO film on various metal electrodes*," Ultramiscroscopy, Vol. 108, pp. 1288-1291, 2008

[13] T. Yokoyama, T. Nishihara *et al.*, "*New electrode material for low-loss and high Q FBAR filters*," IEEE Ultrason. Symp., pp. 429-432, 2004

[14] S. Tanifuji, Y. Aota et al., "*Spurious vobration suppression by film thickness control FBAR*" IEEE Ultrason. Symp., pp. 2193-2196, 2008

[15] M. Dragoman, A. Muller et al., "*High performance thin film bulk acoustic resonator covered with carbon nanotubes*," App. Phys. Lett., Vol. 89, 143122, 2006

[16] L. Garcia-Gancedo, F. Al-Naimi *et al*., "*Fabrication of High-Q Film Bulk Acoustic Resonator (FBAR) Filters with Carbon Nanotube (CNT) electrodes*", IEEE Ultrason. Symp. 2010

[17] M. Mann, Y. Zhang *et al*., "*Controlling the growth of carbon nanotubes for electronic devices*," Microelectronic Engineering, Vol. 87 (5-8), pp. 1491-1493, 2010

[18] S. Liu, Q. Shen *et al.*, "*Chemical functionalization of single-walled carbon nanotube field-effect transistors as switches and sensors*," Coord. Chem. Revs., Vol. 254, pp. 1101-1116, 2010

[19] A. Star, J.P Gabriel *et al*. "*Electronic detection of specific protein binding using nanotube FET devices*," Nanoletters, Vol. 3, pp. 459-463, 2003.

[20] P. Qi, O. Vermesh *et al*., "*Toward large arrays of multiplex functionalized carbon nanotube sensors for highly sensitive and selective molecular detection*," Nanoletters, Vol. 3, pp. 347-351, 2003.

[21] J.F. Lu, L. Gu *et al*., "*Advances in bioapplications of carbon nanotubes*," Advanced Mat., Vol. 21, pp. 139-151, 2009.

[22] A. Modi, N. Koratkar *et al*., "*Miniaturized gas ionization sensors using carbon nanotubes*," Nature, Vol. 424, pp. 171-174, 2003.

[23] J. Li, A.M. Cassell *et al*., "*Carbon nanotubes as AFM tips: Measuring DNA molecules at the liquid/solid interface*," Surf. & Interf. Anal., Vol. 28, pp. 8-11, 1999

[24] A. Devos, E. Iborra *et al*., "*Picosecond ultrasonics as a helpful technique for introducing a new electrode material in BAW technology: the iridium case*," IEEE Ultrason. Symp., pp. 1433-1436, 2007.

[25] J. Olivares, E. Wegmann *et al*., "*Sputtered SiO_2 as low acoustic impedance material for bragg mirror fabrication in BAW resonators*", IEEE Trans. Ultrason. Ferr. Freq. Control., in press.

[26] K. J. Coakley, J. D. Splett *et al*., "*Estimation of Q factors and resonant frequencies*," IEEE Trans. Micro. Theo. & Tech., Vol. 51, pp. 862 – 868, 2003

A History of the Development of CMOS Oscillators: The Dark Horse in Frequency Control

M. S. McCorquodale and V. Gupta

Silicon Frequency Control, Integrated Device Technology, Inc.,
Sunnyvale, CA 94085 USA
{michael.mccorquodale, bhusan.gupta}@idt.com

***Abstract*—Microelectronic technologies developed over the past decade to replace quartz crystal resonators (XTALs) and oscillators (XOs) are discussed. A new figure-of-merit (FOM) is proposed to compare these emerging technologies. It is shown that CMOS oscillators (COs) have been an underestimated technology amidst these efforts as the performance of COs is comparable to existing and emerging technologies and continues to scale. Motivations and concepts behind COs are presented along with native and package-induced frequency drift mechanisms. Design and packaging approaches that minimize drift are demonstrated. A brief history of CO product embodiments is presented. State-of-art CO performance is reported where better than ±25ppm frequency drift is achieved over 0 to 70°C and less than 950fs$_{RMS}$ phase jitter integrated from 12kHz to 20MHz is realized.**

I. Introduction

Quartz crystal resonators (XTALs) and oscillators (XOs) serve as the de facto frequency reference in nearly all electronic platforms. Yet quartz cannot be integrated into microelectronic form. Achieving that goal would likely enable frequency control devices with a smaller form-factor, lower cost, greater flexibility, higher reliability and a myriad of unforeseen potential benefits. Consequently, efforts to replace quartz with such a technology date back to the 1960's. In 1967, a transistor was presented with a mechanically resonant gate [1]. Similarly, the frequency stability of solid-state oscillators, such as that in [2], was also explored. Those efforts were limited as the net performance of frequency control devices based on these technologies was inferior to that which could be achieved with quartz.

In 1980, silicon was presented as a viable mechanical material [3]. Since then, and particularly in the 1990's to the present, several researchers have aimed to develop technologies leveraging silicon microelectromechanical systems (MEMS) to replace quartz. In [4], one of the first MEMS oscillators (MOs) was presented. These results spawned a variety of academic and commercial efforts in the field. However, as these efforts captured the limelight, radio frequency (RF) CMOS became mainstream [5]. With this development, the possibility of high-performance, self-referenced, *LC* CMOS oscillators (COs) became a reality. As opposed to XOs and MOs, COs are electrically resonant and entirely solid-state. In this manuscript, a brief history of the development of this technology will be presented up to the state-of-the-art. It will be shown that COs now achieve performance that rivals both MEMS and quartz technology on several dimensions.

II. Performance Benchmarking and a Proposed FOM

MEMS resonators can be fabricated in a batch silicon process and typically exhibit a high quality (Q) factor, which is what makes the technology interesting for frequency control [4]. In [6] and [7], efforts to benchmark the performance of XO, MOs and COs were presented. Several performance metrics were reported including frequency stability, timing and phase jitter, single sideband (SSB) phase noise power spectral density (PSD) and power dissipation. In that work, it became clear that the native high-Q of MEMS resonators was not realized in the complete device. This is due, predominantly, to the fractional-N phase-locked loop required for frequency-trimming and temperature-compensation of the resonator. Additionally, other work has shown that non-linear transduction of the resonator results in upconversion of flicker noise in MOs [8]. Nevertheless, the native high-Q of MEMS resonators continues to be cited as a critical performance metric despite the fact that frequency control devices referenced to MEMS resonators do not exhibit performance consistent with high-Q.

Addressing these efforts to benchmark different technologies while avoiding use of less relevant metrics, the authors propose a figure-of-merit (FOM) to capture the multi-dimensional nature and aggregate performance of frequency control devices. The radar plot in Fig. 1, adapted from that in [9], illustrates the concept that the best-performing frequency con-

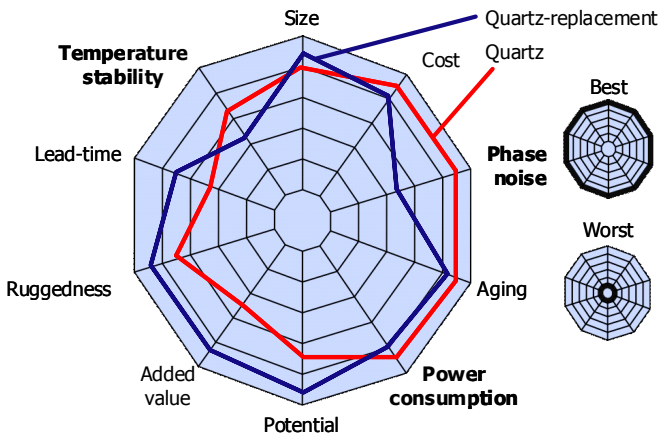


Figure 1. Radar plot illustrating the performance dimensions of a complete frequency control device. Dimensions in bold are considered significant and technology-dependent. The figure is adapted from [9].

trol device achieves high-performance across all dimensions simultaneously. As shown, XOs typically achieve high-stability, low-noise and low-power. Emerging technologies for quartz replacement are generally inferior along these dimensions, but performance continues to scale and these technologies are positioned to offer benefits over quartz including smaller size and shorter lead-times.

In this work, the authors consider the three most significant and technology-dependent performance metrics: frequency stability, phase jitter (integrated from 12kHz to 20MHz) and power dissipation. The goal is to define a measure that is frequency and technology-independent while capturing trade-offs in the embodiment of the device. To begin with, frequency stability is frequency-independent by definition. Similarly, phase jitter over the bandwidth presented is approximately frequency-independent. In emerging technologies, high frequency stability is often achieved through active temperature-compensation due to the fact that the native temperature coefficient of frequency (TC_f) of the reference resonator is high, particularly compared to quartz. Consequently, power is often a penalty in achieving high stability.

Considering these factors, the proposed FOM is computed by plotting the value of each metric along mutually-orthogonal axes and solving for the volume (V_{FOM}) of the irregular tetrahedron defined by these points. The volume is then normalized to 1ppm-ps_{RMS}-mW and taken in dB. Thus, lower FOM indicates better performance. Such an approach equally-weights all three metrics. Fig. 2(a) illustrates ideal performance, which is a single point at 0ppm-ps_{RMS}-mW. COs have traditionally demonstrated performance similar to that in Fig. 2(b) where low-power is achieved, but with poor stability and phase jitter. Fig. 2(c) illustrates the opposite case where high-stability and low-noise are achieved through increased power.

The proposed FOM is particularly useful in benchmarking various technologies. Table I presents the performance of a TCXO, XO and several generations of MOs and COs. Consistent with Fig. 1, the FOM shows that the TCXO and XO achieve performance superior to the emerging technologies. Further, it is interesting to note that the performance of several MOs and COs are comparable. One of the MOs achieves stability and phase jitter that is comparable to the TCXO, but at the expense of power, which is captured by the FOM.

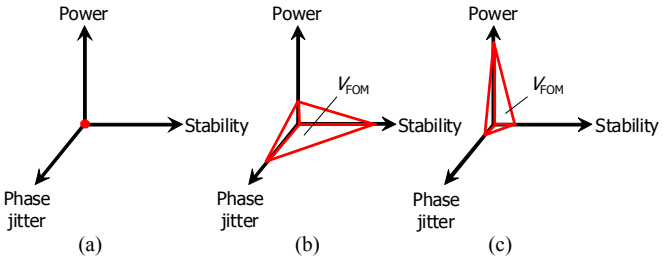


Figure 2. Visualization of the proposed FOM for frequency control devices. (a) Ideal. (b) A device exhibiting low-power, but with poor stability and phase jitter. (c) A device exhibiting high frequency stability and low phase jitter at the expense of power.

III. CMOS Oscillator Technology

A. Motivations and Concepts

A high-accuracy, low-noise and low-power CO would enable frequency control devices to be migrated to a standard and batch silicon manufacturing process with nearly infinite capacity. This would enable the lowest possible cost structure while introducing the possibility of complete microelectronic integration of the frequency source. However, CMOS technology does not include high-Q components. For example, and despite the advent of RF-CMOS, the Q-factor of an integrated inductor is in the range of 10-20. However, CMOS technology supports the design of high-frequency oscillators. Thus, the Q-frequency (Q-f) product in COs is high. Consequently, COs can achieve low-noise, despite the low-Q of the resonator compared to MOs and XOs [10]. The remaining challenge involves minimizing frequency drift and achieving high frequency stability.

B. Native Frequency Drift Mechanisms

Native frequency drift mechanisms in COs have been presented and discussed in [10]. The natural resonant frequency of an LCO is $\omega_o = \sqrt{1/(LC)}$ where L is the net tank inductance and C is the net tank capacitance. Due to resistive losses in both the inductor and capacitor, the actual resonant is given by,

$$\omega_1 = \omega_o \sqrt{\frac{CR_L^2 - L}{CR_C^2 - L}}, \quad (1)$$

Table I. Benchmarking the performance of various frequency control devices based on different technologies

Metric	TCXO (TCF4)	XO (SG-210)	MEMS (SiT8002)	MEMS (SiT5002)	MEMS (DSC1121)	CMOS (Si500)	CMOS (IDT3CP02)	CMOS (IDT3LV04)
Stability (ppm)	2	50	50	2.5	10	150	100	50
Phase jitter (ps_{RMS})	0.25	0.25	250	0.6	1.7	2	2	1
Power (mW)	3.6	3	73	58	50	25.2	3.6	10.8
FOM (dB)	**2.6**	**15.6**	**59.6**	**19.4**	**29.3**	**38.8**	**28.6**	**27.3**

where R_L and R_C are the losses in the coil and the capacitor respectively. Typically, R_L is significantly larger than R_C. Consequently, (1) can be simplified to,

$$\omega_1(T) \approx \omega_o \sqrt{1 - CR_L^2(T)/L}, \quad (2)$$

where it is clear that the native TC_f is negative and concave-down.

In [10], additional drift mechanisms were presented and include harmonic work imbalance and the TC_f of both the fixed capacitance and the variable capacitance presented by the transistors in the sustaining amplifier. A closed-form solution for drift arising from HWI was presented in [10] and is dependent on the spectral content of the current delivered to the tank by the sustaining amplifier. These additional drift mechanisms are small compared to the drift induce by loss in the tank. However, these mechanisms can introduce non-linearity in the temperature-compensated response of the CO.

C. Design Approaches

In [10], an orthogonal frequency-trimming and temperature-compensation approach was presented. Referring to Fig. 3, an array of switched thin-film capacitors, C_f[X:0], serves to trim the nominal frequency against process variation. Temperature-compensation is achieved with an active approach. Here, either the power-supply or a temperature-dependent voltage, $v_{ctrl}(T)$, is applied to a switched array of varactors, C_v[Y:0]. When the varactor is biased to the power supply, it does not vary whereas when switched $v_{ctrl}(T)$, the TC_f is compensated. The number of active varactors sets the amount of compensation.

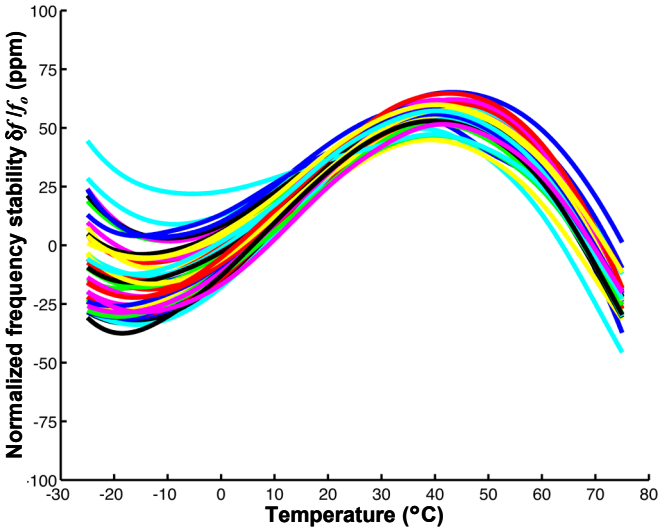


Figure 4. Measured absolute frequency stability of 40 randomly selected COs from production test. No device exceeds ±75ppm error from -20 to 70°C.

Referring to (1), it is clear that if a loss, R_C, is deliberately introduced to the tank capacitance, C, the drift due to the loss in the coil, R_L, can be cancelled. This observation led to the development of the passive compensation approach in [11], which is illustrated in Fig. 3. This approach has enabled COs to achieve total frequency stability under ±100ppm over -20 to 70°C, all operating conditions and lifetime while dissipating less than 4mW. Fig. 4 presents the TC_f of 40 devices which were selected randomly from the production test flow. No device exceeds ±75ppm frequency error against temperature.

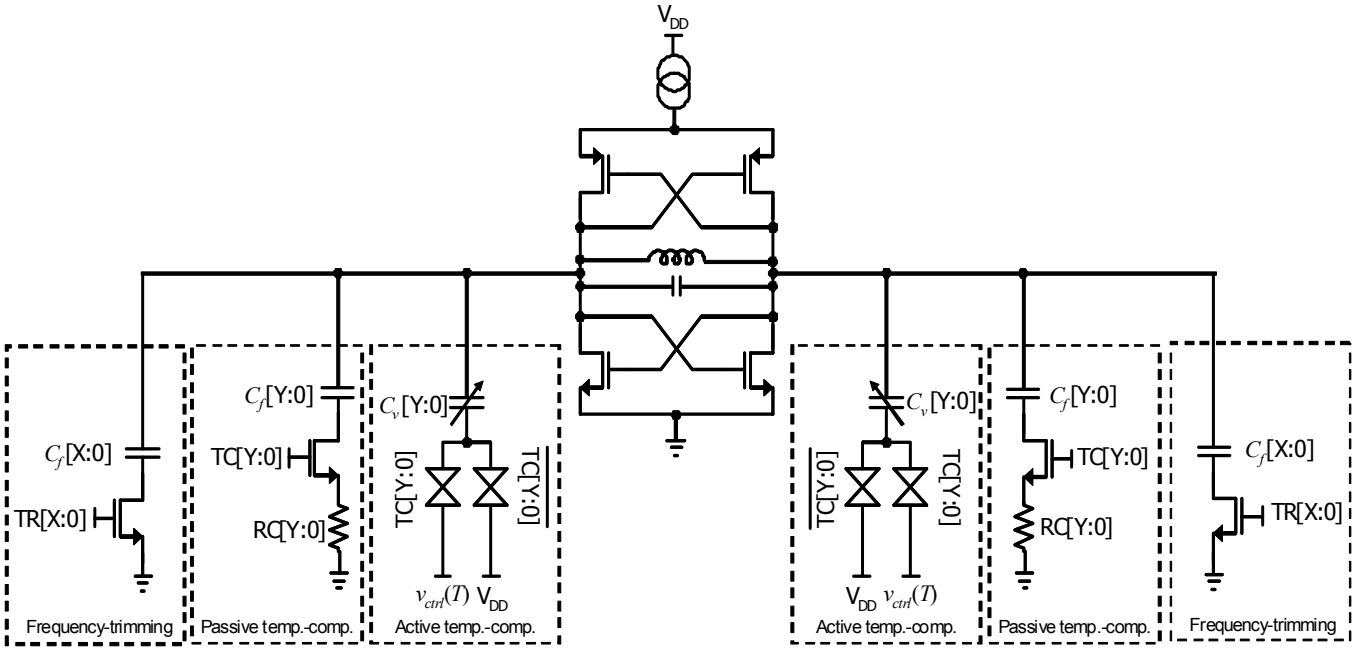


Figure 3. Simplified schematic illustrating frequency-trimming and active and passive approaches to temperature-compensation of the CO.

D. Package-Induced Frequency Drift

The work in [10] and [11], among other work, presented different design approaches to contain the native frequency drift of the CO. However, in [11], it was shown that fringing electromagnetic fields emanating from the die can be perturbed by the surrounding package or environment and induce frequency drift. Referring to Fig. 5(a), the *B*-field that radiates from the coil can extend beyond the boundary of the package. Consequently, if the field is modulated by a permeable material or terminated with a metal, the frequency will drift. Similarly, a parasitic capacitance is created by the stray *E*-field from the device. Any changes in the permittivity of the molding compound of the package can induce frequency drift. Left uncontained, each mechanism can induce frequency drift exceeding hundreds of ppm.

A post-processed wafer-scale Faraday shield was developed in [11] to contain these package-induced frequency drift mechanisms. Fig. 5(b) illustrates the latest rendition of this shield. A thick dielectric mesa is photo-patterned on top of the die, excluding the bond pad windows. Next, metal is electroplated over the mesa. The metal is patterned such that it encapsulates the mesa and contacts the over-glass layer of the die which creates a hermetic seal. Additionally, the backside of the die is metallized. Metal thicknesses are selected to contain the known strength of the fields. As illustrated in Fig. 5(b), the shield terminates the fringing *B*-field on both sides of the die. Additionally, the hermetically-sealed dielectric material presents a constant permittivity to the fringing *E*-field. This approach enabled the devices in [11] to achieve ±300ppm stability over all operating conditions including a panel of industry-standard stress and reliability tests. This latest generation of the Faraday shield, pictured in Fig. 6, enables COs to achieve ±100ppm frequency stability over the same tests.

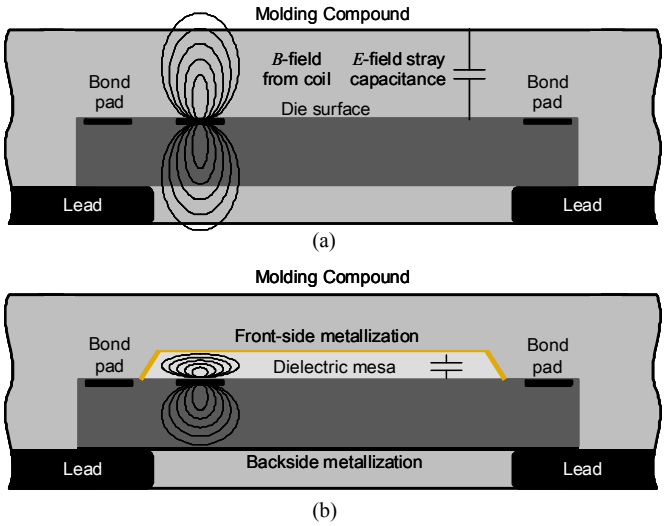


Figure 5. (a) Illustrated cross-section of the CO assembled in a standard plastic package. Electromagnetic fields fringe from both sides of the die. (b) Illustrated cross-section of the post-processed wafer-scale Faraday shield for containing and terminating the fringing fields.

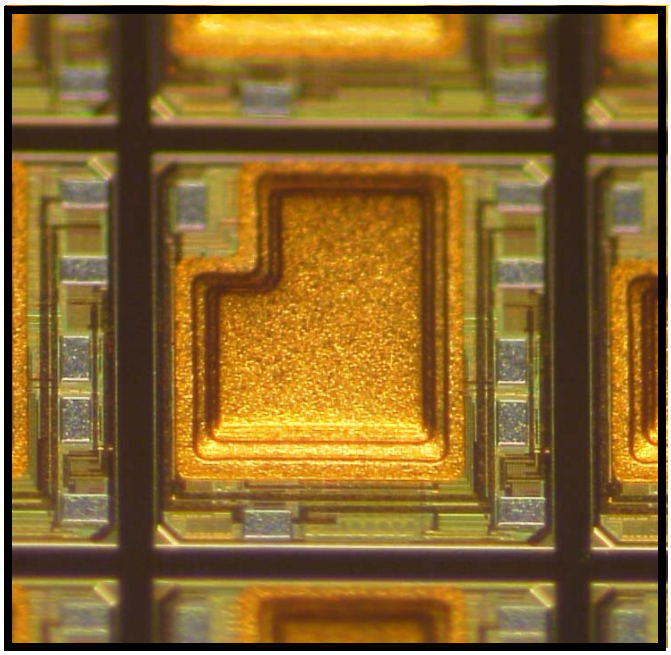


Figure 6. Micrograph of the post-processed Faraday shield at the wafer-level.

IV. Commercialization of CMOS Oscillators

A. Hard Intellectual Property Macro

In [10], a hard intellectual property (IP) macro of a CO was presented. The macro was embedded in a USB bridge controller and was fabricated in a 0.35µm 2P4M CMOS process. The CO occupied 0.22mm^2, dissipated 31mW and maintained ±400ppm stability from -10 to 85°C. The device was a commercial success where tens of millions of units have shipped in production. However, and as discussed in [11], IP is difficult and expensive to support due to costly re-targeting of the macro to various foundries and technology nodes.

B. A Silicon Die as a Frequency Source

In [11], a silicon die as a frequency source was presented for XTAL resonator replacement and in an effort to address the challenges associated with IP while also demonstrating significant differentiation of COs when compared to MOs and XOs. As described in [11], this device utilized the passive temperature-compensation approach shown in Fig. 3. Further, the ability to deliver the device in unpackaged form was enabled by the development of the Faraday shield presented previously. The device achieved ±300ppm total frequency stability over all operating conditions including industry-standard stress and reliability tests. Further, it dissipated under 4mW while achieving 2ps$_{RMS}$ phase jitter integrated from 12kHz to 20MHz.

This device has achieved the greatest commercial success to date as it is well-suited to wireline applications such as USB, S-ATA and PCIe where stability requirements do not exceed ±300ppm. The die frequency source can be assembled

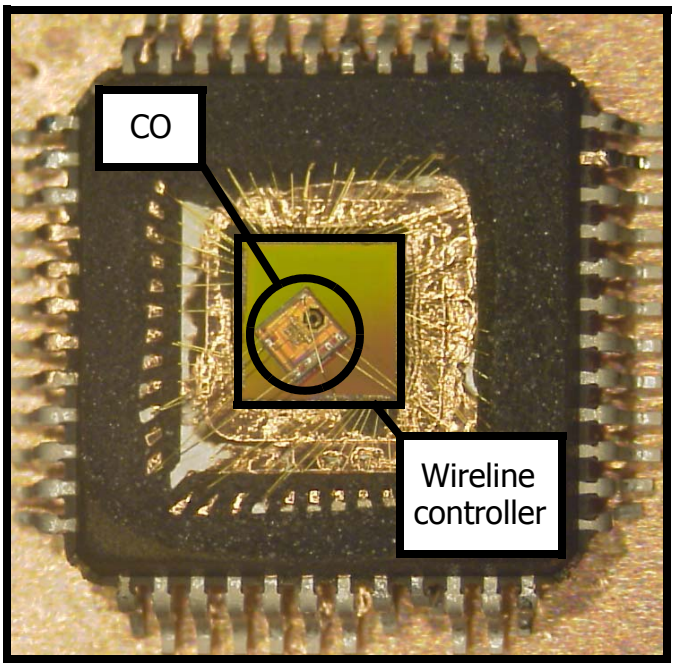


Figure 7. The CO presented in [11] assembled in a 48-pin QFP MCP atop a wireline controller. The MCP eliminates the need for an external XTAL. To the end-user, the device appears to include a hard IP macro. The Faraday shield has been dissolved with the package. The CO die was fabricated in a 1P6M 0.13µm CMOS technology and measures 920µm × 880µm × 250µm.

with a chip-on-board (CoB) process, as shown in [11], or it can be stacked in a multi-chip package (MCP) as shown in Fig. 7. In the latter case, the final assembled device appears as if it contains the hard IP macro in [10]. However, the die frequency source is fabricated in a fixed technology node, thus overcoming the challenges associated with IP development while achieving the same goals of IP integration.

C. Standard Components

Though COs can be integrated as IP and assembled in an MCP or with a CoB process, that does not preclude the device from being assembled as a standard component. COs have been assembled in standard plastic packages that are pin-compatible with XOs and MOs. The frequency stability data shown in Fig. 4 were captured from these standard devices. Historically, these devices have had modest commercial success due to the fact that frequency stability was inferior to that of quartz. However, as Fig. 4 shows, the stability of these devices now rivals XOs. Consequently, commercial traction has increased substantially and is growing exponentially.

D. Reliability

Initial reliability data for COs was presented in [11]. Since then, several enhancements have improved the performance of COs markedly. For example, Fig. 8 shows the measured frequency stability of a population of over 200 COs. It is shown

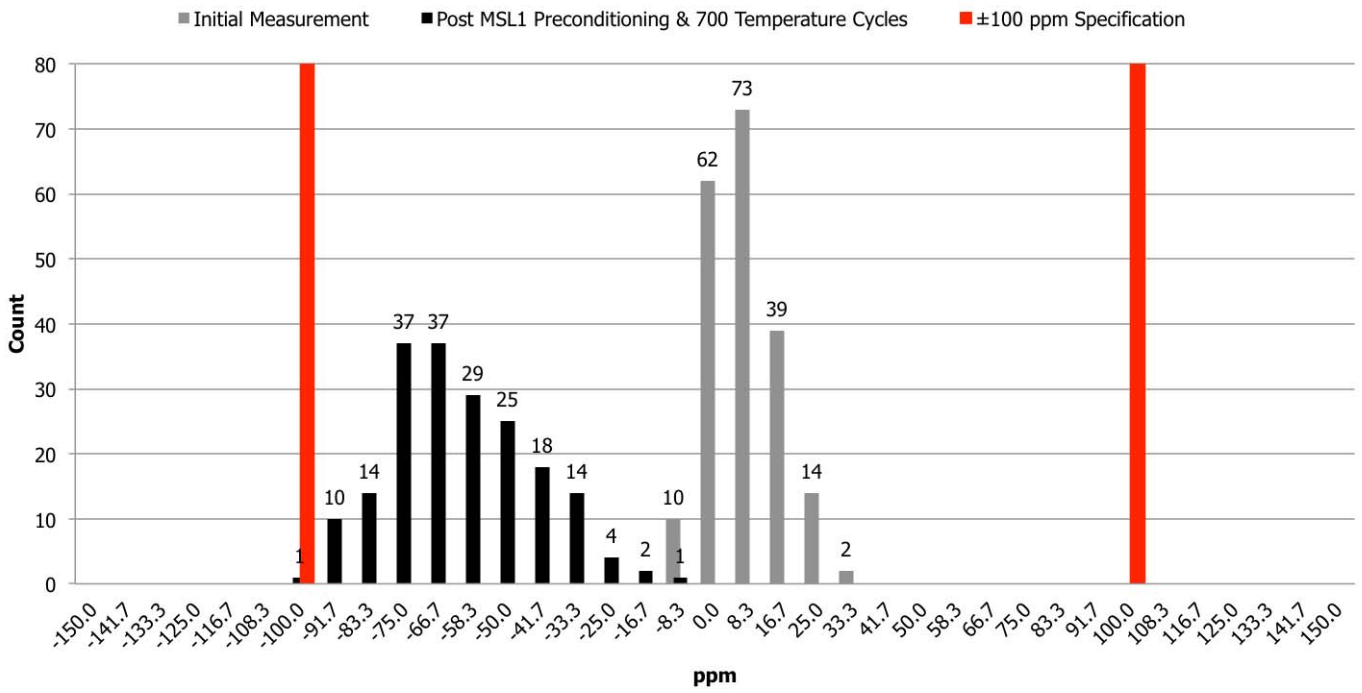


Figure 8. Measured frequency accuracy of COs assembled in standard plastic packages before and after JEDEC MSL1 preconditioning and 700 temperature cycles from -55 to 125°C. All devices maintain ±100ppm frequency stability after stress. The corresponding lifetime is 29 years. Other industry-standard stress tests yield similar results.

that ±100ppm frequency stability is maintained for the entire population after MSL1 pre-conditioning and 700 temperature cycles from -55 to 125°C, which corresponds to a lifetime of 29 years. Other stress and reliability tests yield similar results.

E. State-of-the-Art Performance

The latest generation of COs leverages a frequency-trimming and temperature-compensation approach that is a hybrid of the approaches presented in Fig. 3. Further, a next-generation Faraday shield has also been developed, though it is not shown here. With these advances, state-of-the art performance has improved significantly. In Fig. 9, the temperature-stability of 40 devices selected at random is less than ±25ppm from 0 to 70°C. Further, referring to Fig. 10, the phase jitter integrated from 12kHz to 20MHz is less than $1ps_{RMS}$, which is a significant achievement because this meets the requirement for optical networking.

V. Future Directions and Conclusion

Significant efforts continue in developing frequency control technologies to replace quartz. Given these various technologies, a new FOM was proposed to facilitate benchmarking. As shown, the performance of COs continues to improve to the point that the technology now rivals the performance of XOs and MOs. The temperature-stability of state-of-the art COs is under ±25ppm while phase jitter is less than $1ps_{RMS}$ integrated from 12kHz to 20MHz. Additionally, CO technology has enabled several new embodiments of frequency control devices including IP, singulated die and standard components in plastic packages. Many of these devices have achieved significant commercial success. It is expected that further enhancements to CMOS process technology, new design approaches and new post-process technology will enable CO performance to continue to scale into the future.

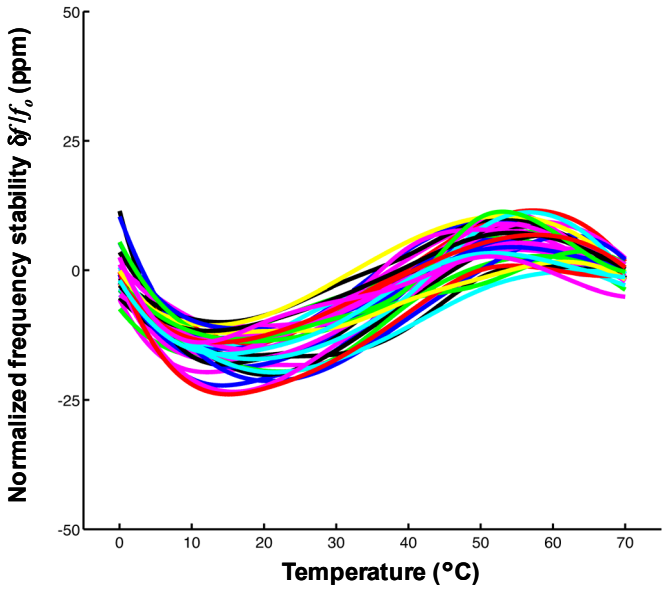


Figure 9. Measured absolute frequency stability of 40 randomly selected next-generation COs against temperature. No device exceeds ±25ppm frequency error from 0 to 70°C.

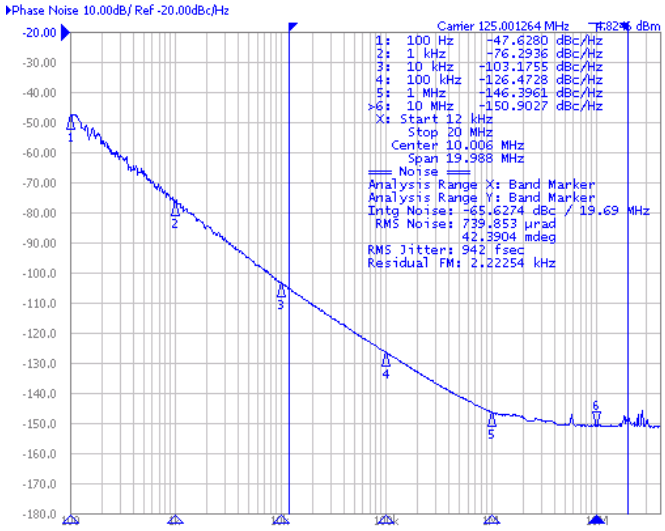


Figure 10. Measured SSB phase noise PSD of the a next-generation CO. Phase jitter, integrated from 12kHz to 20MHz, is $942fs_{RMS}$. The device maintains $<1ps_{RMS}$ phase jitter over the same bandwidth over all process corners.

VI. Acknowledgement

The authors acknowledge all members of Silicon Frequency Control at Integrated Device Technology for their dedication and persistence in developing the technology and products presented herein.

References

[1] H. C. Nathanson, *et al*., "The Resonant Gate Transistor," *IEEE Trans. on Electron Devices*, vol. ED-14, no. 3, pp. 117–133, Mar. 1967.

[2] R. F. Adams and D. O. Pederson, "Temperature sensitivity of frequency of integrated oscillators," *IEEE J. Solid-State Circuits*, vol. SC-3, no. 4, pp. 391–396, Dec. 1968.

[3] K. E. Petersen, "Silicon as Mechanical Material," *Proc. of the IEEE*, vol. 70, no. 5, pp. 420–457, May 1982.

[4] C. T.-C. Nguyen and R. T. Howe, "An integrated CMOS micromechanical resonator high-Q oscillator," *IEEE J. Solid-State Circuits*, vol. 34, no. 4, pp. 440–455, April 1999.

[5] A. A. Abidi, "RF CMOS comes of age," *IEEE J. Solid-State Circuits*, vol. 39, no. 4, pp. 549–561, April 2004.

[6] D. Kenny and R. Henry, "Comparative Analysis of MEMS, Programmable and Synthesized Frequency Control Devices versus Traditional Quartz Based Devices," *in Proc. of IEEE Int. Freq. Control Symp.*, May 2008, pp. 396–401.

[7] M. S. McCorquodale, "Self-Referenced, Trimmed and Compensated RF CMOS Harmonic Oscillators as Monolithic Frequency Generators," *in Proc. of IEEE Int. Freq. Control Symp.*, May 2008 pp. 408–413.

[8] S. Lee, *et al*., "A 10-MHz micromechanical resonator Pierce reference oscillator for communications," *in Dig. of Technical Papers, Solid-State Sensors & Actuators*, June 2001, pp. 1094–1097.

[9] Schoepf, *et al*., "TCMO: A versatile MEMS oscillator timing platform," *in Proc. of the 41st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting*, Nov. 2009.

[10] M. S. McCorquodale, *et al.*, "A Monolithic and Self-Referenced RF LC Clock Generator Compliant with USB 2.0," *IEEE J. of Solid State Circuits*, vol. 42, no. 2, Feb. 2007, pp. 385–399.

[11] M. S. McCorquodale et al., "A silicon die as a frequency source," *in Proc. of IEEE Int. Freq. Control Symp.*, June 2010, pp. 103–108.

A Highly Stable CMOS Self-Compensated Oscillator (SCO) Based on an LC Tank Temperature Null Concept

A. Ahmed, B. Hanafi, S. Hosny, N. Sinoussi, A. Hamed, M. Samir, M. Essam, A. El-Kholy, M. Weheiba, A. Helmy
Timing Products Division, Si-Ware Systems, Cairo, Egypt

***Abstract*—A highly stable all silicon Self-Compensated Oscillator (SCO) reference source is presented. Self compensation across temperature is achieved using a newly discovered phenomenon in LC-tanks named the Temperature Null (TNULL) which has been described and briefly analyzed. A quadrature oscillator architecture has been used to implement the SCO in a 0.18μm CMOS technology. The SCO has on-chip infrastructure facilitating low cost trimming and frequency calibration at room temperature. Excellent frequency stability of ±50ppm has been measured across temperature (0 – 70°C), 3.0 – 3.6V supply and 1 – 15pF load. The SCO has a programmable frequency range of 1 – 133MHz, consumes 7.1mA at 25MHz and has an RMS period jitter of 2ps at 125MHz.**

I. INTRODUCTION

The demand for clocks has been continuously fueled by increased consumption of many feature rich consumer gadgets. A clock is generated from a reference frequency source which has to date been dominantly based on a quartz crystal. This industry de-facto of quartz crystals has been rightfully earned through reliable performance of quartz crystal oscillators (XO) in terms of excellent frequency stability across temperature, low power consumption and low noise. Moreover, quartz crystals enjoy very high manufacturing maturity that has been gained over the past decades which has helped sustain a steady, yet slow, decrease in the price of quartz crystal-based products. As new applications emerge, a need for new reference frequencies arise which unfortunately require long lead times of up to 14 weeks to be developed in quartz crystals leading to longer development cycles and slower market deployment. There has been always a need, thus great value, in decreasing the size of any solution through higher levels of integration while adding more features and functionality. The initial trend of decreasing package footprint and thickness has gained acceptance and interest. However, the benefits of having a monolithic CMOS timing solution are unbeatable. Such a solution will benefit from the continuously decreasing price of CMOS technologies that can produce thousands of reference clocks from a single wafer in comparison to hundreds of quartz crystals from a panel. This is directly translated into lower cost and very high volume capability.

The dominant deterring factor in clock selection for very high volume consumer products will be price as long as the required performance is secured. The most important performance metric is frequency stability across temperature, supply, load and aging. A total frequency stability of ±100ppm across a temperature range of 0 – 70°C and -20 – 70°C is considered the critical threshold that gives access to a large percentage of the market with heavy emphasis on very high volume consumer applications. Aggressive price reduction can only be achieved through a quantum reduction in the cost structure of clocks which requires the search for alternative technologies to quartz. The cost of an XO is composed of: Quartz Crystal, CMOS die, Ceramic Package, Assembly and Test. A programmable reference clock based on a silicon micro-electro mechanical system (MEMS) resonator, [1], [2] has merits in lowering the lead time but fails to address higher integration levels. The cost of a MEMS resonator is lower than a quartz one owing to very high number of resonators on a single MEMS wafer yet still requires wafer level packaging to conserve its performance and reliability. Expensive ceramic packaging is also dropped to low cost plastic packaging [3], [4]. However, the assembly cost of a MEMS-based solution requires stacking of the wafer level packaged MEMS die and CMOS die. In addition, the CMOS die in a MEMS-based solution is larger, compared to the die in an XO, thus more expensive. Furthermore elaborate production testing is required to trim each part to the required performance. Overall, the cost benefits of a MEMS-based solution is not evident thus cannot easily compete in the price sensitive lower performance consumer market.

Another alternative technology is an all silicon CMOS reference clock using an on-chip LC-tank. Recently several solutions have been reported [5], [6], [7] and commercialized. Such solutions are by definition highly integrated and have short lead times due to their programmability. Furthermore, the cost structure is very promising since it has been reduced to: CMOS die, low cost plastic package, simple assembly and production testing. The die in this case is still larger than the one found in an XO yet smaller than the die in a MEMS-based solution. It is then obvious that such a solution is conceptually of lowest cost as long as the production test cost is low. This work presents a new all silicon CMOS reference clock that uses an on-chip LC-tank designed to operate at a very specific low temperature-sensitivity phase operating point thus self compensating across temperature. The part is trimmed using a low cost single insertion at room temperature test procedure. Such a solution is promising very low cost and has achieved a total frequency stability of ±50ppm that can address a large

This work was partially sponsored by ITIDA through the ITAC program, PDP project grant.

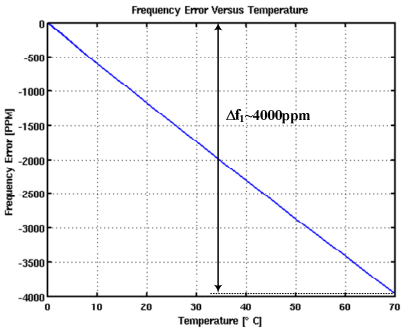


Figure 1. Temperature dependence of an LC-tank resonance frequency.

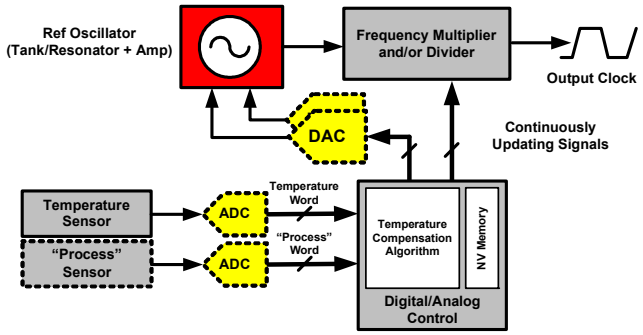


Figure 2. Generic block diagram of open loop temperature compensation.

percentage of high volume consumer applications such as SATA, PCI-E, HDD, LCD TV, printers and many others.

II. Temperature Compensation Techniques

The frequency stability of classical LC-tank based oscillators is highly dependent on temperature and typically may have a temperature coefficient ranging from -50ppm/^{o}C to 100ppm/^{o}C. The main source of temperature dependence comes from the LC-tank and in particular the inductive part. Fig. 1 shows simulation results of the frequency variation of an LC-tank across a 0 – 70^{o}C temperature range of ~4000ppm which is not acceptable as a reference source. This large temperature dependence of LC-tanks has been the main challenge in designing an LC-tank based reference which requires sophisticated compensation techniques to neutralize the frequency variation. There are two main types of compensation techniques; either an open loop technique or a closed loop technique. Closed loop techniques are commonly used in oven-compensated oscillators where the ultimate objective is to tightly control the oscillator temperature by placing it in a temperature controlled oven and in turn stabilizing its frequency. Alternatively, one may opt to design the frequency determining element of the oscillator, the resonator mainly, to be temperature insensitive thus the oscillator becomes self-compensated.

A. Open Loop Compensation

An open loop temperature compensation technique aims to negate the temperature generated frequency shift of the oscillator by imposing a frequency shift in the opposite direction of equal magnitude. To achieve this one can modify the oscillation frequency using one or more frequency tuning controls of the oscillator. Alternatively one can follow the oscillator with a frequency multiplier and control the multiplication factor such that the output frequency is temperature independent. In all cases one must measure the temperature accurately using a suitable sensor and use it to generate the temperature dependent control(s) of the oscillator or frequency multiplier. To successfully achieve the required accuracy in compensation it is imperative not just to have an accurate temperature measurement but to also have precise knowledge of the oscillator frequency across temperature and its frequency tuning control(s). Fig. 2 shows a generic block diagram of an open loop compensation system indicating different variations of where the temperature sensor analog output may be directly used or converted to a digital word. Similarly the temperature compensation algorithm may be implemented using analog or digital circuitry and can generate analog or digital controls which may be converted back to analog signals to control the oscillator.

Clearly this is a knowledge based process since exact cancellation of the temperature dependence will require deterministic knowledge of the complete system behavior to build an effective compensation algorithm. Otherwise one must rely on characterization of the system performance across temperature for every oscillator unit to generate the required data. However, this may still not provide perfect cancellation of frequency shifts unless the system has the ability of synthesizing controls with high accuracy. Inevitably this adds to the overall system complexity that may be translated into cost due to a large die size and/or long production test times using two or more temperature insertions. Nonetheless, open loop compensation has been used to compensate MEMS-based clocks [3] which uses a fractional-N Σ-Δ PLL frequency multiplier that is digitally temperature controlled. On the other hand, a CMOS LC-tank based reference oscillator reported in [6] uses a digital open loop compensation technique that utilizes an integrated microprocessor and [5] uses analog temperature dependent circuitry to control the behavior of the oscillator across temperature. All open loop techniques will require the storage of calibration coefficients on-chip during the calibration process for future retrieval during the lifetime of the oscillator.

B. Self Compensated Oscillator

A self-compensated oscillator across temperature has intrinsically low temperature sensitivity. Such oscillators do not require any real time compensation and temperature sensors and are thus simpler. Usually self compensated oscillators have a specific temperature range where the oscillation frequency has tolerable variation. An excellent example of a self compensated oscillator is an XO built using a quartz crystal. An AT-cut quartz crystal if designed and manufactured correctly can exhibit low temperature sensitivity. This is achieved by selecting the exact cutting angle that achieves lowest temperature sensitivity. This was made possible by excellent control on the manufacturing process of quartz crystals which is today a very mature and stable technology. Similarly there have been trials for producing MEMS resonators that are self compensated by designing stiffness compensated and geometric compensated resonators as noted by [1] and summarized in Fig. 3. Such

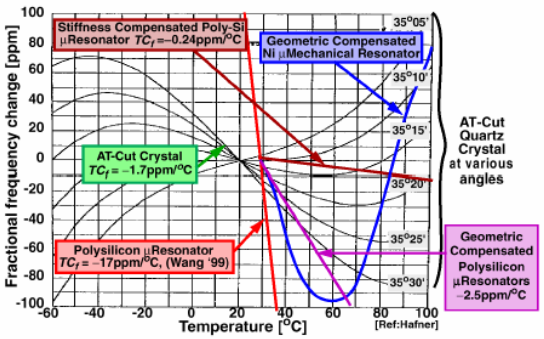
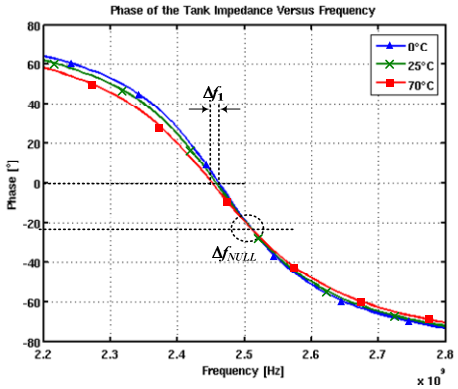


Figure 3. Frequency stability over temperature of self-compensated AT-cut quartz crystal and MEMS resonators.

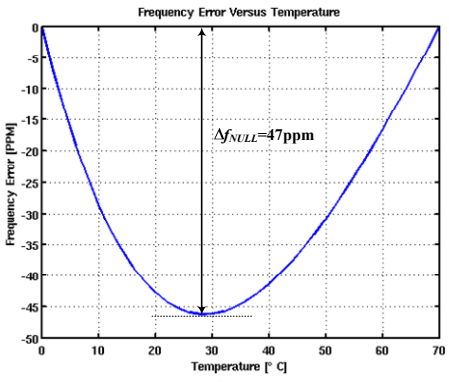
designs depend on materials with opposite temperature coefficients to cancel. However, the manufacturing accuracy and purity of the materials may impact the overall temperature dependence resulting in a wide variation in performance that is not industrially accepted. Such resonators end up with very non-linear temperature dependence of wide variation that make the use of open loop compensation techniques to correct them very challenging. In effect commercially viable self compensated MEMS resonators up to date are not available.

III. LC-Tank Temperature Null Concept

Building an oscillator operating at the LC-tank resonance frequency with variation over temperature shown in Fig. 1

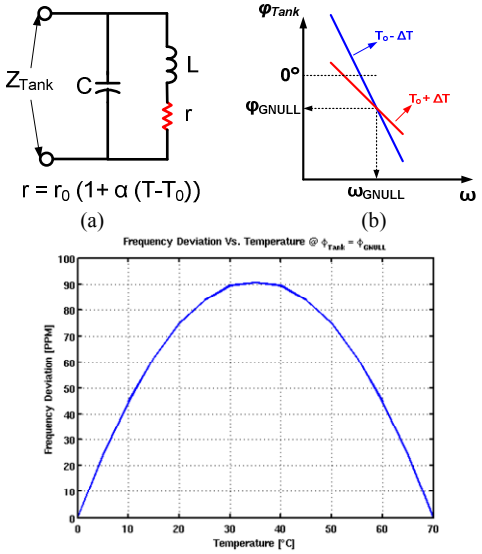


(a)



(b)

Figure 4. (a) LC-Tank impedance phase versus frequency for different temperatures and (b) frequency variation at TNULL.



(a) (b) (c)

Figure 5. (a) Simple first order model of an LC-tank, (b) Global Null of an LC-tank and (c) Frequency deviation at GNULL phase.

would result in very poor frequency stability. This may be easily understood by studying the tank impedance Z_{Tank} phase versus frequency at different temperatures shown in Fig. 4(a). A classical oscillator operates with almost zero phase across the tank. A zero phase intersection with the tank phase curves over temperature results in a large frequency variation Δf_1. However, studying the tank impedance phase curves carefully one can notice that tank phase curves across temperature intersect at a negative phase where frequency variation across temperature is minimized by a factor of ~100 to be Δf_{NULL}=47ppm as shown in Fig. 4(b). This phenomenon of minimum temperature sensitivity of frequency will be named as the LC-tank Temperature Null (TNULL) and the phase across the tank as φ_{NULL}. This non-traditional phase operating point is very desirable since it is an example of a Self Compensated Oscillator (SCO) across temperature. The challenge is to design an LC-tank with a TNULL that exhibits good frequency stability and to successfully oscillate at φ_{NULL}.

The TNULL phenomenon may be mathematically derived using a simple first order model of an LC-tank shown in Fig. 5(a) where the only temperature dependent element is r, the ohmic losses of the inductor with a first order temperature coefficient α. The tank impedance phase φ_{Tank} is given by:

$$\varphi_{Tank} = \angle Z_{Tank} = \tan^{-1}\left(\frac{\omega L}{r}\right) - \tan^{-1}\left(\frac{\omega r C}{1-\omega^2 LC}\right) \quad (1)$$

The Global Null (GNULL) is defined across a temperature range of interest as the intersection of the tank phase curves at the temperature extremes as shown in Fig. 5(b). The operating phase φ_{GNULL} is the point of minimum temperature sensitivity across this temperature range. For a temperature range of $2\Delta T$ around T_0 using the simple first order tank model GNULL phase and frequency are given by:

$$\omega_{GNULL} = \frac{1}{\sqrt{LC}}\sqrt{1+\left(\frac{1-\alpha^2\Delta T^2}{L/Cr_o^2}\right)}, \quad \varphi_{GNULL} = -\tan^{-1}(2r_0 C\omega_{GNULL}) \quad (2)$$

Fig. 5(c) shows the frequency deviation across temperature at the GNULL phase. Since the simple first order model does not take into consideration many second order affects one can note the difference between plots in Fig. 4(b) and Fig. 5(c) yet the TNULL concept is evident.

IV. SCO Implementation Based on TNULL

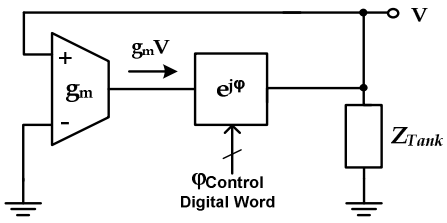
This section illustrates the architecture of the SCO based on the aforementioned LC-tank TNULL concept. The SCO functional block diagram is shown in Fig. 6. The main objective of the architecture is to operate the LC-tank at precisely the TNULL phase. The oscillator is formed of a transconductor stage and a phase shift stage. At steady state, the transconductor and phase shifter provide a phase that is the negative of the required TNULL phase. Thus, at steady state, the LC-tank is forced to oscillate at the required non-zero TNULL phase. One of the most important design aspects of the phase shift circuitry is to provide a phase that is accurate and temperature independent, otherwise frequency deviation will increase. Phase magnitude is controlled via a digital phase control signal ($\varphi_{Control}$) the accuracy of which will determine how close the final operating point to the GNULL will be. The value of $\varphi_{Control}$ is trimmed to tune the phase operating point and is stored on chip.

Fig.7 illustrates one of the methods of injecting a positive phase into the oscillator loop using a quadrature oscillator architecture with variable coupling m between the two oscillators in quadrature. Two identical LC oscillator tanks are forced to oscillate at the same frequency but in quadrature phase with respect to one another. Transconductor stages a and c inject two quadrature currents g_m*V and $g_m*V\angle 90$ into their respective tanks I and Q while stages b and d couple the oscillators together by injecting $m*g_m*V\angle 90$ and $-m*g_m*V$ currents into tanks I and Q respectively. The value of the coupling factor m is variable and is controlled via a digital control word. A gain block produces a 180 degrees phase shift to force both oscillators to be in quadrature. Therefore, the tank total current phase shift φ is a simple function of the coupling ratio m and is given by:

$$\varphi = \tan^{-1}(m) = -\varphi_{GNULL} \qquad (3)$$

The value of m is varied by scaling the bias current of the transconductor stages and the dimensions of the active devices. This guarantees producing a stable temperature independent coupling ratio and phase across both tanks.

The architecture includes an automatic amplitude control (AAC) block. The AAC block senses the oscillation amplitude of both tanks and compares it to a reference voltage (V_{REF}). The amplitude level is adjusted such that at steady state oscillator amplifiers or transconductor stages operate at a fairly linear region. This minimizes the oscillation harmonic content and the impact of this harmonic content on frequency stability. The AAC block controls all four amplifier stages to ensure that all stages have a tracking gain and a constant coupling ratio m during operation.



$$Z_{Tank} = \frac{1}{g_m} e^{-j\varphi}$$

Figure 6. SCO functional block diagram.

The SCO is trimmed using a low cost room temperature only (RTO) trimming and frequency calibration algorithm. The algorithm modulates the temperature of the LC tank using on-chip resistors, varies the value of m and detects the phase with minimum oscillator temperature sensitivity (zero slope) which may be related to the required GNULL phase. The trimming and calibration routines are done only once during production testing and the value of m is stored on chip. At power-up, the value of m is loaded to operate the oscillator at the TNULL to become a self-compensated oscillator.

V. Measured Performance

The SCO architecture described has been fabricated in a 0.18µm CMOS process with thick aluminum top metal. A set of programmable dividers produce a CMOS output frequency from 1 to 133MHz. The chip has been packaged in a standard 4-pin package. A large number of parts have been tested and characterized. Tests include total frequency stability across temperature, supply and load, single side band (SSB) phase noise, current consumption and period jitter.

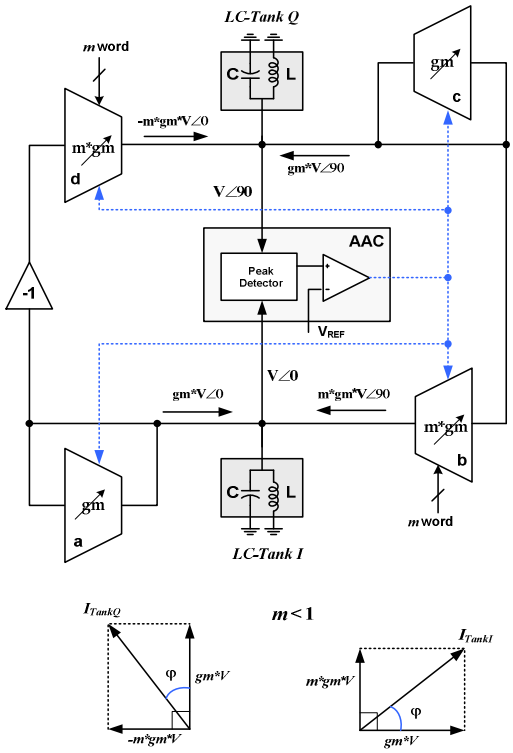


Figure 7. Block diagram of an SCO implemented using a quadrature oscillator architecture.

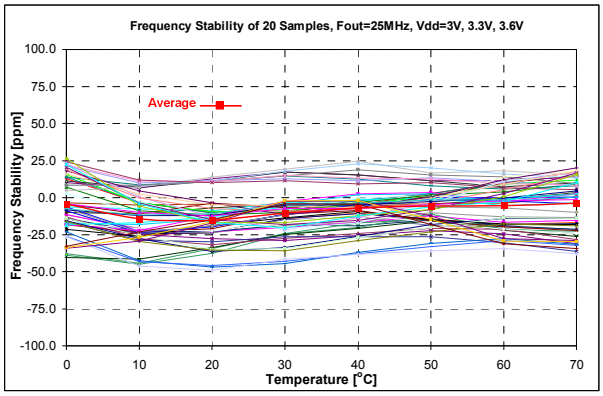


Figure 8. Frequency stability of 20 randomly selected samples.

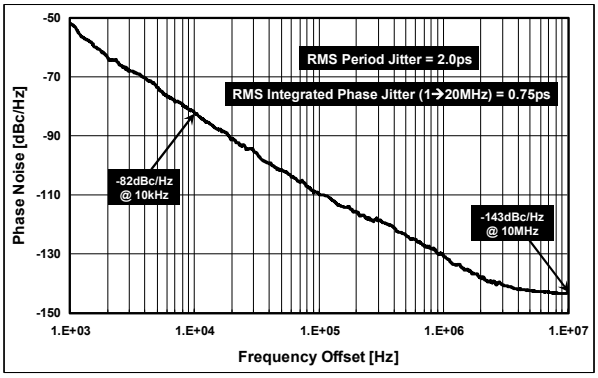


Figure 9. Phase noise measurement of a 125MHz output clock

Parts were randomly selected from a single wafer, trimmed and calibrated to a frequency of 25MHz at the CMOS output. Parts were inserted in a temperature chamber where the frequency was measured from 0 to 70°C in steps of 10°C. The power supply was varied by ±10% relative to a nominal 3.3V value. The measured total frequency stability for 20 parts is plotted in Fig. 8. Results show very good correlation between measurements and simulations indicating the robustness of the architecture and validity of the TNULL concept. Frequency stability across a capacitive load varying from 1pF to 15pF was tested and showed a maximum frequency variation across temperature of <10ppm. Frequency stability measurements illustrate clearly a very tight performance spread across the measured parts with an overall frequency stability of better than ±50ppm. It is to be noted that further improvements in the trimming and calibration routines can improve the initial accuracy to achieve an overall tighter frequency error of ±25ppm. The no load current consumption of the oscillator is 7.1mA at a 25MHz output frequency. The phase noise of a 125MHz output frequency was measured and is shown in Fig. 9 indicating a phase noise of -82dB/Hz at an offset of 10kHz. Processing the phase noise measurements yields 0.75ps RMS integrated phase jitter from 1MHz to 20MHz and a period jitter of 2ps. The performance of the SCO performance is compared to other all silicon CMOS LC-tank based commercially available oscillators in Table I. It is to be noted that the reported SCO frequency stability does not include aging while the other clocks include aging. It is expected that aging would increase the frequency stability by ±25ppm. Overall SCO performance is very competitive in terms of frequency stability and period jitter. Current consumption may be lowered by migrating to a smaller feature size technology such as a CMOS 0.13µm and use of higher quality factor copper integrated inductors will help decrease current and improve phase noise.

TABLE I. SCO PERFORMANCE COMPARISON

Specification	3CP0C02[8]	Si500S[6]	This work
CMOS Technology	0.13µm, Cu	0.13µm, Cu	0.18µm, Al
Frequency Range	4 – 133MHz	0.9 – 200MHz	1 – 133MHz
Total Freq Stability 0 – 70°C	±100ppm	±150ppm	±50ppm
No Load Current	2.2mA	9.7mA @ tri state o/p	7.1mA @ 25MHz
RMS Period Jitter	5ps @ 75MHz	1ps @ 200MHz	2ps @ 125MHz

VI. CONCLUSION

Self-compensation found in quartz crystals through precise selection of a mechanical cutting angle is analogous to the electrical TNULL phase. The LC-Tank TNULL concept has been analyzed, verified in simulation and validated by measurements. A Self Compensated Oscillator (SCO) has been successfully implemented based on the TNULL concept using a quadrature oscillator architecture which facilitated accurate temperature independent tank phase control and successful low cost trimming and calibration. The overall achieved frequency stability of ±50ppm across temperature (0 – 70°C), supply and load is very competitive with reported performance of all CMOS oscillators and may be further improved to achieve ±25ppm. Period jitter, phase noise and current consumption are reasonable and there is room for improvement by going to a smaller feature size technology such as a CMOS 0.13µm process with copper metal layers.

REFERENCES

[1] W.-T. Hsu, "Vibrating RF MEMS for clock and frequency reference applications," Technical Digest, International Microwave Symposium 2006, San Francisco, June 10-15, 2006.

[2] "MEMS Replacing Quartz Oscillators," SiTime Corporation, SiT-AN10010 Rev. 1.1, Sunnyvale, CA, 10 March 2009.

[3] "DSC8002 Series PureSilcon™ Programmable CMOS Oscillator," Discera Corporation, MK-Q-B-P-D-031809-06-6, San Jose CA, 16 March 2010.

[4] "SiT8003 Low Power 1-110 MHz Oscillator," SiTime Corporation, Rev. 1.5, Sunnyvale, CA, 30 May 2010.

[5] Michael S. McCorquodale, "Self-Referenced, Trimmed and Compensated RF CMOS Harmonic Oscillators as Monolithic Frequency Generators," Proc. of IEEE IFCS, May 19–21, 2008, pp. 408–413.

[6] "Si500S Single-Ended Output Siliocn Oscillator," Silicon Laboratories Inc., Rev. 1.0 5/11, Austin, TX.

[7] M. S. McCorquodale, et al., "A Silicon Die as a Frequency Source," Proc. of IEEE IFCS, June 1-4, 2010, pp. 103-108.

[8] "3CP0C02 CrystalFree™ Oscillator Preliminary Data Sheet," Integrated Device Technology Inc., San Jose, CA, April 1 2010.

A 4.6-GHz, 15-mW Frequency Synthesizer CMOS ASIC with Milli-Hertz Frequency Resolution for Miniature Atomic Clocks

Yazhou Zhao, Steve Tanner and Pierre-André Farine
Electronics and Signal Processing Laboratory
Institute of Microengineering, EPFL
2000 Neuchatel, Switzerland
yazhou.zhao@epfl.ch

Arnaud Casagrande
ASULAB
Swatch Group R&D
2000 Neuchatel, Switzerland

***Abstract*—In this paper, a 4.6 GHz frequency synthesizer integrated circuit for Cesium-based miniature atomic clocks is presented. Based on a fractional-N phase-locked loop (PLL) with sigma-delta modulator, the chip features a frequency resolution below 10^{-3} Hz at its output, allowing a tuning resolution of 10^{-13}. This value was measured on the circuit which is integrated into a 130 nm CMOS process. It consumes 15 mW, features an output phase noise of -76 dBc/Hz @ 1 kHz offset from the 4.6 GHz carrier, and provides a programmable output power from -10 to 0 dBm.**

I. INTRODUCTION

Modern wireless communication and navigation systems depend heavily on stable frequency references for synchronization and positioning. Atomic clocks are mostly used for this purpose, however they are large devices (~ 100 cm^3), power-hungry (~ several Watts) and too expensive to be deployed in a large scale or used in portable applications, such as GNSS receivers and future mobile communication devices. Recent efforts for miniaturization have led to the so-called miniature atomic clocks, offering reduced dimensions (~ 10 cm^3) and power consumptions (~ hundreds of mW) [1] as well as high precisions and capabilities of low-cost mass productions.

A low phase noise, high resolution frequency synthesizer is a needed block to implement atomic interrogation. For miniature systems, low power operation is also required. Previously, several papers discussed about different topologies of frequency synthesis for atomic interrogation. One traditional technique is depicted in [2], employing frequency multiplication with step recovery diodes (SRD). Ref. [3] describes a topology realized by a frequency chain using sampling mixers to avoid narrow-band filtering, while [4] and [5] utilize non-linear transmission line (NLTL) circuits instead of SRD during the frequency multiplication. However, the disadvantage of frequency multiplication is the large energy dissipation during the operation, leading to very low energy efficiency and high power consumption.

Fractional-N phase-locked loop (PLL) integrated frequency synthesizers are widely used in communication, and have proven high frequency tuning resolution and low phase noise ([6]-[8]). With the objectives of miniaturization, low-power operation as well as low phase noise level, this paper presents the design and realization of a 4.596 GHz frequency synthesizer integrated circuit, fabricated into a 130 nm CMOS process. Compared to [9] and [10], the chip features an un-common frequency resolution of about 10^{-13}, which allows its use in high-performance Cesium-based atomic clocks.

The paper is organized as follows: in Section II, the circuit general architecture is presented, as well as the detailed design of the main blocks. Section III highlights the realization and experimental results of the chip, while Section IV concludes the paper.

II. CIRCUIT DESIGN

Fig. 1 illustrates the general architecture of the frequency synthesizer, which includes the PLL blocks (a phase and fre-

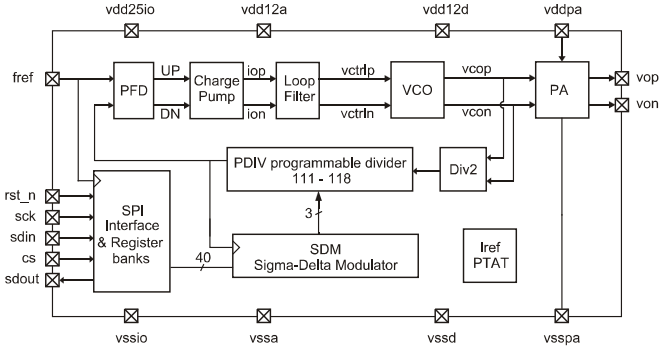


Figure 1. Systematic diagram of the frequency synthesizer.

This work was funded by the EU 7th framework program FP7, under contract no. 224132.

quency detector, a charge pump and a loop filter), a voltage controlled oscillator (VCO), a power amplifier (PA), a divide by two prescaler, a programmable divider controlled by a sigma-delta modulator, a PTAT reference for generating the bias currents, and a serial peripheral interface (SPI) to access a register bank for controlling the chip.

A. Voltage Controlled Oscillator

The VCO is a challenging block in the design of frequency synthesizers because its phase noise has a big impact on the performance of the whole chip. A LC tank VCO is naturally selected for its relatively good quality factor and low power operation. However, to improve the noise immunity of this sensitive structure, a fully-differential scheme is chosen for all elements, including the varactor, permitting to reduce the common-mode noise and to maintain a symmetrical topology of the circuit as shown in Fig. 2.

The VCO LC tank consists of a fully differential inductor with a varactor load for frequency fine tuning, and a 5-bit capacitor bank for coarse tuning. NMOS transistors form the cross-coupled pair to feature the function as a negative impedance convertor for sustaining the oscillation. PMOS transistors are used to construct the current biasing block. In comparison to NMOS transistors, they can provide lower flicker noise. Filtering capacitors are connected between the PMOS gate and VDD to stabilize V_{GS} in order to reduce the noise influence from the voltage supply.

Several other measures are taken to improve the VCO performance. The symmetry in the varactor allows to improve its linearity. For the switches of the capacitor bank, bias resistors are added to control the DC voltage at the switch terminals to ensure a minimal and constant parasitic capacitance of both the source and drain diffusions.

The varactor dimensioning is the result of a trade-off: on one side, it should be large enough to ensure continuous frequency tuning between successive coarse tuning codes and to allow an operation in the middle of its dynamic range, where the linearity is optimal. On the other side, the varactor value should be minimized to reduce phase noise degradation due to circuit noise (substrate, etc.) coupling into it. A compromise is selected with a varactor tuning range corresponding to three coarse tuning steps.

Concerning the cross-coupled transistors, the minimal length is chosen to reduce the parasitic capacitance, while the width is calculated to get sufficient transconductance g_m for sustaining the oscillation, according to Barkhausen's criterion. For the current biasing transistor MS, its size is selected to limit the flicker noise as well as to mitigate the layout mismatch influence.

B. Frequency Prescaler

The divide-by-2 prescaler, shown in Fig. 3, is directly connected to the RF output of the VCO; therefore, it requires a high frequency operation and symmetrical structure. To meet these constraints, a current-mode logic (CML) D-flip-flop is used since its reduced switching amplitude allows the circuit to be operated in high-speed conditions.

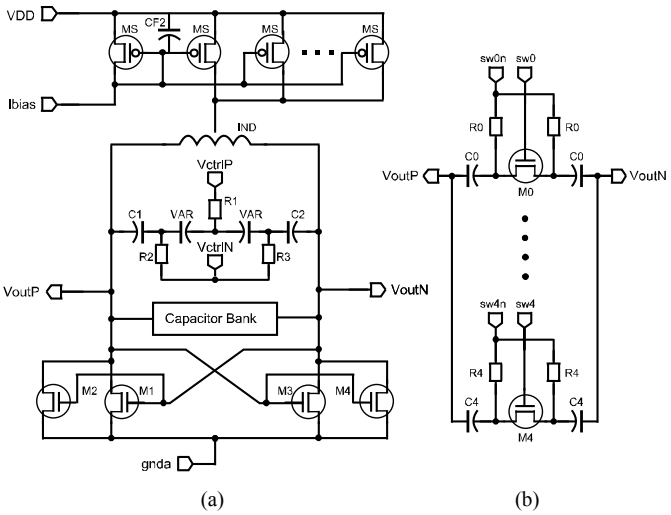


Figure 2. Schematics of (a) the fully differential VCO and (b) the coarse tuning capacitor bank.

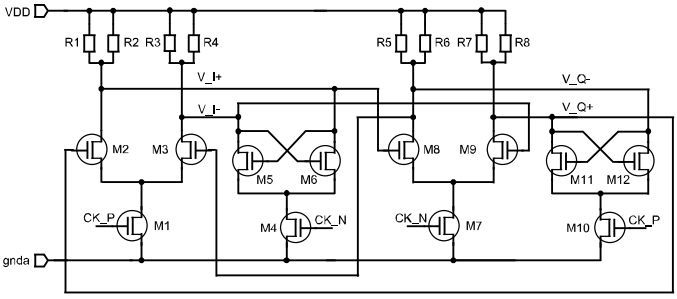


Figure 3. Schematic of the divide-by-2 prescaler.

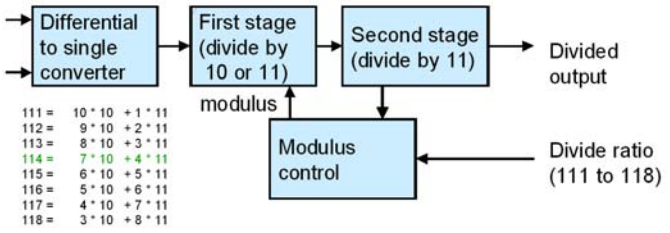


Figure 4. Circuit diagram of the programmable divider.

The two complementary outputs of each D-flip-flop are used to drive a CML-to-CMOS converter so that the power of these four quadrature signals is exploited, thus enhancing the power efficiency. This also provides an equal load to both D-flip-flop outputs that improves the symmetry of the circuit.

To improve the output dynamic swing, the tail current transistor which is normally found in this circuit is suppressed. Instead, a bias current is applied to a single transistor which provides the two functions of switching and biasing.

C. Programmable Divider

A programmable divider is needed for the design of a fractional-N synthesizer. Fig. 4 shows the architecture of the circuit which comprises three stages: a dual-modulus divider 10/11, a programmable counter (divide-by-11) and a modulus

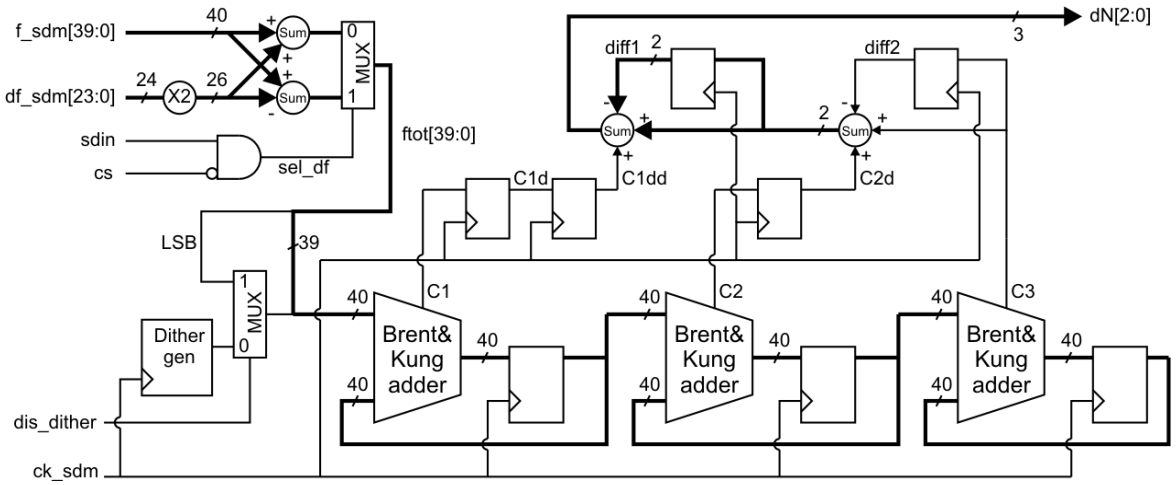


Figure 5. Schematic of the sigma-delta modulator structure.

control (A-counter). The dual-modulus divider divides the RF signal either by 10 or 11 depending on the modulus control input from the A-counter. Its output signal is divided by certain ratio (<11) again after passing the programmable counter to obtain the wanted output frequency, which is also used to reset the A-counter for restart. The A-counter is reduced by 1 each time it is triggered by the clock and sends different modulus signals to control the division ratio of the dual-modulus divider. Considering the high operating frequency of 2.3 GHz, the circuit of this block is realized with dynamic logic.

In this chip, the programmable divider takes the 2.3 GHz signal from the divide-by-2 prescaler and further divides it with an average ratio of about 114.9 to get the desired 20 MHz feedback frequency. Its integer division ratio varies from 111 to 118, controlled by a 3-bit word provided from the sigma-delta modulator.

D. Sigma-Delta Modulator

The sigma-delta modulator (SDM) is used to modulate the programmable divider to achieve a fractional division ratio. Fig. 5 shows the schematic of a 3rd-order, MASH SDM structure made with three 40-bit accumulators, which are built of a Brent & Kung adder and a 40-bit register for each. The 3-bit carry outputs C1, C2 and C3 are combined to calculate the SDM output dN[2:0].

The designed SDM is simulated and provides a low power spectrum density within the PLL bandwidth (up to 250 kHz). Above, the noise rises rapidly but can be filtered out by the PLL loop filter.

III. CIRCUIT REALIZATION AND EXPERIMENTAL RESULTS

The frequency synthesizer integrated circuit is implemented into a 130 nm RF CMOS process. It occupies an active area of about 2 mm^2. The chip microphotograph is shown in Fig. 6. The total power consumption of the chip is 15 mW at the nominal supply voltage of 1.2 Volt.

The tuning range of the RF signal is measured to be from 4.07 GHz to 5.17 GHz, well centered on the 4.6 GHz nominal frequency. The K_{VCO} is measured from 113 MHz/V to 203 MHz/V, which is around three times larger than the expected

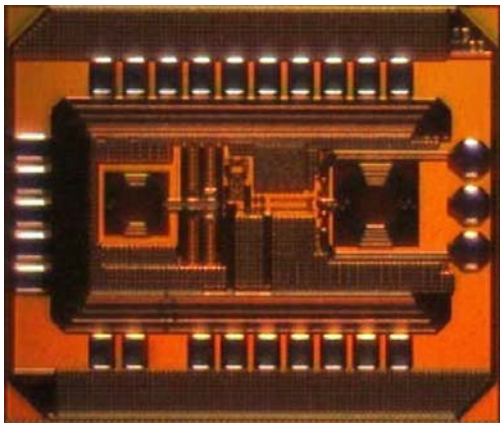


Figure 6. Microphotograph of the die with an active area of about 2 mm^2.

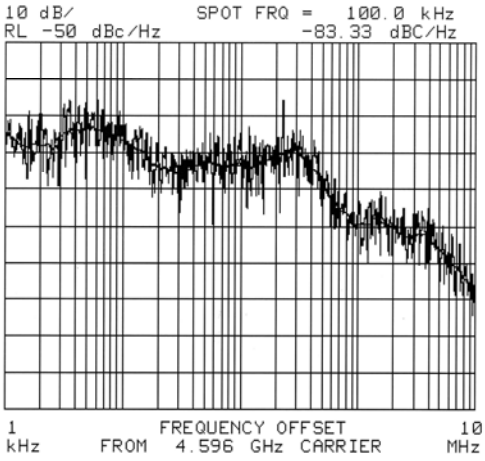


Figure 7. Curve of the synthesizer phase noise.

value. The reason is due to an over-estimation of the varactor value (no model available).

The phase noise curve is shown in Fig. 7, giving a noise level of -76 dBc/Hz @ 1 kHz offset and -83 dBc/Hz @ 100 kHz offset from the carrier. This noise is higher than the ex-

TABLE I. CRITICAL EXPERIMENTAL RESULTS OF THE CHIP COMPARED TO THEIR EXPECTED VALUES.

	Expected Values	**Experimental Results**
PLL BW	120 kHz ~ 300 kHz	100 kHz ~ 400 kHz
Freqency Resolution	10^{-13}	10^{-13}
Power Consumption	15 mW	15 mW
Phase Noise at 1 kHz offset	-77 dBc/Hz	-76 dBc/Hz
Phase Noise at 100 kHz offset	-89 dBc/Hz	-83 dBc/Hz

pected value of -89 dBc/Hz @ 100 kHz offset. It has been verified that the excess of phase noise inside the PLL bandwidth is generated by digital switching noise coming from the on-chip logic and the input level shifter on the 20 MHz reference clock. Indeed, lower phase noise is found when the sigma-delta modulator is switched off. Furthermore, measurements of the VCO alone give a phase noise of -92 dBc/Hz @ 200 kHz offset from the carrier, which is much lower than the closed-loop noise at the same offset position in Fig. 7. This shows that the limited performance comes from the PLL sys-tem, especially from digital switching noise instead of the VCO.

The frequency resolution of the synthesizer is measured by means of a self-down-mixing system. Since the specified frequency resolution value is about 10^{13} times smaller than the 4.596 GHz signal (<1 mHz) it is difficult to make a direct measurement for such a tiny frequency variation to be observed at the RF output. Therefore, down-mixing the 4.596 GHz into the baseband frequency regime can lead to a much easier measurement procedure. In order to eliminate the need for an accurate reference frequency generator, two ASICs using the same input reference clock are used to generate two slightly different RF frequency signals, which are mixed together. The resulting beat frequency is measured in function of the change in the sigma-delta modulator input register. Frequency differences of less than 0.5 mHz are measured, corresponding to a resolution of 10^{-13}, which meets the specifications. The critical experimental results of the full chip measurement are summed up in Table I.

IV. CONCLUSION

A 4.6 GHz frequency synthesizer implemented into a 130 nm RF CMOS process for Cesium-based miniature atomic clocks application is presented. A 3rd-order, 40-bit sigma-delta modulator is used to provide the fractional ratio for the synthesizer to generate the 4.596 GHz signal with a frequency resolution higher than 10^{-13}. The power consumption of the circuit is less than 15 mW and the chip occupies an active area of 2 mm^2. The measured phase noise is -76 dBc/Hz @ 1 kHz offset and -83 dBc/Hz @ 100 kHz offset. These performances allow using this chip in chip-scale atomic clocks with stabilities in the order of 5×10^{-11}. For better stabilities, the phase noise needs further improvements. This can be done by using a smaller varactor and decreasing the on-chip digital noise.

ACKNOWLEDGMENT

This work was funded by the EU 7th Framework Program (FP7), under contract No. 224132. The authors thank ASULAB SA and the Laboratory of Time and Frequency (LTF) of the University of Neuchâtel for measurement support.

REFERENCES

[1] R. Lutwak, P. Vlitas, M. Varghese and M. Mescher, "The MAC – a miniature atomic clock", in *Proc. IEEE Int. Freq. Control Symp. Expo.*, 2005, pp. 752 – 757.

[2] R. Barillet, V. Giordano, J. Viennet and C. Audoin, "Limitation of the clock frequency stability by the interrogation frequency noise: Experimental results", *IEEE Trans. Instrum. Meas.*, vol. 42, no. 2, pp. 276 – 280, Apr. 1993.

[3] G. D. Rovera, G. Santarelli and A. Clairon, "Frequency synthesis chain for the atomic fountain primary frequency standard", *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 43, no. 3, pp. 354 – 358, May 1996.

[4] R. Boudot, S. Guérandel and E. de Clercq, "Simple-design low noise NLTL-based frequency synthesizers for a CPT Cs clock", *IEEE Trans. Instrum. Meas.*, vol. 58, no. 10, pp. 3659 – 3665, Oct. 2009.

[5] F. R. Martinez, M. Lours, P. Rosenbusch, F. Reinhard and J. Reichel, "Low phase noise frequency synthesizer for the trapped atom clock on a chip", *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 57, no. 1, pp. 88 – 93, Jan. 2010.

[6] S. Pamarti, L. Jansson and I. Galton, "A wideband 2.4-GHz delta-sigma fractional-N PLL with 1-Mb/s in-loop modulation", *IEEE J. Solid-State Circuits*, vol. 39, no. 1, pp. 49 – 62, Jan. 2004.

[7] H. R. Rategh, H. Samavati and T. H. Lee, "A CMOS frequency synthesizer with an injection-locked frequency divider for a 5-GHz wireless LAN receiver", *IEEE J. Solid-State Circuits*, vol. 35, no. 5, pp. 780 – 787, May, 2000.

[8] W. Rhee, B. Bisanti and A. Ali, "An 18-mW 2.5-GHz/900-MHz BiCMOS dual frequency synthesizer with < 10-Hz RF carrier resolution", *IEEE J. Solid-State Circuits*, vol. 37, no. 4, pp. 515 – 520, Apr. 2002.

[9] W. Rhee, B. Song and A. Ali, "A 1.1-GHz CMOS fractional-N frequency synthesizer with a 3-b third-order ΔΣ modulator", *IEEE J. Solid-State Circuits*, vol. 35, no. 10, pp. 1453 – 1460, Oct. 2000.

[10] T. P. Kenny, T. A. D. Riley, N. M. Filiol and M. A. Copeland, "Design and realization of a digital ΔΣ modulator for fractional-N frequency synthesis", *IEEE Trans. Vehicu. Tech.*, vol. 48, no. 2, pp. 510 – 521, Mar. 1999.

A Novel High Frequency Synthesizer Using Adaptive Injection Mode-Coupled VCSO for Low Jitter and Low Noise Applications

Ajay K. Poddar[1,2], *Senior Member IEEE*
[1]Synergy Microwave Corp., NJ 07504, USA
[2]Technical University Munich, Germany

Ulrich L. Rohde[1,2], *Fellow IEEE*
[1]Univ. of Cottbus, BTU Cottbus 03046, Germany
[2]Synergy Microwave Corp., NJ 07504, USA

Abstract— **Injection mode-coupled voltage-controlled SAW oscillators (VCSOs) are versatile performers in low-noise frequency synthesizers, both as high-frequency sources and as reference oscillators. A novel high frequency digital synthesizer is reported with the intent of lowering following: size, cost, power-consumption, phase noise, jitter, and vibration sensitivity. The measured phase noise is typically better than -100 dBc/Hz, -115 dBc/Hz, -129 dBc/Hz, -133 dBc/Hz at 100 Hz, 1 kHz, 10 kHz, and 100 kHz offset frequencies respectively from the carrier frequency of 5200 MHz. The reported synthesizer can generate frequencies from 5 GHz to 5.3 GHz in steps of 100 kHz with typical spurious of -65 dBc and the power consumption is less than 300 mill watt for a given 5 dBm output power.**

I. INTRODUCTION

Frequency synthesizers come in many shapes and sizes. They range from tiny system-on-a-chip (SoC) devices and modules to rugged military-grade rack mount systems and bench-top instruments [1]-[20]. They are also based on a wide array of technologies, from analog to digital in nature, starting with many different signal sources, including voltage-controlled oscillators (VCOs) and surface-acoustic-wave (SAW) resonators. One of the more novel approaches combines both analog and digital techniques: a frequency synthesizer based on an injection mode-coupled voltage-controlled saw oscillator (VCSO) circuits. The novel VCSO using injection mode-coupling approach is part of a frequency synthesizer capable of typical measured phase noise of -152 dBc/Hz offset 10 kHz from a 1-GHz carrier, with less than 6 fs RMS jitter for a given power consumption of 60 mill watt (DC bias operating condition: 5V, 12 mA).

Historically synthesizer designers have relied on crystal based multiplied reference oscillator featuring excellent phase noise performance at 1GHz and above frequencies. Although, frequency multiplier crystal oscillator can offer low phase noise performance at 1 GHz and above frequencies but at the cost of large size, high noise floor and high power-consumptions. The most dramatic change will occur when optimum loop bandwidth extends to a few megahertz, where adaptive injection mode mode-coupled VCSO's phase noise performance becomes competitive or supersede the performance of the conventional multiplied crystal oscillator at 1GHz and above. Furthermore, these characteristics, accompanied with the low cost, low power consumption and compact size inherent to the VCSO topologies, are likely to secure their domination in the foreseeable future at higher frequency synthesizers and reference frequency standards.

II. STATE-OF-THE-ART SYNTHESIZER TOPOLOGIES

A typical phase-lock-loop (PLL) synthesizer consists of a reference phase detector, loop filter, VCO and VCO divider (N divider) as shown in Fig. 1. The PLL is typically available as an integrated circuit (IC) with a reference divider, phase detector or phase frequency detector (PFD) with charge pump, and the VCO divider. The relationship between the output frequency 'f_{out}' and reference frequency 'f_{ref}' can be described by [14]

$$f_{out} = N\left[\frac{f_{ref}}{R}\right] \quad (1)$$

As shown in Fig.1, the PFD compares the two input signals (f_{ref} /R) and (f_{out} /N) produces an error voltage proportional to the phase difference between them. The loop filter attenuates the high-frequency noise components from the PFD output and limits the bandwidth of the error signal. The filtered error voltage is applied to the tuning port of the VCO; the error signal drives the VCO frequency, f_{out} higher or lower in frequency until the error voltage at the PFD output is zero.

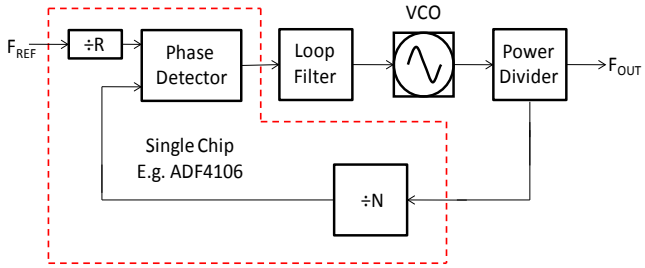


Fig.1. This simplified block diagram shows a simple, single-loop integer-N PLL module.

Typically, the VCO divider is implemented as a dual-modulus counter to obtain large continuous division of the VCO output. To vary the output frequency f_{out} of the synthesizer, the value of N is changed, and it can be seen from (1) that the minimum frequency step (defined as the step size) obtained at the output is given by f_{ref}/R.

Smaller step size can be obtained by operating the PFD at low frequency - i.e., by increasing the R value (and thus increasing N). The problem with increasing the N value is the noise of the PFD increases. The close in phase noise of the synthesizer is estimated by noise of the synthesizer itself (provided by the manufacturer), PN_{SYN} and adding 20logN (where N is the N divider value) and 10log (f_{PFD}) [14]:

$$PN_{TOT} = PN_{SYN} + 10log(f_{PFD}) + 20\log(N)$$

$$= PN_{SYN} + 10\log\left(\frac{f_{REF}}{R}\right) + 20\log(N) \qquad (2)$$

From (2), to minimize the phase noise, the value of N should be kept small. But for fine tuning resolution, the PFD must operate at low frequency, requiring a high value for N and resulting poor noise performances.

The conflict eliminates the conventional PLL design for use in low-noise, high-resolution applications. The loop filter in a PLL synthesizer design also limits switching speed. A higher loop bandwidth yields faster switching speed but with excessive noise. A narrow loop bandwidth lowers phase noise but sacrifices switching speed. The tradeoff is between phase noise and switching speed. Typical phase noise for a conventional PLL synthesizer operating from 500 to 1200 MHz in 500-kHz steps includes switching time of less than 5 ms and better than -70 dBc spurious levels with a 10-MHz reference (Fig. 2).

To improve the performance of an integer-based PLL (as shown in Fig. 1) the 'N' divider can be implemented as a fractional divider instead, rather than an integer divider. This enables the use of higher PFD frequency and lower N values, thus improving the phase noise performance of the synthesizer.

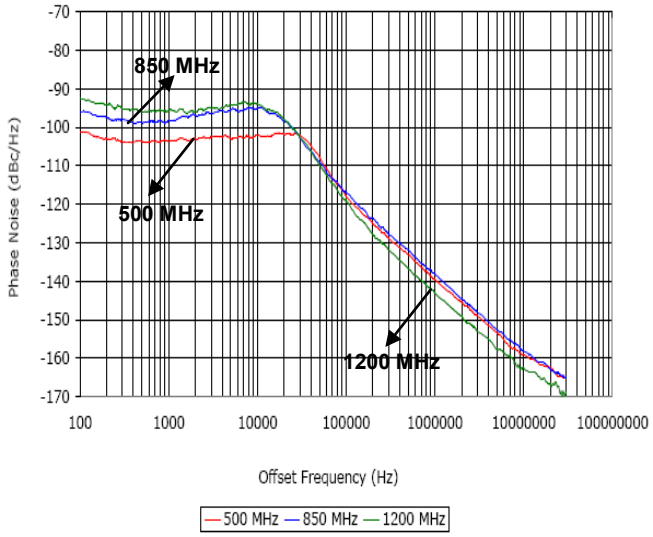


Fig.2. This plot shows the typical measured phase noise for an integer-N frequency synthesizer.

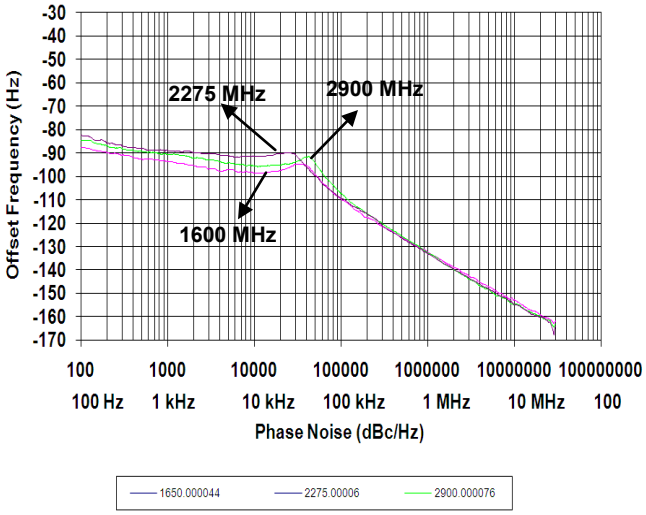


Fig.3. This plot shows the typical measured phase noise for a fractional-N frequency synthesizer,

But a fractional-N divider also introduces spurious at the output of the synthesizer. In addition to this, even with the fractional N the step-size of the synthesizer cannot go below 1 Hz. Typical fractional-N synthesizer performance from 1600 to 2900 MHz with a 10-MHz includes better than 1-ms settling time and spurious levels of -60 dBc or better (Fig. 3).

Multiloop synthesizer architectures with direct-digital synthesizers (DDSs) have been used for tuning resolution better than 1 Hz. Figure 4 shows a typical dual-loop synthesizer where a DDS module combines with a conventional PLL synthesizer. The DDS clock is fed by a selectable frequency synthesizer which provides a fixed set of frequencies based on the parallel select lines.

Unfortunately, the DDS approach can generate high spurious output levels. But these spurs can also be predicted and shifted by proper choice of clock frequency. If spurious products fall within the loop bandwidth, by switching the clock frequency, the spurs can be shifted to a band where they can be attenuated. In this dual-loop approach, the output synthesizer's PFD can be high to improve phase noise while still delivering 1-Hz resolution with a 10-MHz reference. Spurious levels are typically better than -70 dBc from 1100 to 2500 MHz with 1-Hz step size, better than 1-ms switching speed, and low phase noise characteristics (Fig. 5).

As briefly discussed above, designers have often relied on multiplying a low-noise crystal oscillator based VCO to achieve higher-frequency (1 GHz and higher) output signals.

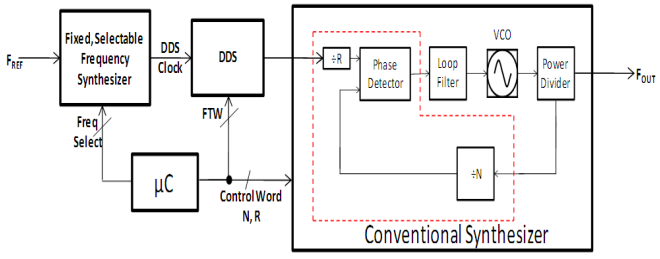


Fig.4. This simplified block diagram shows a DDS-based dual-loop frequency synthesizer.

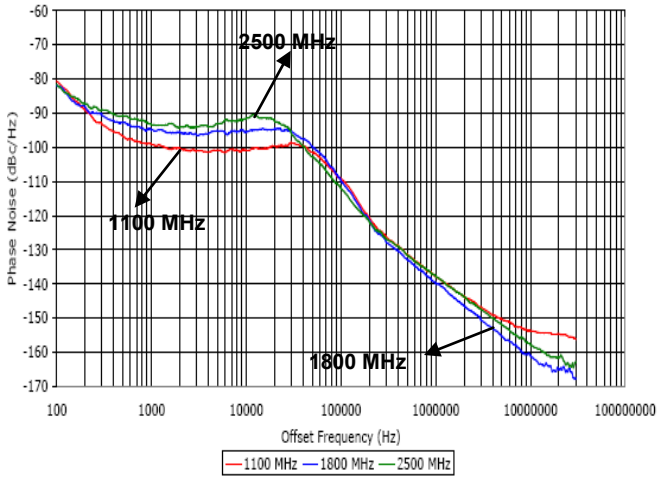


Fig.5. This is typical measured phase noise for a DDS-based dual-loop frequency synthesizer

The crystal oscillator multiplier VCO approach can provide excellent phase noise, but such synthesizers tend to consume a great deal of power, generate high levels of noise, and are large in size. Commercially available crystal oscillators at 100 MHz and below are available with phase noise of -160 to -176 dBc/Hz offset 20 to 100 kHz from a 100-MHz output. When multiplied, the phase noise translates to -140 to -156 dBc/Hz offset 20 to 100 kHz from 1 GHz but degrades the noise floor by as much of 20 dB.

The availability of low-cost, low-noise VCSOs clears the way for small, low-noise synthesizers. Part of using these sources involves modeling the surface acoustic wave resonators under large signal drive conditions for the better insights about noise dynamics at close-in phase noise, then developing manufacturable methods for producing high-purity and temperature stable oscillators in chip form.

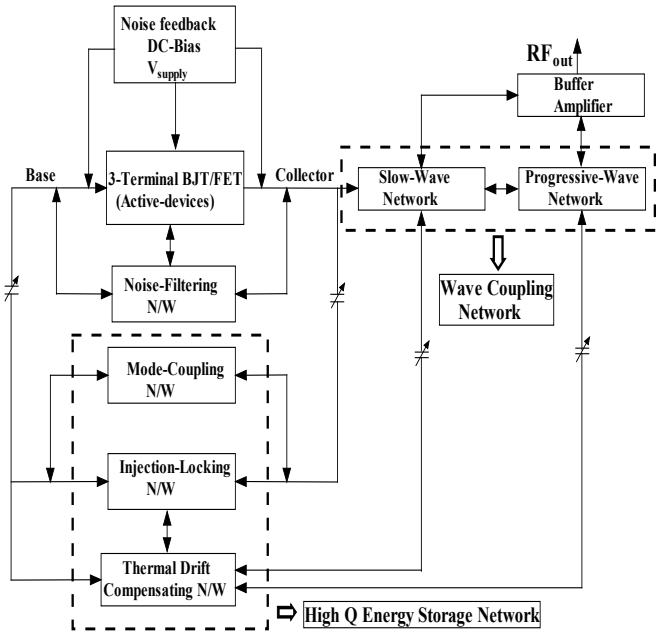


Fig.6. This block diagram shows the functional parts of an adaptive injection mode-coupled VCSO for low cost high performance synthesizer circuit.

Fig. 6 shows a typical block diagram for an adaptive injection mode-coupled VCSO optimized for low phase and noise with 400 kHz tuning require for compensating the frequency drift due to temperature and ageing, while Fig. 7 compares its phase noise at 1 GHz to that of a commercially available multiplied crystal oscillator circuit.

An adaptive mode-coupled VCSO combined with a digital PFD can be used to form a low-cost, low-power synthesizer module. The digital PFD has a wide capture range, although with high noise floor. An analog PFD has lower noise floor, but narrow capture range. To create a low-noise synthesizer with a wide capture range, a digital PFD can be used in parallel with an analog phase detector. The digital PFD is used for initial acquisition, and the analog detector for final tuning.

For example, within a DDS architecture, the VCSO module can be used to create a synthesizer (Fig. 8) with fast switching speed (about 200 µs). The DDS is used as the feedback divider for high resolution and flexibility in choosing a reference frequency. From 530 to 630 MHz, spurious levels are better than -75 dBc with low phase noise (Fig. 9). Since the analog phase detector has a low noise floor, the synthesizer bandwidth was kept large for fast switching time with low phase noise.

III. Frequency Synthesizer Using Adaptive Injection Mode-Coupled VCSO Circuit

A high-performance synthesizer based on an adaptive mode injection-locked VCSO and multiple loops, including a sampling PLL, is shown in Fig. 10.

The main PLL (Fig. 10b) achieves a coarse lock to the desired output frequency using a 120-MHz reference signal, which can be derived from an internal OCXO or low-noise 800-MHz VCSO.

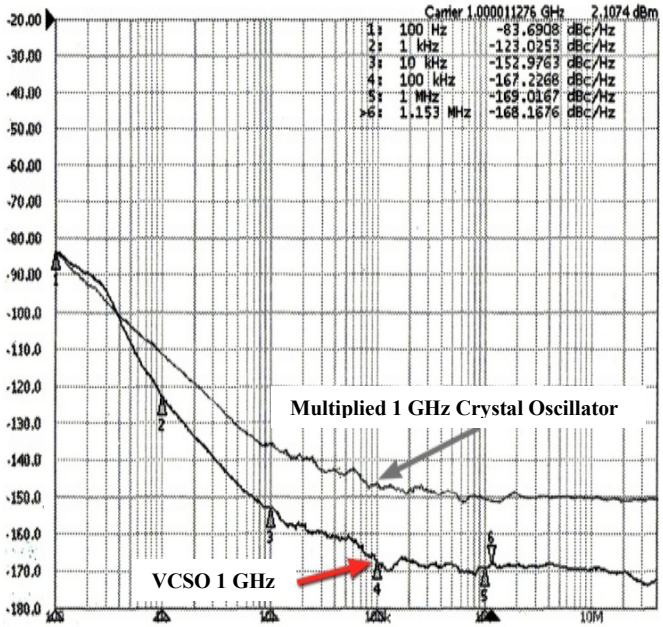


Fig.7. This phase noise plot compares the performance of a 1-GHz VCSO with 1-GHz achieved by multiplying a crystal oscillator.

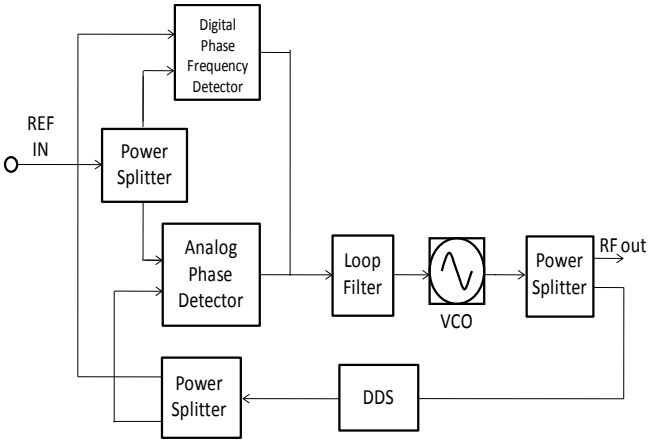


Fig.8. Shows the typical block diagram of DDS based high performance fast switching synthesizer.

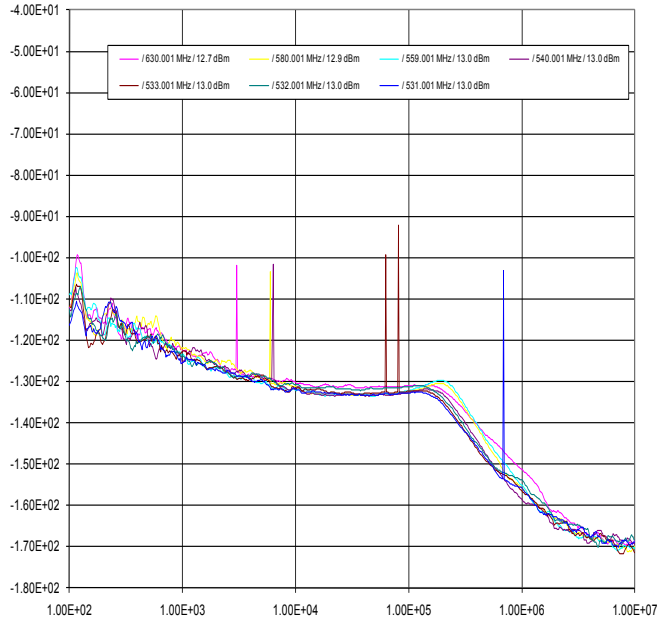
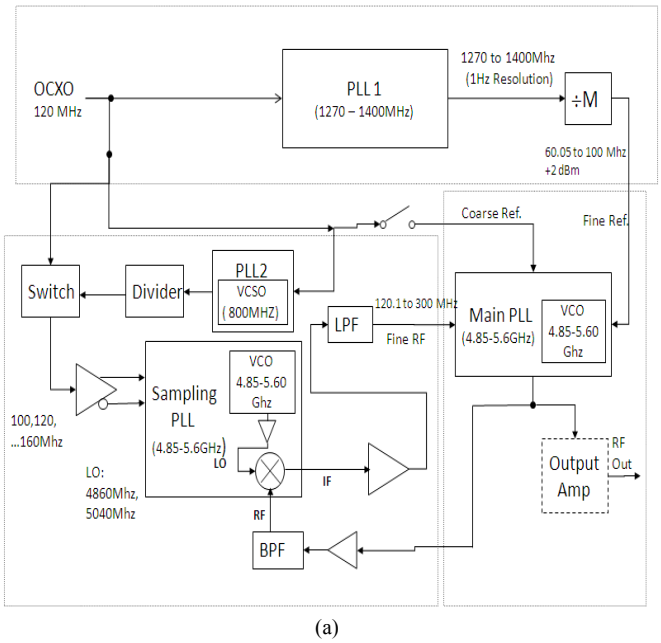


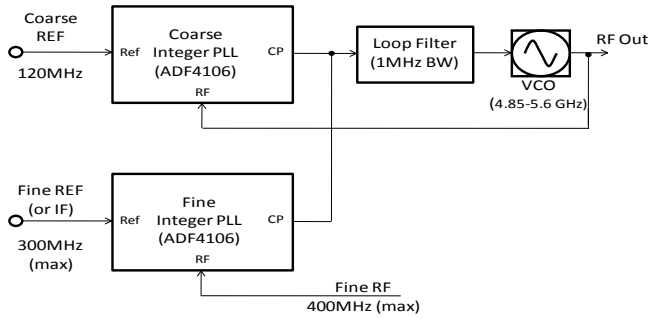
Fig.9. This plot shows the phase-noise of the DDS-based fast-switching frequency synthesizer.

The fine lock is achieved by switching to a fine loop with a reference from PLL1 and fine RF signal from the sampling PLL (Fig. 10c). The sampling PLL features a sampling phase detector with low noise floor. It synthesizes frequencies 120 to 300 MHz away from a desired output frequency by multiplying a reference input signal (from the OCXO or VCSO).

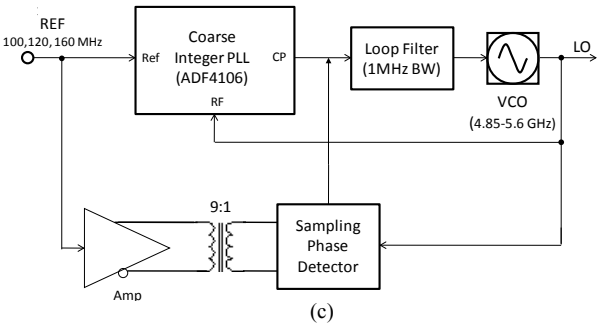
Software is used to choose the appropriate reference for best output spurious performance. The fine RF signal is the difference between RF output and the sampling PLL output (generated by mixing the two and filtering the difference). The fine reference to the main PLL is generated from PLL1. The software determines the PLL1 output frequency and appropriate division ratio (M divider) to achieve the desired output frequency. The synthesizer generates 5000 to 5300 MHz in 100-kHz steps (Fig. 10d) with better than -65 dBc spurious (Fig. 11) and exhibits low phase noise characteristics (Fig. 12).



(a)



(b)



(c)

(d)

Fig.10. This synthesizer architecture (a) uses a sampling PLL with a VCSO providing reference signals, with a (b) main PLL and (c) sampling PLL as part of the design (d) layout and package of the synthesizer module (1.2 x 0.9 inches).

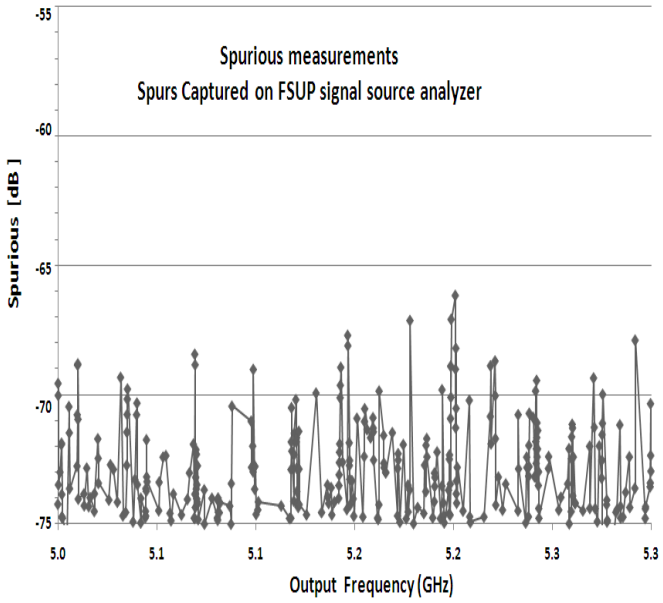


Fig.11. The spurious for the VCSO-based synthesizer with sampling PLL is specified better than -65dBc over the frequency bands (5000-5300MHz)

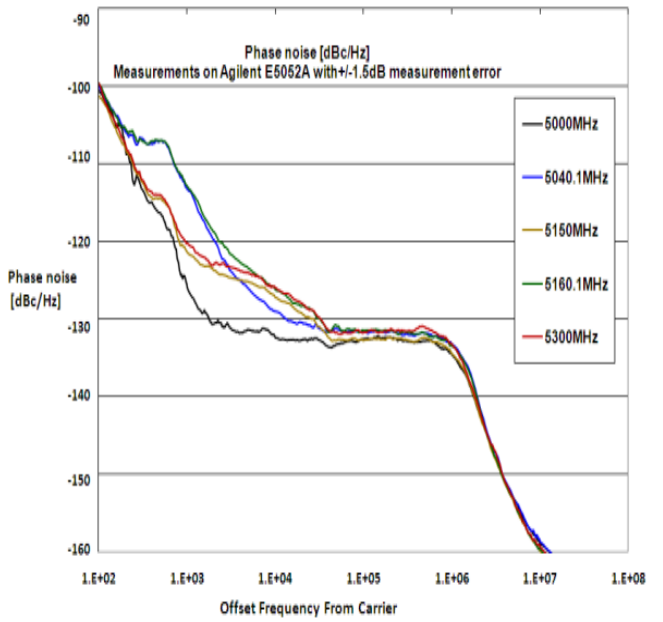


Fig.12. The phase noise for the VCSO-based synthesizer with sampling PLL is specified better than -130 dBc/Hz at 100 kHz from the carrier over the frequency bands (5000-5300 MHz).

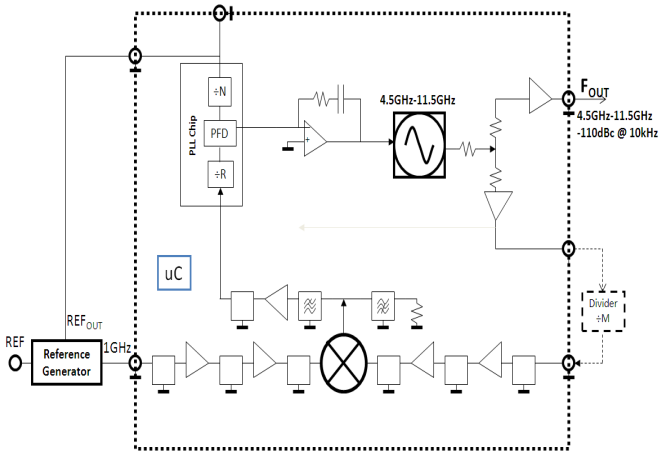
Adaptive mode-coupled VCSOs can also be used in low-spurious synthesizer designs for applications that may require stabilized signals with low levels of spurious noise (Fig. 13).

In such a frequency synthesizer design, the reference generator multiplies a reference frequency (such as 10 MHz) to 1 GHz through a 1-GHz VCSO-based PLL. This 1-GHz signal is mixed with a signal (20 to 150 MHz) from a DDS or flying-adder-based synthesizer to produce a reference signal for the main PLL from 1020 to 1150 MHz [7]-[11].

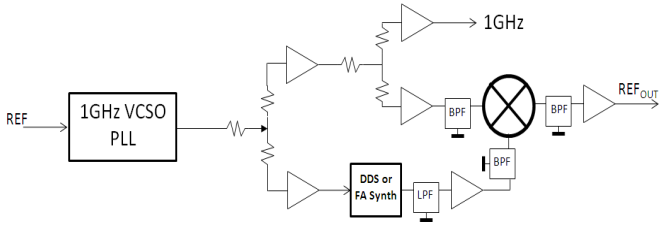
The main PLL is basically a loop that translates a low-noise reference to higher frequencies from 4.5 to 11.5 GHz. The programmable divider (divide-by-M or M-divider) can divide by 4, 5, 6, 7, 8, or 9 for frequencies ranging from DC to 15 GHz. The divider output is fed to the translational loop where this signal is mixed with the 1-GHz low-noise reference and fed as a reference to a PLL chip. The reference from the reference generator is applied to the RF port of the PLL chip. A required frequency is obtained at the output by means of proper reference (R) and divider (N) settings.

It can be seen from Table 1, the division ratios required to implement a high frequency output is low leading to very low phase noise performance of the synthesizer. The advantages of this topology are: design is simple and small enabling small size packaging, performance over size ratio delivered by the architecture is very high, and architecture uses comparatively low cost components.

The drawbacks of this topology are: phase noise is limited by the noise floor of the digital PFD, and even though the division ratios used in the design is low, the noise floor of the dividers degrades the phase noise characteristics.



(a)



(b)

Fig.13. (a) Shows the block diagram main PLL, and (b) the reference generator for the main PLL.

Table1. Describes various settings for different output frequencies

Output Frequency (GHz)	RF Divider (UXM15P)	DDS/FA output frequency (MHz)	R divider	N divider
4.5	4	62.5	2	17
5	4	83.33	3	13
8	7	71.428	2	15
9	8	62.5	2	17
10	8	83.33	3	13
11	9	111.11	3	15

IV. CONCLUSION

By the use of a high performance adaptive synthesis and injection mode-coupling mechanism, low cost, high frequency, and high performance digital frequency synthesizer module is reported in compact package without compromising the phase noise performance, jitter and vibration characteristics.

REFERENCES

[1] Vankka and K. Halonen, Direct Digital Synthesizers: Theory, Design *and Applications*. New York: Springer-Verlag, 2006.

[2] Cicero S. Vaucher, *Architectures for RF frequency synthesizers*, Boston: Kluwer Academic Publishers, 2002.

[3] U. L. Rohde and A. K. Poddar, "Frequency Generation and Synthesis: Configurable, Concurrent, Cost-Effective and Power-Efficient Solutions", ISMOT-2009, 16 - 19, Dec.r 2009, New Delhi, India.

[4] U. L. Rohde and A. K. Poddar," Low Thermal Drift Tunable Frequency Voltage Controlled Oscillatory", US Patent No.7, 262,670 B2, Aug 28, 2007.

[5] U. L. Rohde and A. K. Poddar," Low Noise, Hybrid Tuned Wideband Voltage Controlled Oscillator", *US Patent No*. 7,365,612 B2, April 29, 2008.

[6] U. L. Rohde and A. K. Poddar, "User-Definable, Low Cost, Low Phase Hit and Spectral Pure Tunable Oscillator", U.S. Patent No. 7,586,381 on September 08, 2009.

[7] P. Sotiriadis, "Intrinsic jitter of flying-adder frequency synthesizers," in Proc. IEEE Int. Sarnoff Symp., 2009, pp. 1–4.

[8] P. Sotiriadis, "Timing and spectral properties of the flying adder frequency synthesizers," in Proc. IEEE Int. Freq. Control Symp., Jul. 2009, pp. 788–792.

[9] Paul P. Sotiriadis, "Theory of Flying-Adder Frequency Synthesizers—Part I: Modeling, Signals' Periods and Output Average Frequency". IEEE Transactions on circuits and systems-I: Regular papers, VOL. 57, NO. 8, AUGUST 2010

[10] Paul P. Sotiriadis, "Exact Spectrum and Time-Domain Output of Flying-Adder Frequency Synthesizers", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, no. 9, September 2010.

[11] Paul Peter Sotiriadis, "Cascaded Diophantine Frequency Synthesis", IEEE Transactions on circuits and systems-I: Regular papers, VOL. 55, NO. 3, APRIL 2008

[12] A. Chenakin, " Frequency Synthesis: Current SolutionsAnd New Trends", Microwave Journal, pp. 256-266, May 2007.

[13] J. Browne, " Frequency Synthesizers Tune Communication Systems, " Microwaves & RF, March 2006.

[14] U. L. Rohde, " Microwave and Wireless Synthesizers: Theory and Design," John Wiley & Sons Inc., NJ.

[15] W. Egan, Frequency Synthesis by Phase Lock, Juhn Wiley & Sons Inc., Hoboken, NJ 2000.

[16] S. Pamarti and S. Delshadpour, "A spur elimination technique for phase interpolation-based fractional-N PLLs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 6, pp. 1639-1647, Jul. 2008.

[17] Y. Sun, *et al.*, "Dual-path LC VCO design with partitioned coarsetuning control in 65nm CMOS," to appear in *IEEE MWCL*, Mar. 2010.

[18] X. Yu *et al.*, "A ⊗© fractional-*N* frequency synthesizer with customized noise shaping for WCDMA/HSDPA applications," *IEEE J. Solid-State Circuits,* vol. 44, Aug. 2009, pp. 2193-2201.

[19] T. Wu, P. Hanumolu, K. Mayaram, and U. Moon, "A 4.2 GHz PLL frequency synthesizer with an adaptively tuned coarse loop," in *Proc. IEEE CICC,* Sept. 2007, pp. 547-550.

[20] M. Z. Straayer, M. H. Perrott, "A Multi-Path Gated Ring Oscillator TDC With First-Order Noise Shaping", *IEEE J. Solid-State Circuits*, vol. 44, April pp. 1089-1098, 2009.

2-IDT Double-Mode SAW Filters with Decreased Insertion Loss and Advanced Functionalities

Doberstein S. A.,

Omsk Research Institute of Communications & Electronics (ONIIP), Omsk, Russia
e-mail: info@oniip.ru

Abstract— **This paper presents new 2-IDT double-mode SAW (DMS) filters with the decreased insertion loss and advanced functionalities: low input and output impedances, balanced operation and impedance conversion on 42º YX $LiTaO_3$, 64º YX and 41º YX $LiNbO_3$. To decrease the input and output impedances and resistive losses of the IDT electrodes the parallel connection of the 2-IDT DMS filters in the different acoustic tracks is used simultaneously reducing the aperture of each track. For an impedance conversion a series connection of the output IDTs in the 2 parallel acoustic tracks of the filters is used. To increase the selectivity of the filters the cascaded connection and phase weighting are employed. The constructional and topological optimization of the SAW filters is provided with a computer simulation using an equivalent circuit model. The 240-367 MHz samples of the balanced SAW filters have shown an insertion loss of about 1 dB, stopband attenuation of about 20 dB, impedance conversion from 50-75 Ω at the input to 200-300 Ω at the output in a wide fractional bandwidth of 2-6%. Two cascaded filter has provided an insertion loss of less than 2 dB and stopband attenuation of about 50 dB. The filters are housed in the 5x5x1.35 mm SMD packages.**

I. INTRODUCTION

The 3-IDT DMS filters become major as the RF filters for the telecommunication equipment [1]. In spite of the simpler design the 2-IDT DMS filters have not found wide application because they have the high input and output impedances of 130-200 Ω, large insertion loss and require the complex topologies for the impedance conversion [2-4]. This paper presents new 2-IDT DMS filters with the decreased insertion loss and advanced functionalities: low input and output impedances, balanced operation and impedance conversion on 42º YX $LiTaO_3$, 64º YX and 41º YX $LiNbO_3$. To decrease the input and output impedances and resistive losses of the IDT electrodes the parallel connection of the 2-IDT DMS filters in the different acoustic tracks is used simultaneously reducing the aperture of each track (Fig.1) [5]. The balanced operation of the filters is provided by symmetrical connection of the parallel input and output IDTs to the loads. In order for a low input impedance of the filters to be converted to a high output impedance a series connection of the output IDTs in the 2 parallel acoustic tracks of the filters is used (Fig.2). The strongly different impedances at the input and output are achieved with the use of this method of the impedance conversion without applying the complex topologies [4] and without disturbing the balanced operation at the output of the filter. In this case the natural symmetry of the output IDTs and the geometrical symmetry of the bus-bars and contact pads are used.

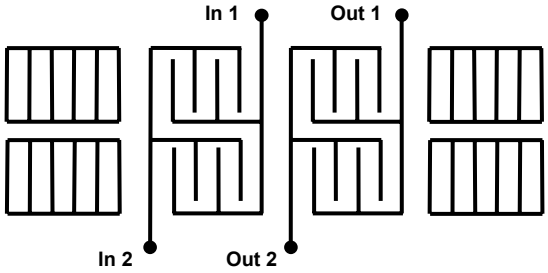


Figure 1. Balanced 2-IDT DMS filter with a parallel connection of the input and output IDTs

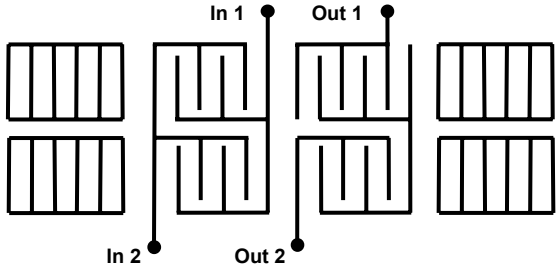


Figure 2. Balanced 2-IDT DMS filter with a parallel connection of the input IDTs and a series connection of the output IDTs

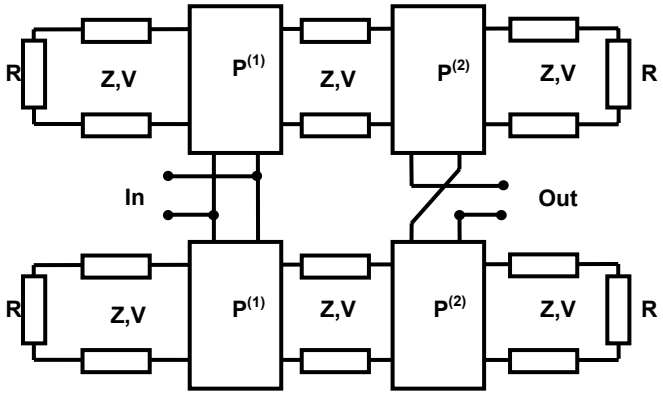


Figure 3. Equivalent circuit of the 2-IDT DMS filter with impedance conversion

II. Simulated and Experemental Results.

A theoretical analysis of the balanced 2-IDT DMS filter was provided using a computer simulation on a basis of an equivalent circuit model [4]. The equivalent circuit of the balanced 2-IDT DMS filter with a parallel connection of the input ITDs and a series connection of the output IDTs is shown on Fig. 3. Here $P^{(1)}$ is a mixed matrix of the input IDT, $P^{(2)}$ is a mixed matrix of output IDT, Z is characteristic impedance of the medium between IDTs and reflectors, V is a SAW velocity, R is an equivalent impedance for a reflector. Fig. 4 shows a simulated frequency response of the balanced 2-IDT DMS filter on 42° YX $LiTaO_3$ in a 50-Ω –200-Ω system. The filter has shown a 3-dB fractional bandwidth of 2% with minimal insertion loss and ripple, stopband attenuation around 20 dB at ±5% offset from a center frequency. The measured frequency response of the balanced 2-IDT DMS filter is shown in Fig. 5. The filter has been symmetrically connected to a 50-Ω input load and to a 200-Ω output load. Measurements were carried out by the network analyzer and balanced transformers.

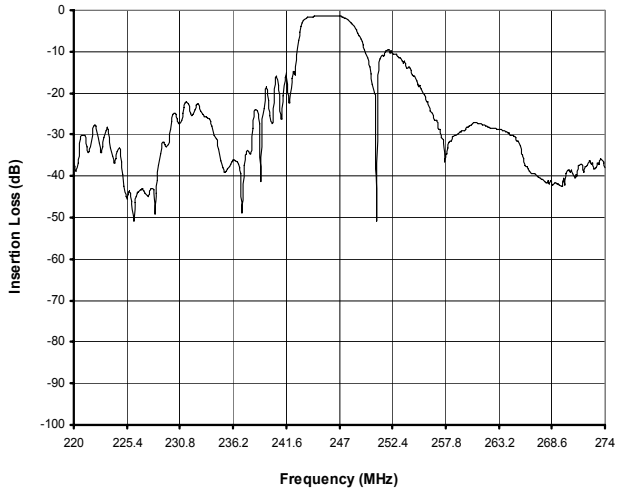


Figure 4. Simulated frequency response of the 2-IDT DMS filter on 42˚ YX $LiTaO_3$

The losses of the balanced transformers were eliminated from the measured insertion loss of the filter. At a center frequency of 246.7 MHz the filter has shown an insertion loss of about 1 dB, 3-dB fractional bandwidth of 5.9 MHz with a low ripple of 0.5 dB and stopband attenuation above 20 dB at ± 12 MHz offset from the center frequency. Fig. 4 and Fig. 5 show good agreement between the simulated and measured responses. A measured input impedance characteristic of the filter is shown in Fig. 6. As will be seen from Fig. 6 at the center frequency of 246.7 MHz the input impedance of the filter is close to real value of 50 Ω. The measured output impedance characteristic of the filter is shown in Fig. 7. As will be seen from Fig. 7 at the center frequency of 246.7 MHz the output impedance of the filter is close to real value of 200 Ω. Thus the impedance conversion from 50 Ω to 200 Ω was obtained experimentally in the 2-IDT DMS filter with a parallel connection of the input ITDs and a series connection of the output IDTs (Fig. 2). For the filter without an impedance conversion (with parallel connection of the input and output IDTs (Fig.1)) a very low insertion loss of less than 1dB was provided in a balanced 50-Ω system. It can be shown that the filters on 64° YX and 41° YX $LiNbO_3$ provide the similar frequency responses with 4% and 6% fractional bandwidth respectively and impedance conversion from 75 Ω to 300 Ω. The measured frequency response of the filter on 64° YX $LiNbO_3$ is shown in Fig. 8. At a center frequency of 367.14 MHz the filter has shown an insertion loss of about 1dB, 3-dB bandwidth of 14.3 MHz, stopband attenuation over 20 dB at ±35 MHz offset from the center frequency in a balanced 75-Ω – 300-Ω system. Fig. 9 shows the measured frequency response of the filter on 41° YX $LiNbO_3$. At a center frequency of 276 MHz the filter has shown an insertion loss of about 1 dB, 3-dB bandwidth of 16.7 MHz, stopband attenuation over 20 dB at ± 33 MHz offset from the center frequency in a balanced 75-Ω – 300-Ω system. To increase the stopband attenuation we can use the phase weighting and cascading two filters. The measured frequency response of the filter on 42° YX $LiTaO_3$ with phase weighting parallel-connected input ITDs is shown in Fig.10.

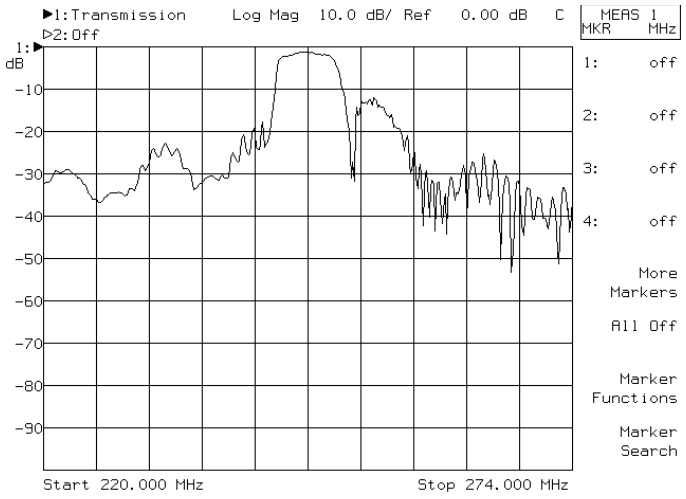


Figure 5. Measured frequency response of the 2-IDT DMS filter on 42˚ YX $LiTaO_3$

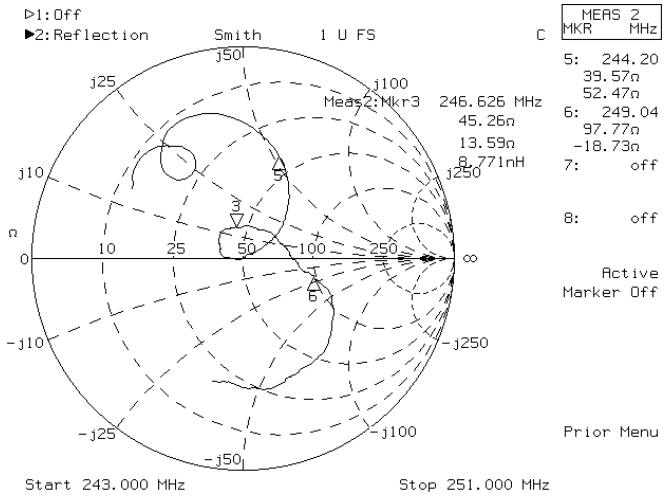


Figure 6. Measured input impedance characteristic of the 2-IDT DMS filter on 42˚ YX $LiTaO_3$

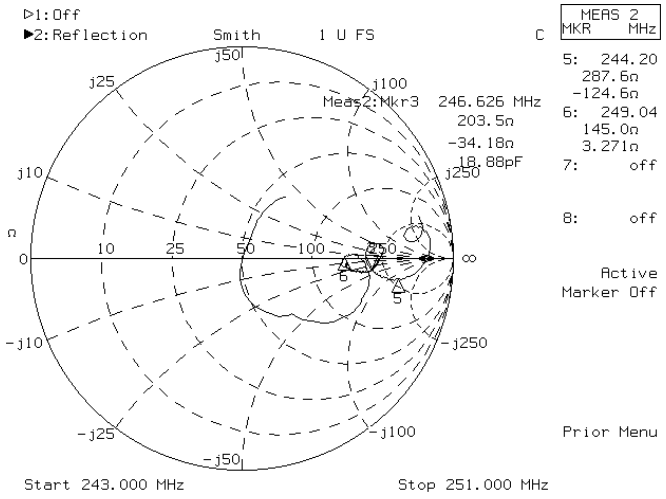


Figure 7. Measured output impedance characteristic of the 2-IDT DMS filter on 42˚ YX $LiTaO_3$

At a center frequency of 276 MHz the filter has shown an insertion loss of about 1 dB, 2-dB bandwidth of 3.3 MHz, stopband attenuation over 40 dB at ±11 MHz offset from the center frequency in a balanced 50-Ω – 200-Ω system. Fig.11. shows the frequency response of the cascade of two balanced filters on 42° YX $LiTaO_3$. The first filter is built without the impedance conversion (with parallel connection of the input and output IDTs). The second filter is built with the impedance conversion (with a series connection of the output IDTs). In a 50-Ω – 200-Ω balanced system the cascade filter has shown an insertion loss of less than 2 dB, 2-dB bandwidth of 4.3 MHz, stopband attenuation of about 50 dB at ±11 MHz offset from the center frequency.

III. CONCLUSION

Our new 2-IDT DMS filters have decreased insertion loss and wider functionalities than their much used prototypes [2-4] and they will be widely employed in the telecommunication equipment. For example these filters offer a single-chip solution for the front-end of the high frequency handheld transceivers for matching the low impedance antenna with modern high impedance double balanced mixer and for suppression of the local oscillator and image frequencies. The filters are housed in the 5x5x1.35 mm SMD packages.

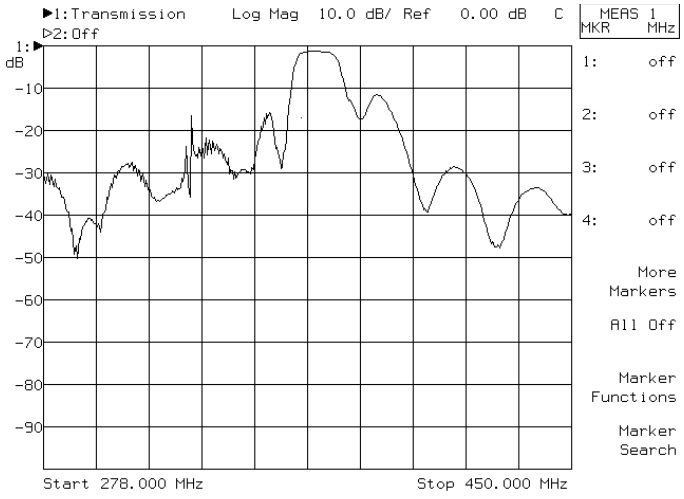


Figure 8. Measured frequency response of the 2-IDT DMS filter on 64˚ YX $LiNbO_3$

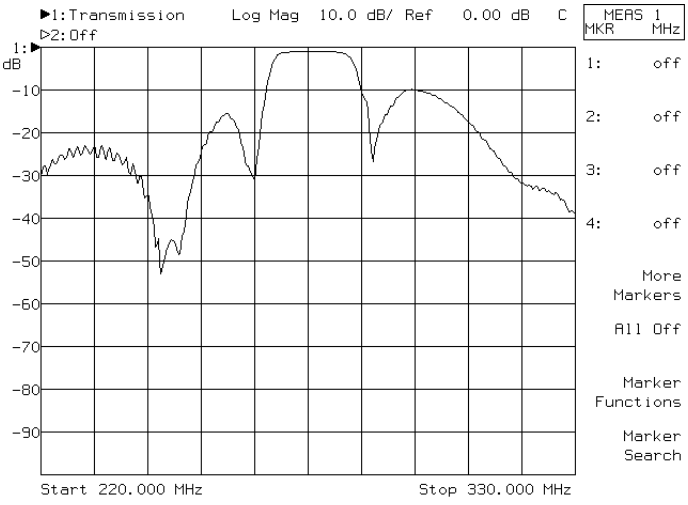


Figure 9. Measured frequency response of the 2-IDT DMS filter on 41˚ YX $LiNbO_3$

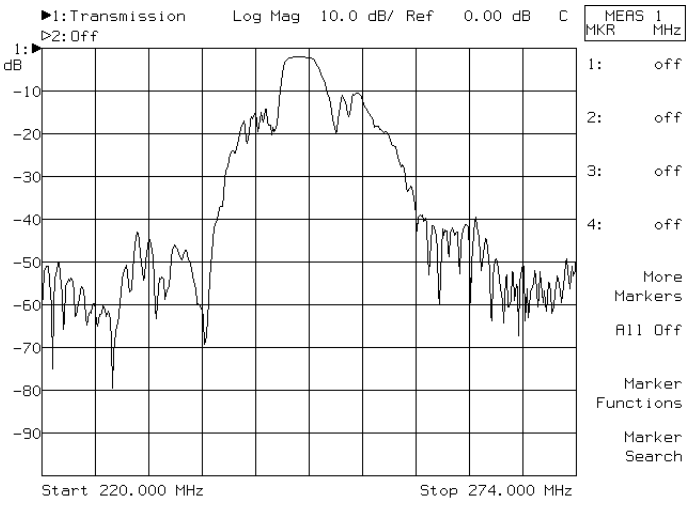


Figure 10. Measured frequency response of the 2-IDT DMS filter with phase weighting on 42˚ YX $LiTaO_3$

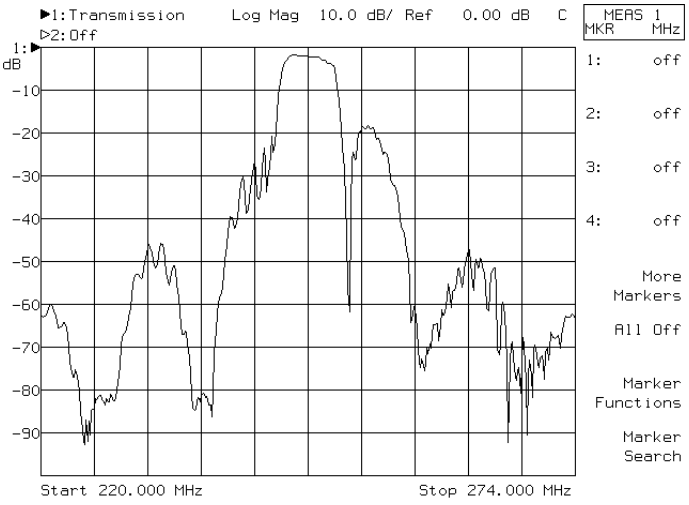


Figure 11. Measured frequency response of the cascaded 2-IDT DMS filter on 42˚ YX $LiTaO_3$

ACKNOWLEDGMENT

Author thanks K. Nikolaenko and N. Ryabova for photomask manufacture, O. Luchina for manufacture of SAW filters, I. Goncharov for editing manuscript.

REFERENCES

[1] O. Kawachi , S. Mitobe, M. Tajima, S. Inoue and K. Hashimoto, “Low-loss and wide-band double-mode surface acoustic wave filters using pitch-modulated interdigital transducers and reflectors”, IEEE Trans. on UFFC, vol. 54, № 10, pp. 2159-2164, 2007.

[2] T. Morita, Y. Watanabe and Y. Nakazawa, “280 MHz wideband low-loss filters” , IEICE of Japan , 1990 Autumn Nat. Con. Record , pp.1-263-264, 1990

[3] T. Morita, Y. Watanabe, M. Tanaka and Y. Nakazawa., “Wideband low loss double mode SAW filters”, Proc. IEEE Ultrason. Symp., pp. 95-104, 1992.

[4] S. A. Doberstein, “Balanced low-loss longitudinally-coupled double-mode resonator SAW filters with impedance conversion”, Proc. IEEE Freq. Control Symp., pp. 199-203, 2008.

[5] J.Meltaus, V.P.Plessky, A.Gortchakov, S.Harma and M.M.Salomaa, “SAW filter based on parallel-connected CRFs with offset frequencies”, Proc. IEEE Ultrason. Symp., pp. 2073-2076, 2003.

Determination and Experimental Verification of High-Temperature SAW Orientations on Langatate

Peter M. Davulis and Mauricio Pereira da Cunha
Dept. of Electrical and Computer Engineering and Laboratory for Surface Science and Technology,
University of Maine, Orono, ME, 04469, U.S.A.
mdacunha@eece.maine.edu

***Abstract*— Langatate (LGT) is a member of the langasite family of crystals appropriate for high-temperature frequency control and sensing applications. This paper identifies multiple LGT SAW orientations for use at high temperature, specifically in the 400 to 900°C range. Orientations with low sensitivity to temperature are desired for frequency control devices and many sensors, conversely large temperature sensitivity is a benefit for temperature sensors. The LGT SAW temperature behavior has been calculated along orientation searches, sweeping the Euler angles (0°, Θ, Ψ), (90°, Θ, Ψ), and (Φ, 90°, Ψ). The temperature coefficient of delay (TCD) and total frequency change over the temperature range were analyzed from 400 to 900°C. Multiple SAW orientations were found with zero-TCD between 400 and 500°C, but no orientations were identified with turn-over temperatures above 500°C. Other orientations have a low variation, −0.8%, in frequency over the range from 400 to 800°C. Additionally, temperature-sensitive orientations were identified with TCD up to 75 ppm/°C at 900°C, for use as temperature sensors at high temperature. The predictions are shown to agree with measured behavior of LGT SAW delay lines fabricated along 6 orientations in the (90°, 23°, Ψ) plane.**

I. INTRODUCTION

There is growing need for high-temperature sensors and frequency control devices in oil and gas extraction, metal and ceramic manufacturing, aerospace, and energy generation industries [1]-[4]. Surface acoustic wave (SAW) devices can meet the demands for high-temperature survivability and wireless multi-sensor operation using substrates such as gallium orthophosphate ($GaPO_4$) and the langasite family of crystals, including langasite ($La_3Ga_5SiO_{14}$, LGS) and langatate ($La_3Ga_{5.5}Ta_{0.5}O_{14}$, LGT) [5]-[16]. Langatate, in particular, has a number of attractive properties for high-temperature operation: i) no phase change until the melting point at 1470°C [17], [18], ii) piezoelectric constants 2-4 times those of quartz [18], iii) higher resistivity than LGS [17], iv) demonstrated temperature-compensated SAW orientations both around room temperature and at elevated temperatures [18]-[22]. However, so far there has been very limited published on high-temperature orientation searches for LGT. In [9], [10] LGS SAW properties were calculated up to 1000°C, but the constants and temperature coefficients employed were taken from [23] and those were measured around at room temperature. Several modern high-temperature frequency control and sensing applications require operation beyond 400°C. For example, sensors are desired for high-temperature fuel cells, which operate in the 600-1000°C range [24], and for gas turbine engines in the compressor, combustor, and exhaust engine sections, which can typically reach temperatures around 450-750°C, 900-1500°C, and 650°C, respectively [24], [25].

In this work, LGT SAW propagation properties are investigated from 400 to 900°C utilizing constants and temperature coefficients extracted at high-temperature [21]. The SAW properties are calculated according to three different orientation sweeps. The search results are used to identify orientations with either low or high sensitivity to temperature. The orientations are investigated along the Euler angles (0°, Θ, Ψ), (90°, Θ, Ψ), and (Φ, 90°, Ψ), in which 2 Euler angles are swept. Multiple zero-TCD orientations are identified between 400 and 500°C. SAW orientations with a small total frequency shift, −0.8%, between 400 and 800°C were also identified, which are anticipated to find use in sensor and frequency-control devices. Orientations with large temperature sensitivity above 400°C are uncovered and those can be utilized as temperature sensors.

The LGT material constants used and the SAW property identification targets are described in Section II. The calculation results of the SAW orientation sweeps are presented in Section III and analyzed to identify high-temperature orientations with either low or high sensitivity to temperature along different temperature ranges. Section IV details high-temperature SAW measurements that provide experimental verification of predicted LGT SAW properties. Section V concludes the paper.

II. MATERIAL CONSTANTS AND SAW PROPERTIES CALCULATIONS

In this work the LGT stiffened elastic constants from [20], [21] and the density and thermal expansion from [26] are used for the high-temperature LGT SAW predictions. The SAW electromechanical coupling (K^2) predictions come from [27], [28], since the high-temperature stiffened elastic constants from [20], [21] do not separately consider the piezoelectric effects. Therefore in this paper all the reported values of K^2 refer to calculations at 25°C. In [21] the authors observed that

Funding was provided in part by DEPSCoR grant # FA9950-07-1-0519, Petroleum Research Fund Grant ACS PRF# 42747-AC10, and Air Force Office AFO Grant # FA8650-05-D-5807.

all the fabricated SAW devices at orientations where $K^2 > 0.2\%$ at room temperature operate up to at least 900°C.

The SAW properties are calculated using the matrix method implemented with MATLAB (The MathWorks, Natick, MA, USA) [19] to sweep across three regions of Euler angles: Region A = (0°, Θ, Ψ), Region B = (90°, Θ, Ψ), and Region C = (Φ, 90°, Ψ). The LGT SAW velocity, frequency change, and TCD were calculated at 400, 500, 600, 700, 800, and 900°C over those three regions by varying two Euler angles, as indicated. These sweeps were selected because they represent cut planes with surface normals in the Y-Z, X-Z, and X-Y planes for easier wafer cutting and X-ray alignment. For the (Φ, 90°, Ψ) sweeps, Ψ was searched only from 0 to 40° because in this region $K^2 \geq 0.1\%$. Only orientations with $K^2 > 0.1\%$ at 25°C are discussed for TCD and frequency change in this work.

The LGT SAW propagation properties analysis at high temperature aimed at identifying potential orientations for three different conditions: (i) narrow temperature range of operation; (ii) large temperature range of operation (400 to 800°C); and (iii) operation of several orientations in a same plane (variation of third Euler angle, Ψ). First, SAW orientations for a narrow operation temperature range were searched for low and high TCD. Next, change in SAW frequency from 400 to 800°C is used to identify orientations for a wider range of operation temperatures. Finally, SAW cuts are investigated for multiple promising orientations in the same plane. Multiple devices on the same wafer with different orientations are useful for sensor applications [29]-[31].

III. Langatate SAW Properties at High Temperature

A. SAW TCD at High Temperature

The calculated LGT SAW TCD for Region A, the Euler angles (0°, Θ, Ψ), at 400°C, 600°C, and 900°C are displayed as contour plots in Fig. 1(a), 1(b), and 1(c), respectively, and are representative of the data calculated at the 6 temperatures (400, 500, 600, 700, 800, and 900°C) mentioned in Section II. The LGT SAW TCD calculated at 400°C for the Regions B, (90°, Θ, Ψ), and C (Φ, 90°, Ψ), are shown in Fig. 2 and Fig. 3, respectively.

For the three regions swept, A, B, and C, orientations with TCD=0 were only found at 400°C and none were found at 500°C or higher; however, multiple orientations have negative TCD at 400°C indicating the existence of turn-over temperatures between 400 and 500°C. In particular, LGT SAW orientations with Euler angles (90°, 40→70°, 60→75°) have $0.1 \leq K^2 \leq 0.2\%$ at room temperature, TCD ≈ −10 ppm/°C at 400°C and TCD ≈ 5 ppm/°C at 500°C. Therefore these orientations have TCD=0 between 400 and 500°C, the highest turnover temperatures for LGT identified in this work for orientations with $K^2 \geq 0.1\%$.

Multiple LGT SAW orientations were found that have close to zero or zero TCD at 400°C and $K^2 \geq 0.1\%$; their Euler angles are listed in Table I along with the respective TCD and K^2.

There are also regions with higher TCD that can be useful for temperature sensing. Specifically, (0°,135→150°, 20→25°) has TCD ≈ 25ppm/°C at 400°C and $K^2 \approx 0.7\%$. Orientations in the region (90°, 20→40°, 110→130°) have TCD between 15 and 35 ppm/°C at 400°C and $0.3\% \leq K^2 \leq 0.7\%$. At 900°C all the calculated orientations were found to have TCD between 40 and 75 ppm/°C.

B. SAW Frequency Change from 400 to 800°C

The LGT SAW behavior for a large temperature span was analyzed by calculating the total SAW frequency change from 400 to 800°C, $\Delta f/f_{400} = (f_{800°C} - f_{400°C})/ f_{400°C}$, shown in Fig. 4 for the Regions A, B, and C.

As can be seen from Fig. 4, the LGT SAW orientations in the range (90°, 55→75°, 55→80°) have the lowest temperature sensitivity from 400 to 800°C, $\Delta f/f_{400} = -0.6\%$, among the calculated orientations with $K^2 \geq 0.1\%$. The SAW orientations (0°, 40→90°, 40→50°) and (90°, 25→80°, 55→65°) have $\Delta f/f_{400} \approx -0.8\%$ with $0.1\% \leq K^2 \leq 0.3\%$. The moderate frequency change is partly due to the fact that these orientations have turnover temperatures between 400 and 500°C.

Conversely, other LGT orientations have larger temperature dependencies. Orientations in the ranges (0°, 10→30°, 65→90°), (0°, 110→160°, 15→45°) and (90°, 15→40°, 95→140°) have $\Delta f/f_{400}$ between −1.2% and −1.8% with $0.2\% \leq K^2 \leq 0.7\%$. These orientations have high TCD through out the temperature range: 15 to 40 ppm/°C at 400°C and 40 to 70 ppm/°C at 800°C.

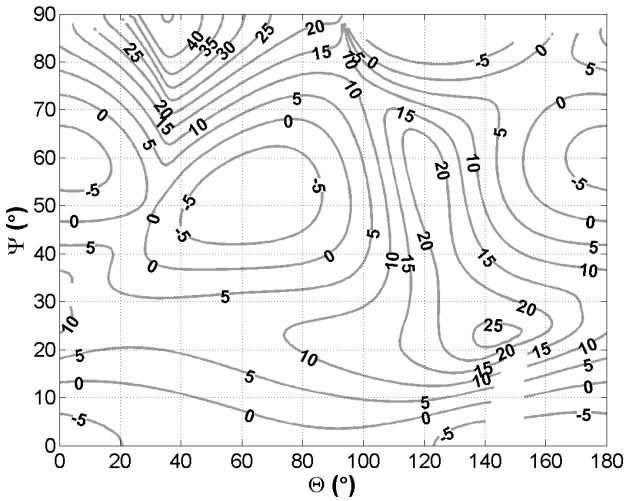
C. SAW Plane Cuts for High-Temperature Sensor Suites

Multiple regions of LGT SAW planes have been identified as desirable for fabricating a suite of devices on the same wafer with $K^2 \geq 0.2\%$ and a variety of temperature sensitivities, useful for sensor applications. LGT planes in the range (0°, 120→140°, Ψ) contain SAW orientations with both low and high temperature-sensitivities, $\Delta f/f_{400}$ ranging from −0.8% to −1.4%, and K^2 ranging from 0.2% to 0.7%.

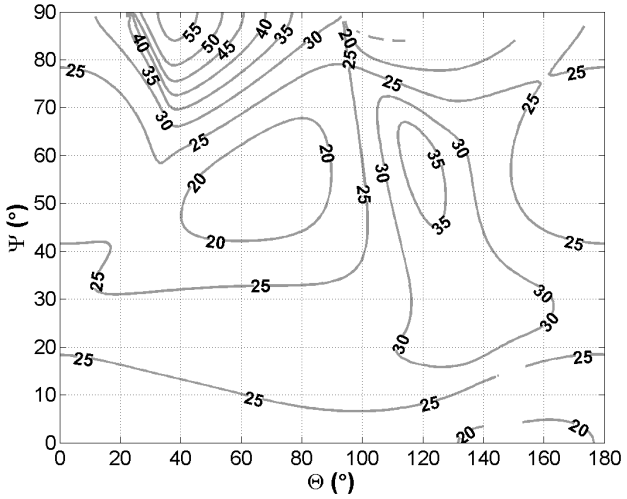
LGT planes in the range (90°, 20→40°, Ψ) have regions in which there are SAW orientations with small temperature sensitivity, $\Delta f/f_{400} = -0.8\%$, and others with large temperature sensitivity, $\Delta f/f_{400} = -1.8\%$, and K^2 varying between 0.2% and 0.7%. Additionally, there are a number of orientations in this range that have $\Delta f/f_{400} = -1.0\%$ and that are widely spaced over the plane, with potential for multiple measurand sensor applications.

IV. Experimental Verification and Discussion

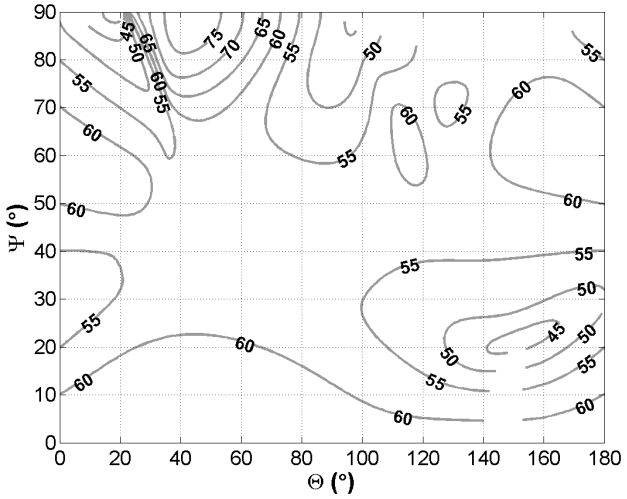
The SAW high-temperature properties calculated in this work were compared to those of LGT SAW devices reported in [21] for experimental validation of the frequency and TCD predictions up to 900°C. Six SAW delay lines were fabricated along orientations (90°, 23°, Ψ) with Ψ = 0, 13, 48, 77, 119, and 123°. The interdigital transducers had periodicity given by wavelength λ = 32 μm and utilized 110-nm-thick Pt/Rh/ZrO_2 electrodes [13] to enable stable high-temperature operation above 900°C.



(a)



(b)



(c)

Figure 1. Contour plots of predicted LGT SAW TCD in ppm/°C for Region A, Euler angles (0°, Θ, Ψ): (a) at 400°C, (b) at 600°C, (c) at 900°C.

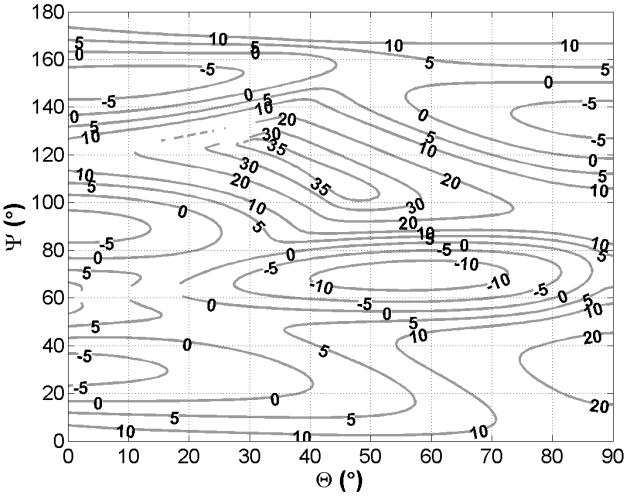


Figure 2. Contour plot of predicted LGT SAW TCD in ppm/°C for Region B, Euler angles (90°, Θ, Ψ), at 400°C.

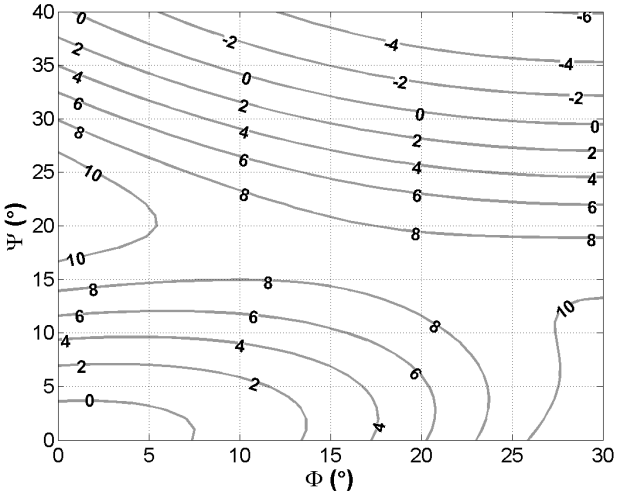
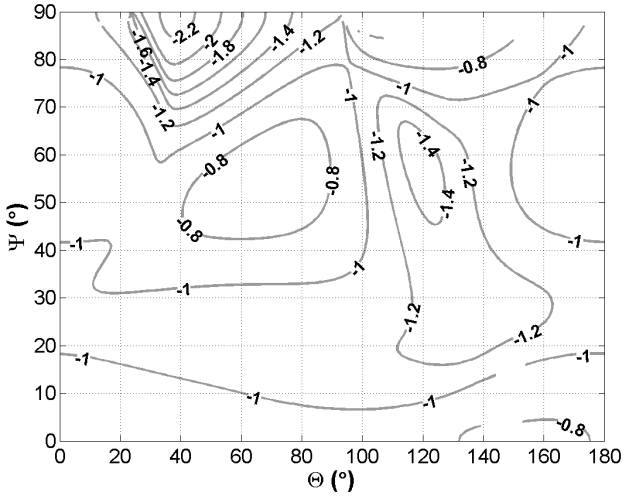


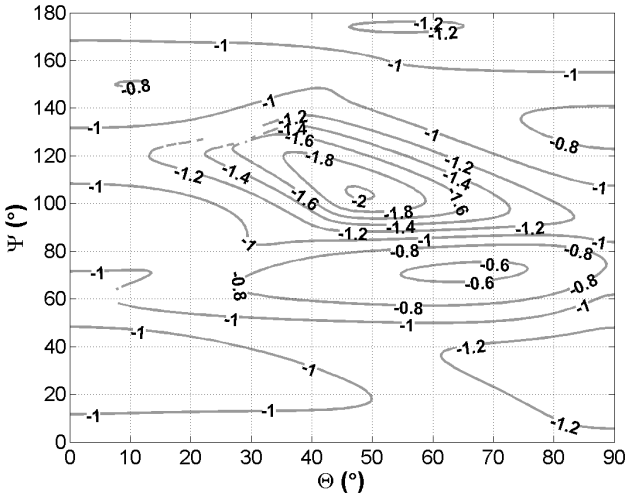
Figure 3. Contour plot of predicted LGT SAW TCD in ppm/°C for Region C, Euler angles (Φ, 90°, Ψ), at 400°C.

TABLE I. LGT SAW ORIENTATIONS WITH TCD ≈ 0 AT 400°C

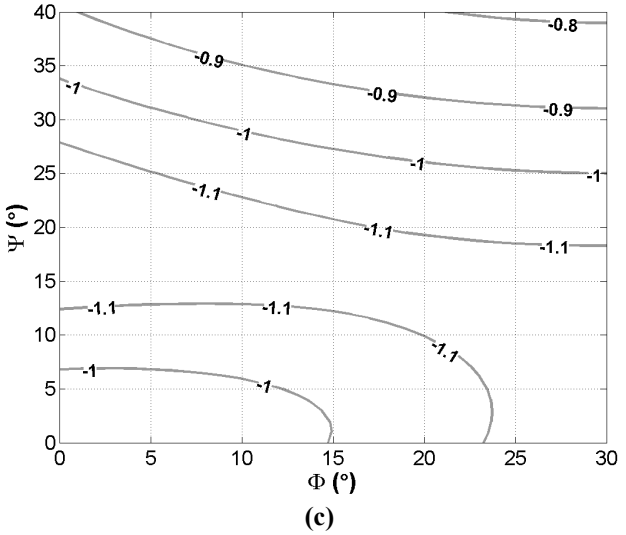
Euler Angle Sweep (°)	Orientation Euler Angles (°)	TCD at 400°C (ppm/°C)	K^2 at 25°C (%)
(0°, Θ, Ψ)	(0, 20→90, 30→50)	−5 to 5	0.1 to 0.4
Region A	(0, 30→130, 0→15)	−5 to 5	0.1 to 0.5
	(0, 110→180, 0→90)	−5 to 5	0.1 to 0.4
(90°, Θ, Ψ)	(90, 10→90, 40→80)	−10 to 5	0.1 to 0.4
Region B	(90, 0→20, 100→105)	≈0	0.1 to 0.2
	(90, 0→20, 35→140)	≈0	0.1 to 0.2
	(90, 0→40, 10→20)	≈0	0.1 to 0.2
	(90, 0→40, 155→170)	≈0	0.1 to 0.2
(Φ, 90°, Ψ)	(0→8, 90, 0→5)	≈0	0.5
Region C	(0→30, 90, 28→40)	≈0	0.1 to 0.2



(a)



(b)



(c)

Figure 4. Contour plots of predicted LGT SAW frequency change (percent) from 400 to 800°C, $\Delta f/f_{400} = (f_{800°C} - f_{400°C})/ f_{400°C}$, for orientation sweep (a) Region A, Euler angles (0°, Θ, Ψ), (b) Region B, Euler angles (90°, Θ, Ψ), and (c) Region C, Euler angles (Φ, 90°, Ψ).

The calculated and measured SAW frequencies responses are compared in Fig. 5(a) and 5(b) at 400 and 900°C, respectively, for the six orientations identified in the previous paragraph. Along the 6 fabricated orientations the measured and predicted operation frequencies agree within experimental and prediction uncertainties in the temperature range of 400 to 900°C. The calculated SAW frequencies are within 0.85% of the measured frequencies at 400°C and within 1.6% at 900°C.

The predicted and measured TCDs on the (90°, 23°, Ψ) plane are compared in Fig. 6 at 400°C. The predicted and measured TCD agree to within the experimental and prediction uncertainty for the 6 tested SAW orientations throughout the temperature range between 400 and 900°C. The discrepancy between the predicted and measured TCD is below 4.7 ppm/°C at 400°C and less than 24.4 ppm/°C at 900°C.

The agreement between the calculated and measured SAW frequency and TCD for the 6 orientations fabricated along the (90°, 23°, Ψ) plane validates the LGT high-temperature constants and coefficients for SAW predictions at high-temperature.

V. Conclusions

This work has investigated the high-temperature SAW phase velocity, frequency change, and TCD properties behavior of LGT up to 900°C for three regions in space, namely, (0°, Θ, Ψ), (90°, Θ, Ψ), and (Φ, 90°, Ψ). New LGT SAW orientations have been identified with zero TCD between 400 and 500°C and with low TCD up to 900°C with potential applications for frequency control, sensing, and signal processing. Alternatively, temperature-sensitive orientations have been found of interest to temperature sensing.

Orientations have been identified for high-temperature applications with: i) a narrow temperature range of operation; ii) large temperature range of operation (400 to 800°C); and iii) multiple orientations in a same plane for temperature measurement and separation of the thermal effect from other measurands, as appropriate in sensor applications.

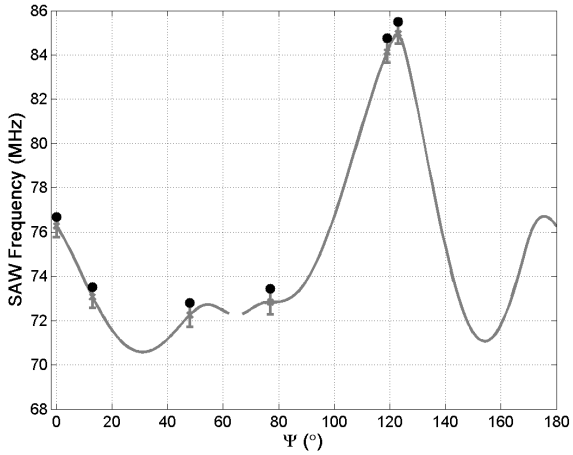
The agreement between the calculated and measured SAW frequency and TCD responses reported in the work validates the LGT constants and temperature coefficients for SAW predictions up to 900°C.

Acknowledgment

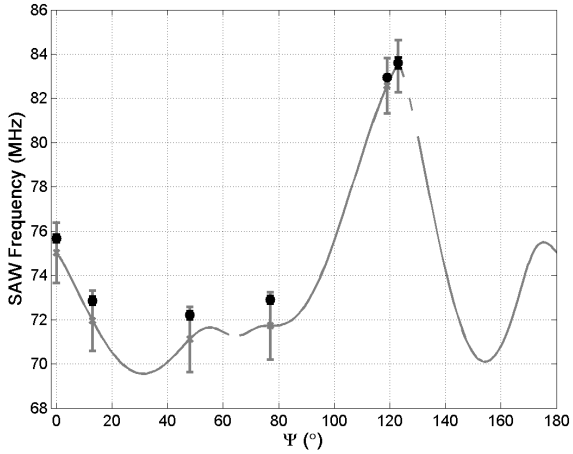
The authors would like express their gratitude to the personnel of the Laboratory of Surface Science and Technology (LASST) and Department of Electrical and Computer Engineering at the University of Maine for valuable technical discussion and for assisting with the equipment used in this work and with the fabrication of the devices.

References

[1] J. Hornsteiner, E. Born, G. Fischerauer, and E. Riha, "Surface acoustic wave sensors for high-temperature applications," *Proc. 1998 IEEE Int'l Freq. Cntrl. Symp.*, pp. 615-620.



(a)



(b)

Figure 5. Measured and predicted LGT SAW frequency for devices with λ = 32 μm: (a) along plane (90°, 23°, Ψ) at 400°C, (b) along plane (90°, 23°, Ψ) at 900°C. Gray line: calculated frequencies, with error bars at the at angles of fabricated devices; black circle: measured frequencies with error bars for Ψ = 0, 13, 48, 77, 119, 123°.

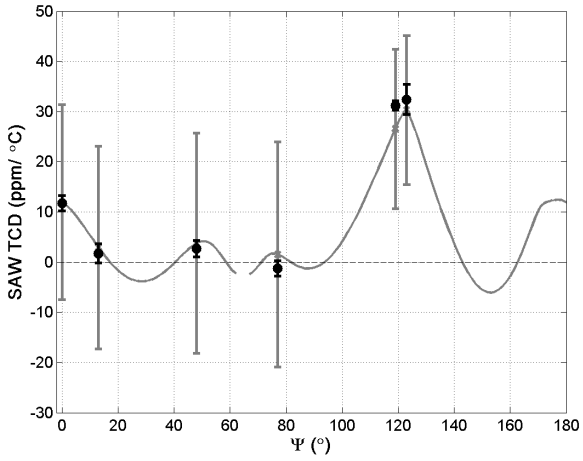


Figure 6. LGT SAW TCD along plane (90°, 23°, Ψ) at 400°C. Gray line: calculated TCD, with error bars at the at angles of fabricated devices; black circle: measured TCD with error bars for Ψ = 0, 13, 48, 77, 119, 123°.

[2] P. Krempl, G. Schleinzer, W. Wallnöfer, "Gallium phosphate, $GaPO_4$: a new piezoelectric crystal material for high-temperature sensorics," *Sens. Actuators A*, A61 (1-3), 1997, pp. 361-363.

[3] J. Détaint, A. Zarka, B. Capelle, D. Palmier, and E. Philippot, "Optimization of the design of the resonators using the new materials: application to gallium phosphate and langasite," *Proc. 1997 IEEE Int'l Freq. Cntrl. Symp.*, pp. 566-578.

[4] J. Thiele and M. Pereira da Cunha, "Platinum and palladium high-temperature transducers on langasite," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont*., vol. 52, no. 4, pp. 545-549, April 2005.

[5] J. Hornsteiner, E. Born, and E. Riha, "Langasite for high temperature surface acoustic wave applications," *Phys. Stat. Sol. A,* vol. 163, pp. R3-R4, 1997.

[6] J. W. Mrosk, L. Berger, C. Ettl, H.-J. Fecht, G. Fischerauer, and A. Dommann, "Materials issues of SAW sensors for high-temperature applications," *IEEE Trans. Ind. Elec.*, vol. 48, no. 2, Apr. 2001.

[7] S.-Q. Wang, J. Harada, and S. Uda, "A wireless surface acoustic wave temperature sensor using langasite as substrate material for high-temperature applications," *Jpn. J. Appl. Phys.*, vol. 42, part 1, no. 9B, pp. 6124-6127, Sept. 2003.

[8] R. Fachberger, G. Bruckner, G. Knoll, R. Hauser, J. Biniasch, and L. Reindl, "Applicability of $LiNbO_3$, langasite, and $GaPO_4$ in high temperature SAW sensors operating at radio frequencies," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 51, no. 11, pp. 1427-1431, Nov 2004.

[9] X. J. Ji, T. Han, W. K. Shi, and G. W. Zhang, "Investigation on SAW properties of LGS and optimal cuts for high-temperature applications," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 52, no. 11, pp 2075 -2080, Nov. 2005.

[10] T. Han, X. J. Ji, and W. K. Shi, "Optimal pressure-sensitive cuts for surface acoustic waves on langasite," *Science in China Series F: Information Sciences*, vol. 49, no. 2, pp 254-261, 2006

[11] E. Ansorge, S. Schimpf, S. Hirsch, J. Sauerwald, H. Fritze, and B. Schmidt, "Evaluation of langasite ($La_3Ga_5SiO_{14}$) as a material for high temperature microsystems," *Sens. and Actuators A.*, vol. 130-131, pp. 393-396, Aug. 2006.

[12] M. N. Hamidon, V. Skarda, N. M. White, F. Krispel, P. Krempl, M. Binhack, and W. Buff, "High-temperature 434 MHz surface acoustic wave devices based on $GaPO_4$," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 53, no. 12, pp. 2465-2470, December 2006.

[13] M. Pereira da Cunha, R. J. Lad, T. Moonlight, G. Bernhardt, and D. J. Frankel, "High temperature stability of langasite surface acoustic wave devices," *Proc. 2008 IEEE Int'l Ultrason. Symp.*, pp. 205-208.

[14] J. Bardong, M. Schulz, M. Schmitt, I. Shrena, D. Eisele, E. Mayer, L. M. Reindl, and H. Fritze, "Precise measurements of BAW and SAW properties of langasite in the temperature range from 25°C to 1000°C," *Proc. 2008 IEEE Int'l Freq. Cont. Symp.*, pp. 326-331.

[15] A. Canabal, P. M. Davulis, E. Dudzik, and M. Pereira da Cunha, "CDMA and FSCW surface acoustic wave temperature sensors for wireless operation at high temperatures," *Proc. 2009 IEEE Int'l Ultrason. Symp*., pp. 807-810.

[16] A. Canabal, P. M. Davulis, G. M. Harris, and M. Pereira da Cunha, "High-temperature battery-free wireless microwave acoustic resonator sensor system," *Electronics Letters*, vol. 46, no. 7, pp. 471-472, 1 April 2010.

[17] J. Bohm, R. B. Heimann, M. Hengst, R. Roewer and J. Schindler, "Czochralski growth and characterization of piezoelectric single crystals with langasite structure: $La_3Ga_5SiO_{14}$ (LGS), $La_3Ga_{5.5}Nb_{0.5}O_{14}$ (LGN), $La_3Ga_{5.5}Ta_{0.5}O_{14}$ (LGT) Part I," *J. Cryst. Growth*, vol. 204, pp. 128–136, 1999.

[18] J. A. Kosinski, "New piezoelectric substrates for SAW devices," *Int'l J. of High Speed Electronics and Systems*, vol. 10, no. 4, 2000, pp. 1017-1068.

[19] M. Pereira da Cunha, D. C. Malocha, E. L. Adler, and K. J. Casey, "Surface and pseudo surface acoustic waves in langatate: predictions and measurements," *IEEE Trans. on Ultrason. Ferroelect. Freq. Control,* vol. 49, no. 9, pp 1291-1299, Sept. 2002.

[20] P. M. Davulis, A. Shyam, E. Lara-Curzio, M. Pereira da Cunha, "High temperature elastic constants of langatate from RUS measurements up to 1100°C," *Proc. 2008 IEEE Int'l Ultrason. Symp.*, pp. 2150-2153.

[21] P. Davulis and M. Pereira da Cunha, "High-temperature langatate elastic constants and experimental validation up to 900°C," *IEEE Trans. on Ultrason. Ferroelect. Freq. Control,* vol. 57, no. 1, pp 59-65, Jan. 2010.

[22] B. T. Sturtevant, M. Pereira da Cunha, "Assessment of langatate material constants and temperature coefficients using SAW delay line measurements," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 57, no. 3, pp. 533-539, Mar. 2010.

[23] A. Bungo, C. Jian, K. Yamaguchi, Y. Sawada, R. Kimura, and S. Uda, "Experimental and theoretical analysis of SAW properties of the langasite substrate with Euler angle (0°, 140°, φ)," *Proc. 1999 IEEE Int'l Ultrason. Symp.*, pp. 231-234.

[24] B. T. Chorpening, D. Tucker, and S. M. Maley, "Sensors applications in 21st century fossil-fuel based power generation," *Proc. 2004 IEEE Sensors*, vol. 3, pp. 1153-1156.

[25] W. Cullinane and R. Strange, "Gas turbine engine validation instrumentation: measurements, sensors, and needs," *Proc. of the SPIE*, vol. 3852, pp. 2-13, 1999.

[26] T. R. Beaucage, E. P. Beenfeldt, S. A. Speakman, W. D. Porter, E. A. Payzant, M. Pereira da Cunha, "Comparison of high temperature crystal lattice and bulk thermal expansion measurements of LGT single crystal," *Proc. 2006 IEEE Int'l Freq. Cont. Symp.*, pp. 658-663.

[27] B. T. Sturtevant, P. M. Davulis, and M. Pereira da Cunha, "Pulse echo and combined resonance techniques: a full set of LGT acoustic wave constants and temperature coefficients," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 56, no. 4, pp. 788-797, April 2009.

[28] B. T. Sturtevant, "Ultrasonic characterization of single crystal langatate," Ph.D. diss., University of Maine, 2009.

[29] L. Reindl, G. Scholl, T. Ostertag, H. Scherr, U. Wolff, and F. Schmidt, "Theory and application of passive SAW radio transponders as sensors," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 45, no. 5, pp 1281-1292, Sept. 1998.

[30] W. Buff, S. Klett, M. Rusko, J. Ehrenpfordt, and M. Goroll, "Passive remote sensing for tempearture and pressure using SAW resonator devices," *IEEE Trans. Ultrason., Ferroelect., Freq. Cont.*, vol. 45, no. 5, pp 1388-1392, Sept. 1998.

[31] V. Kalinin, G. Bown, and A. Leigh, "Contactless torque and temperature sensor based on SAW resonators," *Proc. 2006 Int'l Ultrason. Symp.*, pp. 1490-1493.

Temperature Coefficients of Elastic Constants of Trigonal and Tetragonal Crystals

B.P. Sorokin, A.V. Telichko
Technological Institute for Superhard and Novel Carbon Materials
Troitsk, Moscow region, Russia
E-mail: bpsorokin2@rambler.ru

***Abstract*—Temperature coefficients of elastic constants (TC_{ij}) for trigonal and tetragonal crystals have been obtained. The TC_{ij} of a number of trigonal and tetragonal crystals have been calculated and the comparison with known experimental data has been fulfilled. The nature of anomalous TC_{66} of quartz has been discussed.**

I. INTRODUCTION

Normal temperature behavior of elastic constants for most of the crystals is proportionally decreased with temperature increasing and a small dependence (or lack thereof) at low temperatures. Knowledge of these dependences is important for the solid state physics and applications of crystals.

One of most attractive properties of quartz is the temperature stability observed for some crystalline cuts and modes. Owing to combination of thermostability with piezoelectricity and small attenuation of acoustic waves quartz single crystal has great importance as a material for acoustoelectronic devices. There are only a few piezoelectric crystals possessing the same combination of properties. Nature of the temperature stability is closely connected with elastic anharmonicity of specified elastic modes in such crystals. In practice the search of cuts possessing thermostable orientation was commonly based on the data of temperature coefficients of elastic constants obtained from experiment. Unfortunately there is no exact theory which can be used for explanation of anharmonic behavior of crystals. Consequently a possibility to predict a trend of the search for the new thermostable materials and crystalline cuts is absent. On the basis of the theory explaining the small amplitude bulk acoustic waves (BAW) propagation in crystals under the influence of bias finite fields the temperature coefficients (TC_{IJ}) of elastic constants (C_{IJ}) of cubic crystals have been obtained earlier [1]. Satisfactory agreement between calculation and experimental results for a number of cubic crystals has been achieved. The similar method of calculations of frequency dependences on temperature in quartz piezoelectric resonators has been used by authors [2].

The main objective of this paper is to obtain the temperature coefficients for trigonal and tetragonal crystals.

II. EFFECT OF TEMPERATURE ON THE PROPAGATION OF BULK ACOUSTIC WAVES IN CRYSTALS

In the theory of BAW propagation in crystals under the action of the temperature as a bias field both two effects of geometrical and physical nonlinearity of elastic media should be taken into account. As a result of first effect the temperature action leads to anisotropic dilatation of low symmetry crystals, to the distortion of the sample's shape and therefore to the changing of the path of acoustic wave. Secondly the weakening of chemical bonds between atoms under the thermal expansion usually leads to the decreasing of the macroscopic elastic constants of a crystal. Modified Green-Christoffel's equation for elastic waves of small amplitude in a uniformly deformed elastic medium referred to initial (natural) state can be written as [3]:

$$\left[\Gamma_{BC}(\bar{\eta}) - \rho_0 v^2 \delta_{BC}\right]\tilde{U}_B = 0. \tag{1}$$

Here ρ_0 is the density of undisturbed crystal; v is the phase velocity and $\tilde{U}_B$ is the BAW displacement vector. For a mechanically free sample the static deformation arising due to thermal expansion can be described by the relation:

$$\bar{\eta}_{PQ} = \alpha_{PQ}\Delta T, \tag{2}$$

where α_{PQ} is the tensor of linear thermal expansion. Using the form of Green-Christoffel's tensor obtained earlier [3] and the static deformation (2), in this case we can write:

$$\begin{aligned}\Gamma_{BC}(\Delta T) &= \bar{C}_{FC} C^*_{ABFD} N_A N_D = \\ &= \bar{C}_{FC}\left(C^S_{ABFD} + C^S_{ABFDPQ}\alpha_{PQ}\Delta T\right)N_A N_D,\end{aligned} \tag{3}$$

where $\bar{C}_{FC}$ is Green's tensor of finite static deformation; C^*_{ABFD}, C^S_{ABKL} and C^S_{ABKLPQ} are the effective, 2nd and 3rd order elastic constants (SOEC and TOEC) respectively; N_A is

the unit vector of wave propagation direction. Under the thermal expansion of mechanically free crystal Green's tensor takes a form:

$$\overline{C}_{FC} = \delta_{FC} + 2\overline{\eta}_{FC} = \delta_{FC} + 2\alpha_{FC}\Delta T. \quad (4)$$

Substituting (4) into (3) and taking into account only the terms proportional to ΔT a linearized form of Green-Christoffel's tensor can be obtained:

$$\Gamma_{BC}(\Delta T) = [C^{S}_{ABCD} + (2C^{S}_{ABPD}\delta_{CQ} + C^{S}_{ABCDPQ})\alpha_{PQ}\Delta T]N_A N_D. \quad (5)$$

Solution of system (1) with tensor (5) for the chosen elastic modes in the certain directions of the crystals gives us the required temperature coefficients of elastic constants. So the main relations applying to the 32, 3m and $\overline{3}$m trigonal crystals have a form:

$$\begin{aligned}
\frac{dC_{11}}{dT} &= (2C_{11} + C_{111} + C_{112})\alpha_{11} + C_{113}\alpha_{33},\\
\frac{dC_{33}}{dT} &= 2C_{133}\alpha_{11} + (2C_{33} + C_{333})\alpha_{33},\\
\frac{dC_{44}}{dT} &= (2C_{44} + C_{144} + C_{155})\alpha_{11} + C_{344}\alpha_{33},\\
\frac{dC_{66}}{dT} &= \frac{1}{2}[(4C_{66} + C_{222} - C_{112})\alpha_{11} + (C_{113} - C_{123})\alpha_{33}],\\
\frac{dC_{13}}{dT} &= \frac{1}{2}[(C_{113} + C_{123})\alpha_{11} + (2C_{13} + C_{133})\alpha_{33}],\\
\frac{dC_{14}}{dT} &= (2C_{14} + C_{114} + C_{124})\alpha_{11} + C_{134}\alpha_{33}.
\end{aligned} \quad (6)$$

Similar relations for hexagonal (622, 6mm, $\overline{6}$2m, 6/mmm) crystals have a form:

$$\begin{aligned}
\frac{dC_{11}}{dT} &= (2C_{11} + C_{111} + C_{112})\alpha_{11} + C_{113}\alpha_{33},\\
\frac{dC_{33}}{dT} &= 2C_{133}\alpha_{11} + (2C_{33} + C_{333})\alpha_{33},\\
\frac{dC_{44}}{dT} &= (2C_{44} + C_{144} + C_{155})\alpha_{11} + C_{344}\alpha_{33},\\
\frac{dC_{66}}{dT} &= \frac{1}{2}[(4C_{66} + C_{222} - C_{112})\alpha_{11} + (C_{113} - C_{123})\alpha_{33}],\\
\frac{dC_{13}}{dT} &= \frac{1}{2}[(C_{113} + C_{123})\alpha_{11} + (2C_{13} + C_{133})]\alpha_{33}.
\end{aligned} \quad (7)$$

Note that for trigonal and hexagonal crystals there is the relation between elastic constants $C_{66} = \frac{1}{2}(C_{11}-C_{12})$. Temperature coefficients for the tetragonal (422, 4mm, $\overline{4}$2m, 4/mmm) crystals can be written as:

$$\begin{aligned}
\frac{dC_{11}}{dT} &= (2C_{11} + C_{111} + C_{112})\alpha_{11} + C_{113}\alpha_{33},\\
\frac{dC_{33}}{dT} &= 2C_{133}\alpha_{11} + (2C_{33} + C_{333})\alpha_{33},\\
\frac{dC_{44}}{dT} &= (2C_{44} + C_{144} + C_{155})\alpha_{11} + C_{344}\alpha_{33},\\
\frac{dC_{66}}{dT} &= 2(C_{66} + C_{133})\alpha_{11} + C_{366}\alpha_{33},\\
\frac{dC_{12}}{dT} &= (C_{12} + C_{112})\alpha_{11} + \frac{1}{2}C_{123}\alpha_{33},\\
\frac{dC_{13}}{dT} &= \frac{1}{2}[(C_{113} + C_{123})\alpha_{11} + (2C_{13} + C_{133})\alpha_{33}].
\end{aligned} \quad (8)$$

As a result of analysis of (6), (7) and (8) one can conclude that temperature coefficients of elastic constants substantially depend on the anharmonic 3rd order elastic constants and coefficients of linear thermal expansion. Terms formed as the productions of TOEC and coefficients of linear thermal expansion represent the physical nonlinearity; SOEC multiplying by coefficients of linear thermal expansion define the geometrical nonlinearity.

III. Results and Discussion

Using derived relations we have been calculated first order temperature coefficients of trigonal, hexagonal and tetragonal crystals with known SOEC and TOEC and linear thermal expansion. Calculated results have been compared with known experimental data (tabs. I-III). Material constants and experimental temperature coefficients were taken from [4-39].

In the most cases the agreement of signs between the experimental and calculated temperature coefficients has been obtained. A negative sign of TC_{IJ} is associated with normal temperature dependence. Anomalous positive sign of TC_{66} has been calculated for quartz in agreement with experiment. Especially good agreement between experimental and calculated data takes a place for diagonal constants. Note that some crystals (quartz, lithium niobate, magnesium) have shown the most adequate coincidence both in sign and magnitude of TC_{IJ}. On the contrary the most unsatisfactory coincidence is observed for Zn and Er data (tab. II).

The main origin of discrepancies could be connected with mistakes in measuring of SOEC and TOEC and linear thermal expansion coefficients. Commonly such investigations have been executed by different working groups using the samples of various origins. Especially the effect of real defect structure of various crystals should result in the discrepancy of 3rd order elastic constants. On the other side as one can see from the tables there are mistakes in experimentally obtained temperature coefficients.

Therefore the difference between the experimental and calculated temperature coefficients could be used for independent analysis of mistakes' origin.

TABLE I. COMPARISON OF EXPERIMENTAL AND CALCULATED TEMPERATURE COEFFICIENTS FOR TRIGONAL CRYSTALS

dC_{IJ}/dT, 10^6 Pa/K	$LiNbO_3$			SiO_2			$La_3Ga_5SiO_{14}$			Al_2O_3		
	Calc. [4-6]	*Calc.* [4, 5, 7]	*Exper.* [8]	*Calc.* [9-11]	*Exper.* [12]	*Exper.* [13]	*Calc.* [14, 15]	*Exper.* [16]	*Exper.* [17]	*Calc.* [18-20]	*Exper.* [21]	*Exper.* [18]
dC_{11}/dT	-39,0	10,7	-34,6	-4,9	-4,3	-3,8	-3,9	-8,9	-11,1	-26,4	-37,3	-35,8
dC_{33}/dT	-21,1	-0,3	-36,8	-13,0	-16,9	-20,8	-13,8	-24,6	-35,3	-23,7	-42,3	-29,9
dC_{44}/dT	-18,2	-12,0	-12,2	-4,0	-9,9	-10,1	-2,4	-1,6	-4,3	-13,4	-26,5	-24,2
dC_{66}/dT	-12,8	-4,9	-10,3	2,1	7,0	7,5	-2,6	1,5	5,8	-10,6	-9,7	-4,4
dC_{14}/dT	5,1	-5,2	-1,7	-2,7	-1,8	-1,9	-0,6	-5,2	-2,2	-0,2	-1,7	|2,2|

TABLE II. COMPARISON OF EXPERIMENTAL AND CALCULATED TEMPERATURE COEFFICIENTS FOR HEXAGONAL CRYSTALS

dC_{IJ}/dT, 10^6 Pa/K	Zn		Mg		Er	
	Calc. [22-24]	*Exper.* [25]	*Calc.* [23, 26, 27]	*Exper.* [27]	*Calc.* [28-30]	*Exper.* [31]
dC_{11}/dT	-41,8	-79,2	-15,4	-18,0	-6,3	-20,2
dC_{33}/dT	-47,1	-18,5	-18,2	-19,9	-13,4	-11,0
dC_{44}/dT	-23,7	-32,1	-5,8	-8,4	-2,3	-7,7
dC_{66}/dT	-15,7	-	-7,0	-2,0	-2,9	-4,9
dC_{13}/dT	-11,1	-8,0	-1,0	-0,9	-1,6	1,6

TABLE III. COMPARISON OF EXPERIMENTAL AND CALCULATED TEMPERATURE COEFFICIENTS FOR TETRAGONAL CRYSTALS

dC_{IJ}/dT, 10^6 Pa/K	TeO_2			KH_2PO_4		
	Calc. [32-34]	*Exper.* [34]	*Exper.* [35]	*Calc.* [36-38]	*Exper.* [24]	*Exper.* [39]
dC_{11}/dT	-13,9	-15,1	-20,2	-29,5	-47,0	-52,5
dC_{33}/dT	-19,7	-29,5	-31,5	-30,9	-33,3	-36,7
dC_{44}/dT	0,6	-2,0	-3,5	-1,6	-7,7	-3,8
dC_{66}/dT	-11,8	-29,3	-22,0	-1,3	-3,3	-3,7
dC_{12}/dT	-11,3	-16,9	-21,2	-13,0	-9,4	-
dC_{13}/dT	-2,9	-1,5	4,9	-4,2	-5,0	-

Let's analyze the nature of anomalous value of TC_{66} which can be observed in trigonal crystals. It is convenient to rewrite corresponding relation taken from (6) as:

$$\frac{dC_{66}}{dT} = \left[(2C_{66} + C_{166} + C_{266})\alpha_{11} + C_{366}\alpha_{33}\right]. \qquad (9)$$

Such form is obtained due to relations of coupling between dependent and independent components of TOEC tensor of trigonal crystals. Substituting the data on SOEC and TOEC as well as linear thermal expansion coefficients into (9) one can see that first term (geometrical nonlinearity) in the right side of (9) will always have a positive sign, but the other terms could be both positive and negative. Since linear thermal expansion coefficients have positive signs, the signs of the second and others terms are determined by the TOEC signs only. Usually in crystals the signs of the TOEC are negative, and as a rule the TOEC magnitudes exceed the SOEC ones. Analyzing the TOEC values of quartz among the known data one can see an unusual peculiarity: $|C_{112}| > |C_{111}|$. Calculation of the TOEC values useful for the analysis of (9) gives us $C_{166} = -3{,}775 \cdot 10^{10}$, $C_{266} = +3{,}525 \cdot 10^{10}$ and $C_{366} = +14{,}75 \cdot 10^{10}$ (in Pa). Consequently, the sum of first, third and fourth terms in (9) will exceed the negative contribution of the second one, and TC_{66} will be positive.

It is well known that the transformation of generalized Hook's law from linear to polynomial dependence between deformation and stress is defined by TOEC values [40]:

$$\tau_{AB} = C_{ABCD}\eta_{CD} + C_{ABCDPQ}\eta_{CD}\eta_{PQ}, \qquad (10)$$

where τ_{AB} is the tensor of thermodynamical stress. In the most cases when positive deformation increases and TOEC values are negative, the quadratic term in (10) produces a small decreasing of the stress in comparison with linear dependence in harmonic approximation. If TOEC values are positive, the trend will be opposite and non-equilibrium state of the crystal in respect of one or another elastic mode should arise. Returning to quartz one can conclude that the positive signs of C_{266} and C_{366} nonlinear elastic modules indicate an unusual behavior under the shear deformation in the XY plane.

REFERENCES

[1] B. P. Sorokin, D. A. Glushkov, and K. S. Aleksandrov, "About the temperature dependence of second order elastic constants of cubic crystals," Fizika Tverdogo Tela (in Russian), vol. 41, pp. 235-240, February 1999.

[2] B. K. Sinha, and H. F. Tiersten, "First temperature derivatives of the fundamental elastic constants of quartz," J. Appl. Phys., vol. 50, pp. 2732-2739, April 1979.

[3] B. P. Sorokin, D. A. Glushkov, P. P. Turchin, S. V. Michailyuta, K. S. Aleksandrov, and A. B. Doubovsky, "Elastic anharmonicity and elastic constants temperature dependences of different quality quartz crystals," in Proc. of 2000 IEEE/EIA Intl Freq. Contr. Symp. and Exhibition. Kansas City, USA, 2000, pp. 410-415.

[4] Y. S. Kim and T. Smith, "Themal expansion of lithium tantalate and lithium niobate single crystals," J. Appl. Phys., vol. 40, pp. 4637-4641, October 1969.

[5] J. Kushibiki, I. Takanaga, M. Arakawa, and T. Sannomiya, "Accuate measurements of the acoustical physical constants of $LiNbO_3$ and $LiTaO_3$ single crystals," IEEE TUFFC, vol. 46, pp. 1315-1323, May 1999.

[6] Y. Cho, and K. Yamanouchi, "Nonlinear, elastic, piezoelectric, electrostrictive and dielectric constants of lithium niobate," J. Appl. Phys., vol. 61, pp. 875-887, February 1987.

[7] Y. Nakagawa, K. Yamanouchi, and K. Shibayama, "Third-order elastic constants of lithium niobate," J. Appl. Phys., vol. 44, pp. 3969-3974, September 1973.

[8] R. T. Smith, and F. S. Welsh, "Temperature dependence of the elastic, piezoelectric and dielectric constants of lithium tantalate and lithium niobate," J. Appl. Phys., vol. 42, pp. 2219-2230, May 1971.

[9] R. Bechmann, A. D. Ballato, and T. J. Lukaszek, "Higher-order temperature coefficients of the elastic stiffinesses and compliances of alpha-quartz," Proc. IRE, vol. 50, pp. 1812-1822, August 1962.

[10] R. Stern, and R. T. Smith, "On the third-order elastic moduli of quartz," J. Acoust. Soc. Am., vol. 44, pp. 640-641, February 1968.

[11] A. G. Smagin, and M. I. Yaroslavskii, Piezoelectricity of Quartz and Quartz Resonators. Moscow: Energy Publisching House, 1970.

[12] A. G. Smagin, and B. G. Mil'stein, "Elastic constants of α-quartz single crystal," Sov. Kristallographiya (in Russian), vol. 19, pp. 832-836, July-August 1974.

[13] M. M. Shevel'ko, and L. A. Yakovlev, "Precise measurements of elastic characteristics of synthetic piezoquartz," Akusticheskii Zhurnal (in Russian), vol. 23, pp. 331-332, March-April 1977.

[14] D. C. Malocha, H. François-Saint-Cyr, K. Richardson, and R. Helmbold, "Measurements of LGS, LGN and LGT thermal coefficients of expansion and density," IEEE TUFFC, vol. 49, pp. 350-355, March 2002.

[15] B. P. Sorokin, P. P. Turchin, S. I. Burkov et al, "Influence of static electric field, mechanical pressure and temperature on the propagation of acoustic waves in $La_3Ga_5SiO_{14}$ piezoelectric single crystals," in Proc. of Intl Freq. Contr. Symp. Honolulu, USA, 1996, pp. 161-169.

[16] I. M. Sil'vestrova, Yu. V. Pisarevskii, I. A. Senyushenkov, and A. I. Krupnii, "Temperature dependences of elastic properties of $La_3Ga_5SiO_{14}$ single crystals," Fizika Tverdogo Tela (in Russian), vol. 28, pp. 2875-2878, September 1986.

[17] A. B. Ilyaev, B. S. Umarov, L. A. Shabanova, and M. F. Dubovik, "Temperature dependence of electromechanical properties of LGS crystals," Phys. Stat. Sol. (a), vol. 98, pp. K109-K114, December 1986.

[18] R. E. Hankey, and D. E. Schuele, "Third-order elastic constants of Al_2O_3," J. Acoust. Soc. Am., vol. 48, pp. 190-202, July 1970.

[19] J. B. Wachtman, Jr., W. E. Tefft, D. G. Lam, Jr., and R. P. Stinchfield, "Elastic constants of synthetic single crystal corundum at room temperature," J. Res. Natl. Bur. Std., vol. 64A, pp.213-228, May-June 1960.

[20] E. M. Voronkova, B. N. Grechushnikov, G. I. Distler et al, Optical Materials for Infrared Technics. Moscow: Nauka Publishing House, 1965.

[21] W. E. Tefft, "Elastic constants of synthetic single crystal corundum," J. Res. Nat. Bur. Stand., vol. 70A, pp. 277-280, July-August 1966.

[22] K. D. Swartz, and C. Elbaum, "Third-order elastic constants of zinc," Phys. Rev. B, vol. 1, p. 1512-1517, February 1970.

[23] S. I. Novikova, Thermal Expansion of Solid States. Moscow: Metallurgiya Publishing House, 1976.

[24] Landolt-Börnstein International Tables, Elastic, piezoelectric, pyroelectric, piezooptic, electrooptic constants, and nonlinear dielectric suspectibilities of crystals, New Series, Group III, vol. 11. Berlin-Heidelberg-New York: Springer-Verlag, 1979.

[25] G. A. Alers, and J. R. Neighbours, "The elastic constants of zinc between 4.2° and 670°K," J. Phys. Chem. Solids, vol. 7, pp. 58-64, October 1958.

[26] E. R. Naimon, "Third-order elastic constants of magnesium. I. Experimental," Phys. Rev. B, vol. 4, p. 4291-4296, December 1971.

[27] S. Eros, and C. S. Smith, "Low-temperature elastic constants of magnesium alloys," Acta Metall., vol. 9, pp. 14-22, January 1961.

[28] N. S. Petrenko, V. P. Popov, E. A. Pushkarev, and V. A. Finkel, "Thermal expansion of erbium at low temperatures," Phys. Status Solidi (b), vol. 68, pp. 145-148, April 1975.

[29] S. B. Palmer, E. W. Lee, and M. N. Islam, "The elastic constants of gadolinium, terbium and erbium," Proc. R. Soc. Lond. A, vol. 338, pp. 341-357, June 1974.

[30] R. Ramji Rao, and C. S. Menon, "Phonon dispersion relations and thermal expansion of erbium," J. Appl. Phys., vol. 44, pp. 3892-3896, November 1973.

[31] M. Rosen, D. Kalir, and H. Klimker, "Single-crystal elastic constants and magnetoelasticity of erbium from 4.2 to 300 °K," Phys. Rev. B, vol. 8, pp. 4399-4404, November 1973.

[32] G. Arlt, and H. Schweppe, "Paratellurite, a new piezoelectric maternal," Solid State Communications, vol. 6, pp. 783-784, November 1968.

[33] A. M. Antonenko, M. D. Volnyansky, and A. Yu. Kudzin, "Third-order elastic constants of paratellurite single crystals," Sov. Kristallographiya (in Russian), vol. 24, pp. 1071-1073, September-October 1979.

[34] Y. Ohmachi, and N. Uchida, "Temperature dependence of elastic, dielectric, and piezoelectric constants in TeO_2 single crystals," J. Appl. Phys., vol. 41, pp. 2307-2311, May 1970.

[35] H. Schweppe, "Elastic and piezoelectric properties of paratellurite (TeO_2)," Ultrasonics, vol. 8, pp. 84-87, April 1970.

[36] A. I. Korobov, V. M. Prokhorov, O. Yu. Serdobol'skaya, and P. Hegedush, "Third-order elastic moduli of KDP crystal," Sov. Kristallographiya (in Russian), vol. 23, pp. 566-569, May-June 1978.

[37] L. Boyer, et R. Vache Mesure, "Des constantes élastiques du KDP par diffusion Brillouin," Phys. Status Solidi (a), vol. 6, pp. 105–108, August 1971.

[38] S. Haussuhl, "Elastische und thermoelastische Eigenschaften von KH_2PO_4, KH_2AsO_4, $NH_4H_2PO_4$, $NH_4H_2AsO_4$ und RbH_2PO_4," Z. Kristallogr., Bd. 120, S. 401-414, Juni 1964.

[39] S. I. Chizhikov, N. G. Sorokin, I. Yu. Ledovskaya, and E. V. Makarevskaya, "Elastic propertis of KDP and DKDP crystalls under high temperatures", Sov. Kristallographiya (in Russian), vol. 18, pp 860-862, July-August 1973.

[40] K. S. Aleksandrov, B. P. Sorokin, and S. I. Burkov, Effective Piezoelectric Crystals for Acoustoelectronics, Piezotechnics and Sensors, vol. 2. Novosibirsk: SB RAS Publishing House, 2008.

Large Band-Pass BAW Filter for Space Applications

M. Chatras, L. Catherinot, S. Bila, D. Cros
XLIM, CNRS UMR 6172, University of Limoges
Limoges, FRANCE
matthieu.chatras@unilim.fr

P. Monfraix
TAS
Toulouse, FRANCE

L. Esatgerie
CNES
Toulouse, FRANCE

T. Baron, S. Ballandras
FEMTO-ST
Besançon, FRANCE
thomas.baron@femto-st.fr

***Abstract*— A resonator using shear waves of lithium niobate is used in this paper to achieve large fractional bandwidth BAW filter. Electromechanical couplings in the 20 – 50 % range are obtained for the shear waves of a thin layer of Lithium Niobate suspended on a silicon substrate. By this way a filter with fractional bandwidth over 20% has been designed for space applications.**

I. INTRODUCTION

Acoustic waves, using surface or bulk propagation, are used in numerous applications in frequency generation, control or filtering in modern wireless communication systems [1] [2]. With the growing demand for multimedia and mobile applications, new generations of telecommunication satellites require higher performances, higher functionalities and still stronger cost and size constraints [3] [4]. In that context, Bulk Acoustic Waves (BAW) devices can offer many potentialities for smart RF components or systems. For instance this technology is now used as alternative to Surface Acoustic Waves (SAW) filters in handset duplexers for UMTS and DCS standards around 2 GHz with Aluminum Nitride piezoelectric layers [5]. However, Aluminum Nitride is not suitable for large band applications, due to its electromechanical coupling coefficient. This material is mainly processed for local oscillators or narrowband filtering operations (<5%) [6] [7] [8].

Single crystal-based acoustic resonators for filtering have received a strong interest for many years. Various developments have been particularly achieved using Quartz with either SAW delay lines or resonators. However, most approaches have been developed exploiting quartz machining along standard etching, rarely compatible with batch processes as used for Micro-ElectroMechanical Systems (MEMS) [9][10][11].

That is why Lithium Niobate ($LiNbO_3$) layers bonded on silicon substrates are studied to reach large band pass specifications. It is essential to maximize the values of the electromechanical coupling coefficient in Lithium Niobate, and to use wisely crystallographic cuts to achieve very efficient filters. Thanks to Lithium Niobate shear wave propagation behavior, it is possible to synthesis and to achieve band-pass filters with fractional bandwidth over 10 or 20 %. The resonant frequency of the resonator and the filter proposed in this paper is not given because the industrial partner of this project keeps this value confidential.

II. ELECTROMECHANICAL COUPLING COEFFICIENT

By solving the generalized Christoffel equation in piezoelectric materials, we can use tensorial expressions that give phase velocity and polarization [12]. For "z" propagation (fig1.), three plane waves are created, which have orthogonal polarization with different velocities.

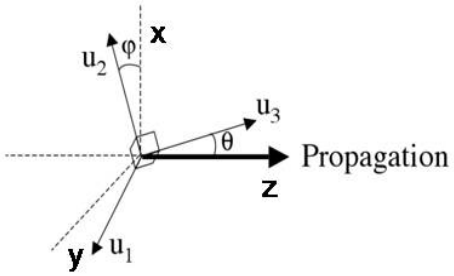


Fig.1: Propagation in anisotropic crystal

It results a thickness wave polarized along the "Z" axis, and two shear waves, polarized in the "XY" waves plane. These velocities permit to define very important parameters for BAW resonators: thickness, electromechanical coupling coefficient kt^2 and shear electromechanical coefficient ks^2.

The higher these coefficients are, the best electromechanical coupling the piezoelectric layer gives. The following expression describes the dependence of kt^2 (or k_s^2) according to resonance and anti-resonance.

$$k_t^{\,2} = \frac{\pi^2}{4}\left(\frac{f_p - f_s}{f_p}\right)$$

It is not possible for Aluminum Nitride to have a coupling coefficient kt^2 higher than 7%. We propose to use two different cuts of Lithium Niobate: the (YXl)/ 36° and the (YXl)/163° cuts. The first one allows for the excitation of high velocity (7000 $m.s^{-1}$) longitudinal modes whereas shear waves are used in the second one, allowing for electromechanical coupling. For these orientations we obtain k_s^2 around 30% for the (YXl)/36° cut and around 45% for the (YXl)/163° cut.

III. 1D Simulation

The first step is to size the thicknesses of the several layers of the structure to reach the targeted frequencies. In the case of figure 2, we used the cut Y+36°; we obtain a coupling coefficient kt^2 equal to 33%. This resonator with very large electromechanical coupling will enable to reach easily fractional bandwidth filter over 10%.

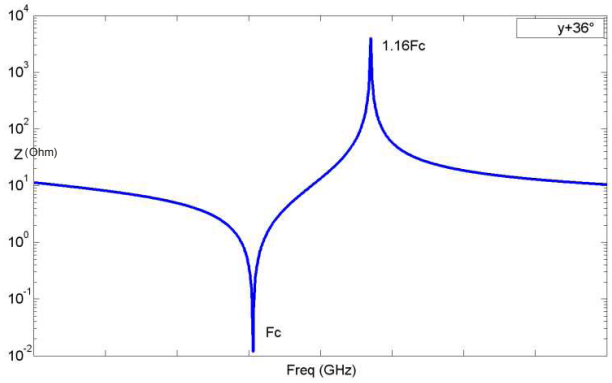


Fig.2: simulation response of a lithium niobate resonator (cut Y+36).

IV. Fabrication and Measurements

In this work, we propose a filter fabricated on $LiNbO_3$/Silicon substrates obtained by Au/Au bonding at room temperature and a lapping/polishing on the upper face of the lithium niobate substrate. This approach allows for a collective and accurate production of filters, the filter frequency being controlled by the membrane thickness of the lithium niobate layer.

We based our device fabrication on gold bonding with a lapping/polishing process to prepare wafer compound to different devices. It allows us to manufacture a BAW resonator on $LiNbO_3$ membrane as shown in fig. 3. A gold thin layer (200nm) is deposited first by sputtering on both $LiNbO_3$ (bottom face) and silicon (top face) wafers. Both wafers are then bonded together via gold layer compression into an EVG bonding machine. During this process, we apply a pressure of 65 $N.cm^{-2}$ to the whole contact surface, yielding a high quality bond. $LiNbO_3$ is subsequently thinned by lapping and polishing steps to an overall thickness of several microns. The resonator is finally suspended by etching the back side of the silicon substrate by DRIE. Aluminum electrodes are deposited on the $LiNbO_3$ upon the cavity to achieve working BAW resonators. A cross section of the resonator is presented figure 3.

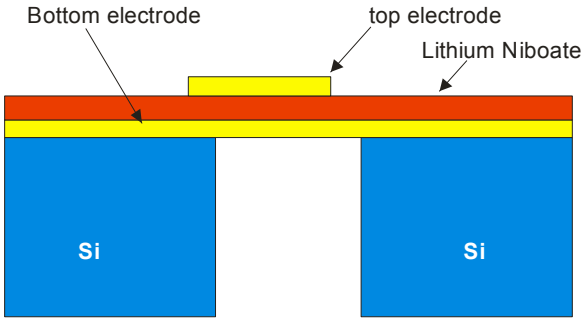


Fig.3: cross section of the structure

Photography of the fabricated resonator is presented figure 4, in a top view. We can see CPW accesses for measurement and apodized shape of the electrodes to prevent from parasitic modes.

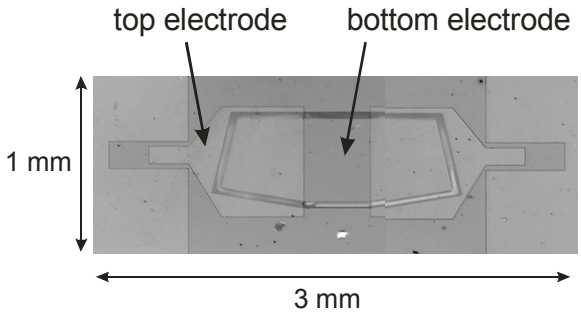


Fig.4: fabricated resonator (top view).

S_{21}-parameter measurements for the structure shown in Fig. 4 were performed using a HP 8510 A network analyzer system. The measured response is presented figure5.

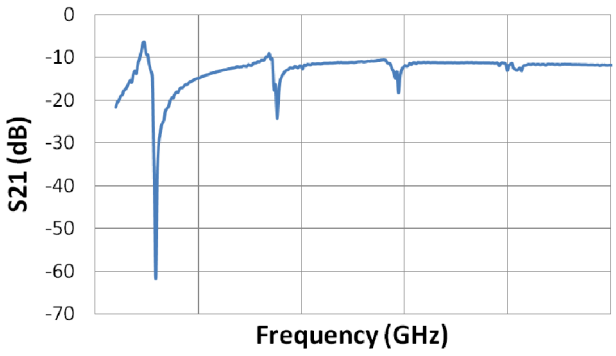


Fig.5: measured resonator.

It exhibits an electromechanical coupling of 31 %, wich is very close to the predicted one (33%). This resonator can then been used to synthesis large fractional bandwidth filter. Figure 6 presents a lithium niobate based BAW 3 pole filter with a 3 dB fractional bandwidth of 22 % in a ladder configuration.

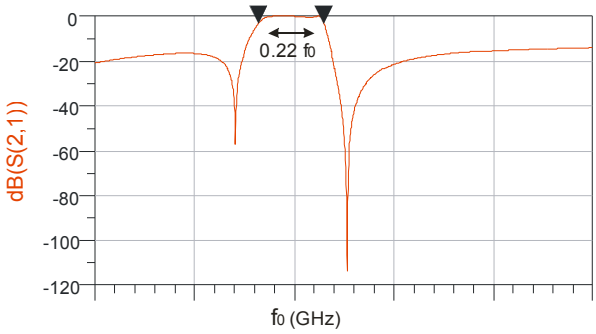


Fig.6: computed ladder 3 pole filter with large fractional bandwidth for space application.

V. Conclusion

In this paper an efficient method to compute BAW filter using shear waves has been presented using a scalar approach. By this way, very important electromechanical coupling coefficients have been obtained with lithium niobate. Large band pass filter become then very easy to achieve for space applications.

Acknowledgment

Authors want to thank the CNES for the program FOVETTES support.

References

[1] K.M.Lakin, G.R. Kline, K.T. McCarron, “High-Q Microwave Acoustic Resonators and Filters”, 1993 IEEE Microwave Symp. Digest, vol. 3, pp 1517-1520.

[2] R. Aigner ‘SAW and BAW Technologies for RF Filter applications: A Review of the Relative Strengths and Weaknesses’, Proceedings of IEEE Ultrasonics Symposium, Beijing/China, 2 – 5 Nov 2008.

[3] R. Aigner, J. Kaitila , J. Ellä, L. Elbrecht, W. Nessler, M.Handtmann, T.-R. Herzog, and S. Marksteiner, “Bulk-Acoustic-Wave Filters: Performance Optimization and Volume Manufacturing”, proceedings of IMS-MTTS, Philadelphia, 2003.

[4] Bar, P.; Giry, A.; Triolet, P.; Parat, G.; Pache, D.; Ancey, P.; Carpentier, J.F. “Full deplex receiver and PA integration with BAW devices”, IEEE SiRF 2008, p 9-12.

[5] S. Giraud, S. Bila, M. Chatras, D. Cros, M. Aubourg, « Bulk acoustic Wave Filter Synthesis and Optimization for Multi-Standard communication Terminals” IEEE TUFFC, pp 57, pp 52-58, JAN 2010.

[6] Reinhardt, A.; Parat, G.; Defay, E.; Aid, M.; Domingue, F., “Acoustic technologies for advanced RF architectures”, 2010, IEEE Newcas conference, pp 161-164.

[7] R.Lanz, P.Muralt, “Solidly Mounted BAW Filters for 8 GHz based on AlN Thin Films “, 2003 IEEE Ultrasonic Symp,pp 178-181.

[8] C. Zuo, J. Van der Spiegel, G. Piazza, “1.05 GHz CMOS Oscillator based on Lateral-Field-Excited Piezoelectric AlN Contour-Mode MEMS Resonators”, Proc. 2009 IEEE Int. Freq. Contr. Symp., pp. 70 -74, 2009.

[9] Volatier-A; Defay-E; Aid-M; Nhari-A; Ancey-P; Dubus-B, Switchable and tunable strontium titanate electrostrictive bulk acoustic wave resonator integrated with a Bragg mirror, Applied Physics Letters; 92(3): 032906 – 2008.

[10] G. Piazza, P.J. Stephanou, A. P. Pisano, "Piezoelectric AluminumNitride Vibrating Contour-Mode MEMS Resonators," Microelectromechanical Systems, Journal of, vol.15, no.6, pp.1406-1418, Dec. 2006.

[11] X. Zhu, V. Lee, J. Phillips, and A. Mortazawi, "Intrinsically switchable contour mode acoustic wave resonators based on barium titanate thin films," Microwave Symposium Digest, 2009. MTT '09. IEEE MTT-S International, vol., no., pp.93-96, 7-12 June 2009.

[12] K. Y. Hashimoto, “RF Bulk Acoustic Wave Filters for Communications”, Artech House, 2009.

Effect of substrate material and electrode surface preparation on stress and piezoelectric properties of aluminum nitride

Sergey Mishin
Research and Development
Advanced Modular Systems, Inc
Goleta, CA/USA
smishin@amssb.com

Michael Gutkin
Technical Marketing
Advanced Modular Systems, Inc
Goleta, CA/USA
mgutkin@amssb.com

Abstract — Intrinsic stress and electromechanical coupling coefficient are the most important parameters of piezoelectric thin films used in membrane as well as cantilever type MEMS such as energy harvesters, transducers and others.

In this paper, practical aspects controlling stress and crystallinity in aluminum nitride (AlN) film used in manufacturing of Bulk Acoustic Wave (BAW)/ Film Bulk Acoustic Resonator (FBAR) filters [1], [2], [4] have been presented. Controlling stress by modifying deposition parameters was compared to controlling stress by surface modification. Stress was found to be an important indicator in predicting crystallinity of AlN films. Impact of the substrate material and electrode deposition on stress and crystal orientation of piezoelectric AlN was studied. In process applications where control of the surface layer thickness is critical and surface topography is variable, use of ion mill to reduce surface roughness is preferable to CMP. It was generally found that for most amorphous materials such as silicon dioxide or silicon nitride that are placed under the electrode or piezoelectric AlN, it is critical to get the best possible surface smoothness. On the other hand for the electrode material such as molybdenum, it is more important to get material with the best crystal orientation. Polishing of the well oriented electrode material did not provide further improvement. It was also observed that if AlN is deposited on a smooth or well oriented layer, orientation of the AlN is improved when the stress in the film is more tensile. On the other hand if the tensile stress is caused by residuals on the surface or surface roughness, tensile stress usually leads to poorly oriented films.

I. INTRODUCION

Early in the development of AlN films, it was observed that stress could vary for variety of reasons. Sometimes most compressive stress in the AlN film produced best crystallinity, other time tensile stress produced best results. Controlling stress was particularly important for FBAR applications where membranes are susceptible to cracking or delaminating due to high stress. Even in BAW technology, high stress can simply cause yield loss in photo-lithography due to high warpage of the wafers. Three major factors must be addressed to control stress in the AlN films:

- Substrate or pre-electrode processing
- Electrode or pre-AlN processing
- Deposition of AlN

We propose some practical guidelines to successful stress control with high electromechanical constant.

II. EQUIPMENT

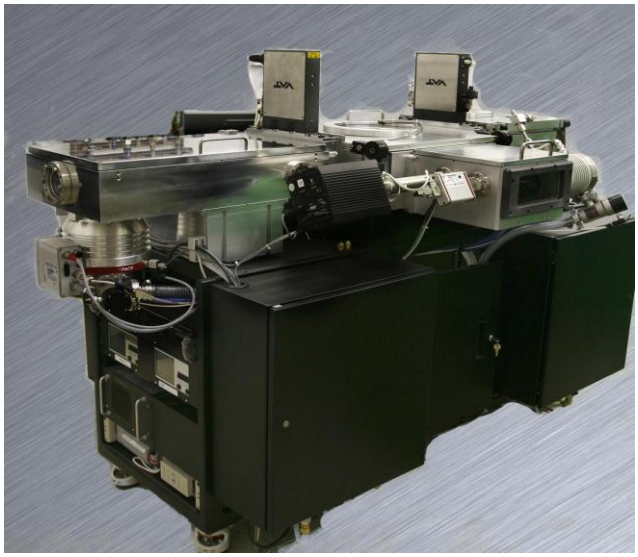
In this investigation we used Advanced Modular Systems cluster tool with three modules: AlN deposition chamber, electrode deposition chamber and ion beam trimming module (shown in Figure 1).

AlN deposition uses a dual magnetron with AC power applied between targets. It is a reactive deposition using aluminum target and, argon and nitrogen process gasses. Electrode deposition chamber is a dual magnetron with DC power supply, electrode target (Mo, W, Pt, etc…) and argon process gas. The trimming module uses DC focused ion source with argon process gas to improve thickness/uniformity of either electrode or AlN films. Use of the trimming module opens up a much wider process window

978-1-61284-110-6/11/$26.00 ©2011 IEEE

for stress control because we don't need to spend too much effort on controlling thickness uniformity during deposition.

Figure 1: AMSystems cluster tool



III. SUBSTRATE OR PRE-ELECTRODE PROCESSING

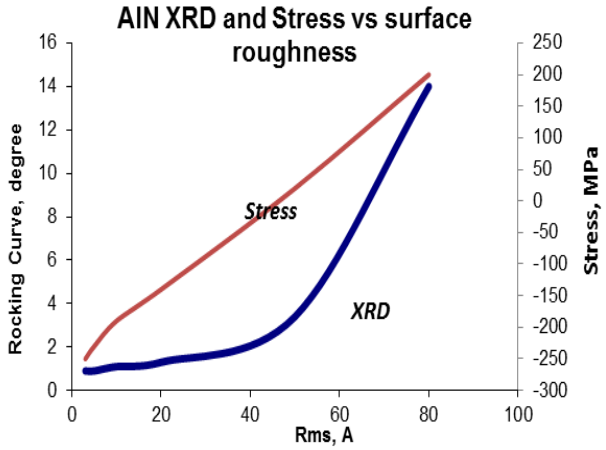
A. *Surface Cleaning*

The surface on which the electrode is deposited can irreparably damage quality of the piezoelectric aluminum nitride. It is very critical to have a smooth and uncontaminated surface before depositing the electrode material. If substrate was exposed to the photoresist, it is critical to have the surface adequately cleaned. Depending on the materials used, either wet chemical or dry strips can be effective. If wafers are not adequately cleaned, it can lead to high tensile stress in the future deposited films with very poor crystal orientation and, in worse cases, delamination.

B. *Surface smoothness*

After adequate cleaning it is important to make sure that the surface is smooth. Typically, 0.6nm surface roughness is acceptable. There are two common methods for the surface smoothing: Chemical Mechanical Polishing (CMP) and ion beam milling. CMP can be performed on most materials and with appropriate selection of polishing slurry it is capable of both planarizing and smoothing the surfaces. One limitation is that it is hard to control the amount of material removed, especially if there is diverse topography of features on a wafer. Ion milling can be precise in the amount of material removed, but is only effective on hard dielectric materials such as SiN, AlN or SiO2. It requires too much material removal on metals to be useful. Figure 2 shows impact of surface roughness on the crystallinity and stress of AlN deposited on the PECVD oxide that is polished to different level of smoothness. Ion beam trimming tool was used to improve the surface roughness.

Figure 2: AlN Rocking curves and Stress as a function of oxide surface roughness

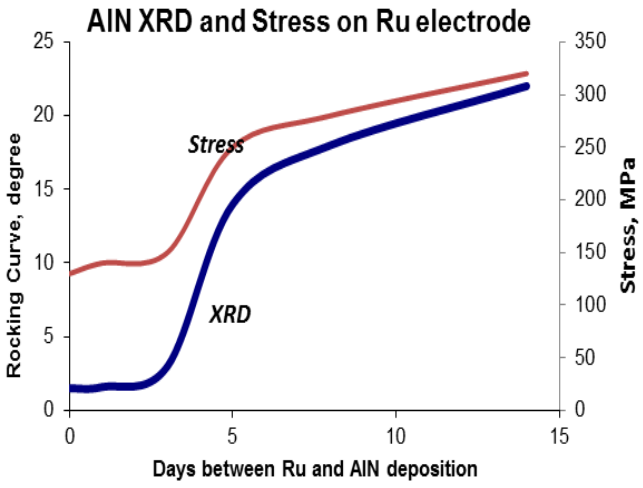


IV. ELECTRODE OR PRE-ALN PROCESSING

Before piezoelectric AlN is deposited, the electrode material is typically deposited and in most cases patterned. It is important to keep track of the stress in the electrode films because it is a good indicator of the problems with the substrate preparation. Erratic stress in the electrode usually results in problems with AlN deposition.

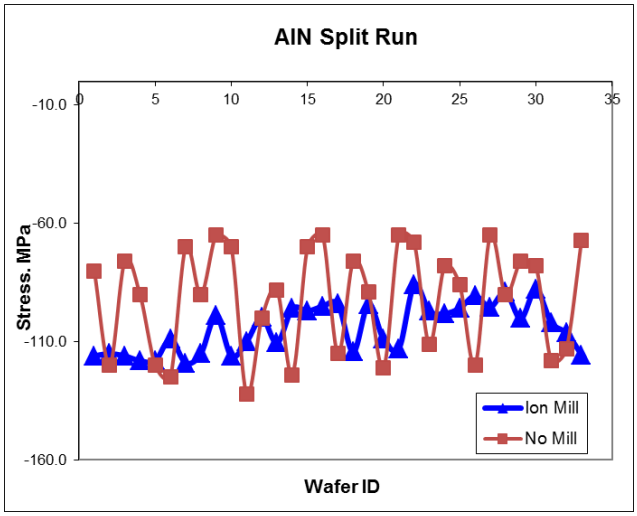
Some electrode materials such as Ru oxidize in the air if AlN is not deposited within relatively short period of time. Significant variations in both stress and rocking curves have been observed if Ru electrode is left in the ambient atmosphere for a long time. Figure 3 shows the impact on AlN parameter as Ru is left exposed for significant amount of time.

Figure 3: AlN properties on differently exposed Ru electrodes



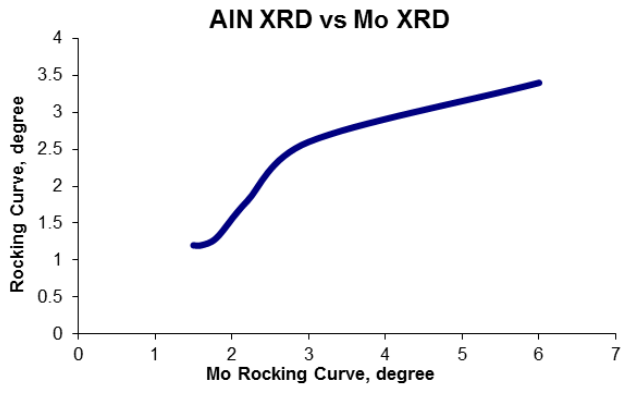
Even though electrodes such as W and Mo are not easily oxidized, it was found that removal of 10nm to 20nm of the electrode material by CMP has beneficial impact on repeatability of stress as can be seen in Figure 4. Polishing the surface of the W also produces improvement in rocking curve, but no improvement was observed for Mo.

Figure 4: Repeatability of AlN stress on polished vs. unpolished electrodes



Measuring rocking curve on Mo deposition has been useful in predicting better crystal orientation in the following AlN deposition, see Figure 5. Rocking curve measurements had no correlation to the rocking curves on AlN or stress repeatability for either W or Ru.

Figure 5: AlN rocking curve depends on Mo electrode property



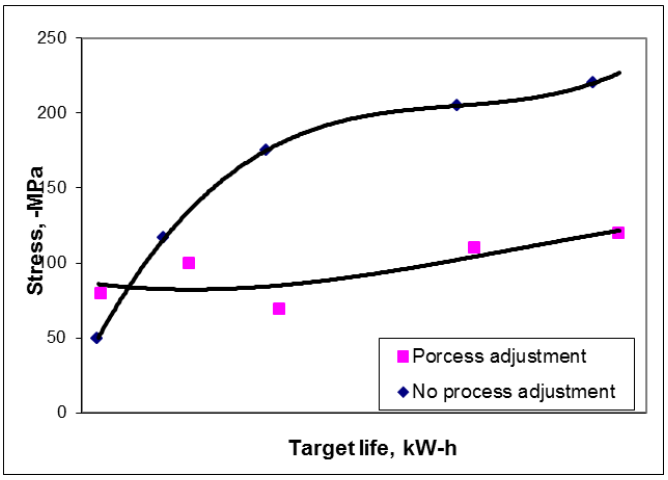
V. DEPOSITION OF AlN

Stress control in the AlN deposition is typically accomplished be controlling either deposition pressure or substrate bias. Higher deposition pressure produces more tensile stress and higher substrate bias more compressive stress [3]. It is important to establish baseline stress repeatability in the deposition module. High resistivity Si is best suitable for such application. Wafers with resistivity of >15ohm-cm are acceptable. Lower resistivity can lead to substrate interaction with the chamber components and variable stress results.

Another common cause of stress variation is a plasma-substrate coupling due to variation of plasma potential and contact through the electrode structure to the wafer holding/clamping mechanism. We have observed as much as 700MPa difference in stress due to the conductive contact to the wafer surface vs. insulated contact.

As the target material gets eroded and a thick insulating layer builds up on the shields, it is not uncommon to see stress changing from the beginning of the target life to the end of the target life. Usually such changes are repeatable from target to target and can be built into automated software adjustment on the deposition system, see Figure 6.

Figure 6: AlN film stress vs. target life



VI. CONCLUSION

In many cases it is more important to prepare the substrate and the electrode for the AlN deposition than trying to control deposition itself. Making sure that both electrode and substrate have been treated correctly will go a long way towards producing excellent AlN films.

REFERENCES

[1] Robert Aigner "Corrective Actions to Meet Extreme Tolerance Requirements for Thin Films" 2007 Octagon Communication Inc. DBA MPDigest

[2] Frank Z. Bi and Bradley P. Barber "Bulk Acoustic wave RF Technology", 2008 IEEE

[3]K. Hashimoto "RF BULK ACOUSTIC WAVE FILTERS FOR COMMUNICATIONS", pp.173-195

[4]Sergey Mishin, Dan Marx, Brian Sylvia, V. Lughi, K. L. Turner, and D. R. Clarke "Sputtered AlN Thin Films on Si and Electrodes for MEMS Resonators: Relationship between Surface Quality Microstructure and Film Properties" 2005, IEEE

Multiple Frequency Solidly Mounted BAW Filters

Alexandre Reinhardt,
Jean-Baptiste David, Christine Fuchs
CEA LETI, Minatec Campus,
Grenoble, France
E-mail: alexandre.reinhardt@cea.fr

Marta Clement, Jimena Olivares,
Enrique Iborra
Grupo de Microsistemas
y Materiales Electrónicos,
Universidad Politécnica de Madrid
Madrid, Spain

Nick Rimmer, Steve Burgess
SPP Process Technology Systems UK Ltd.,
(formerly Aviza Technology)
Newport, South Wales,
United Kingdom

***Abstract*—This paper reports the simultaneous fabrication of Receive and Transmit Bulk Acoustic Wave filters for the WCDMA standard on the same die. Both filters are based on Solidly Mounted Resonators using a common Bragg mirror, but with each having a specific piezoelectric film thickness. Electrical measurements reveal that the process steps required to provide the two different piezoelectric film thicknesses on the same die does not impact the electrical performances of resonators and filters and that this approach could thus be generalised to more than two filters.**

I. Introduction

Acoustic filters are an enabling technology for radiofrequency front-ends, where they are used for RF band pass filtering. Among them, Bulk Acoustic Wave (BAW) filters are commercially available for almost 10 years now, and find their use for stringent filtering applications, where low losses, steep skirts and sharp roll-off are required [1]. They consist in a piezoelectric thin film sandwiched between two electrodes, and operating on a thickness-extensional vibration mode. Therefore, their operation frequency is fully defined by the set of layers from which they are obtained and especially by thicknesses. Unlike other technologies, like Surface Acoustic Wave (SAW), Lamb Wave Resonators, or other Micro Electro Mechanical (MEMS) resonators, their characteristics depend only weakly on their geometric design. This means that a unique set of resonators, operating at a given frequency, can be fabricated at a time. Local addition of mass by depositing an added layer over some resonators enables to shift their frequency through the mass loading effect, and thus enables to obtain the two sets of resonators required to obtain a conventional ladder or lattice type filter. However, it is hardly possible to synthesize two filters for different standards. Instead, only one bulk acoustic wave filter is usually fabricated at a time and current RF front-ends thus require as many filter dies as they include filters. This will surely become a critical problem in mobile phone applications with newer standards making use of even more frequency bands, what will cause an increase in the number of filters necessary in a radiofrequency front-end. Therefore, the integration of these components in systems will become critical.

Previous studies have focused on co-integrating two RF filters on the same die, at least to provide both Receive (Rx) and Transmit (Tx) filters of the same standard on the same die, thus enabling to assemble only a single filter die on a duplexer module. This means being able to provide four sets of resonators having each a different set of resonance frequencies. The most intuitive way is to exploit the mass-loading effect in a way similar as to the one in which two sets of frequencies are defined in a conventional filter [2]. It has been recently applied for example to the fabrication of a Film Bulk Acoustic Resonator (FBAR) based duplexer for the WCDMA standard [3]. By using both a so-called band tuning layer to define resonators at Rx frequencies, and then a second loading layer to define the parallel resonators of both filters, a set of four frequencies was obtained. The same principle could be similarly applied with Solidly Mounted Resonators (SMR). The advantage of this solution is that it requires only an additional layer and an additional mask level used to pattern the band tuning layer compared to a conventional BAW process. However, this additional band-tuning layer normally causes a reduction of the effective electromechanical coupling factor of resonators, making the synthesis of wide bandwidth filters more difficult. This reduction increases as the frequency shifts between the two filter bands becomes large. This means in practice that this approach can only be used where a reduction in effective electromechanical coupling factor for one of the two filters is acceptable. This also means that this solution is only applicable to two close frequency bands.

In order to fabricate simultaneously filters with arbitrary bandwidths and frequency spacing, without sacrificing effective electromechanical coupling factor, i.e., ultimately, insertion losses and ripple, several piezoelectric layer thicknesses need to be provided on the same substrate [4]. A first approach proposed to implement such a device was first to deposit the thinner piezoelectric film, then to protect it where the highest frequency filter will be located, and finally to deposit an additional thickness of the piezoelectric layer where the lowest frequency filter will be located [5]. However, authors have shown that in practice, with Aluminum Nitride (AlN), piezoelectric properties of the thicker film are decreased compared to the initially deposited one. Investigations have proven that the cause of this decreased piezoelectric properties is attributed to either a bad crystallinity of the re-grown layer, or by a regrowth of AlN with a reverse polarity, causing a partial cancellation of the effective piezoelectricity of the piezoelectric film. The origin of this phenomenon is the surface

978-1-61284-110-6/11/$26.00 ©2011 IEEE

degradation of the initial AlN film, caused by reaction to chemical agents during photolithography steps, but it was also noticed, although less significantly, in samples having only been exposed to air for some time. The conclusion of this first approach was that oxidation or reaction with humidity may trigger surface deterioration and cause re-growth with the wrong polarity.

Another approach would be to start with the deposition of the thickest piezoelectric layer first, and then to partially etch this film to provide a thinner layer where the highest frequency filter will be located. The advantage is that it ensures that the structural quality of AlN remains unaffected. However, there are two critical issues: first AlN is still exposed to chemicals during photoresist development or stripping, which are known to damage it, in partially etching it and increasing its surface roughness. Conjugated effects of roughness increase and surface contamination of the piezoelectric film are expected to prove detrimental towards quality or effective electromechanical coupling factors of resonators. Secondly, although the deposition processes of AlN in an industrial environment is very well controlled and reaches unmatched thickness homogeneity in order to reduce frequency dispersions, while offering growth rates compatible with mass production in an industrial line, etching processes are much less optimized. Processes with reasonable etch rates suffer from much lower uniformity and thickness control, especially when it comes to partially etch a layer without proper etch-stop. On the other hand, some equipments, for example dedicated to frequency trimming of resonators, offer an excellent control over etching depth and uniformity, but at the expense of very slow etch rates, making these approaches difficult to implement for mass production. So using direct partial etching of AlN does also not seem to be an interesting alternative.

To avoid the drawbacks of the above-mentioned methods for fabricating several BAW filters simultaneously, we propose in this paper a new approach based on multiple depositions and patterning of the piezoelectric layer [6]. In section II, we discuss the design of two filters compatible with the proposed approach. Then, in section III, we describe the fabrication of both Reception and Transmission filters for the WCDMA standard. Section IV provides the electrical characteristics of the fabricated filters and a comparison with conventional filters with the same design.

II. WCDMA DUPLEXER FILTERS DESIGN

The two filters were designed to match the WCDMA standard, using the procedure described in [7], [8]. Based on a first estimation of achievable resonator electromechanical coupling and quality factors, Butterworth-Van Dyke (BVD) lumped element models of resonators were built and used for filter synthesis. An equivalent filter model, based on an initial filter architecture, is first proposed. The transmission filter is based on a 50 Ω single stage ladder topology, with only three resonators, since most of the rejection will be provided by the duplexer itself. The reception filter is implemented as a double lattice topology, as it requires higher rejection and

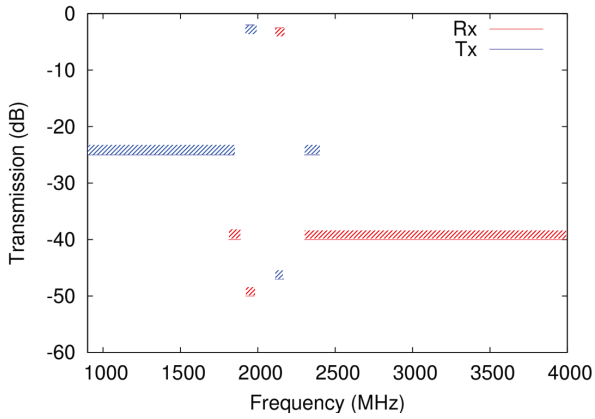


Fig. 1. Filter masks for the WCDMA duplexer : antenna to Rx path (red) and Tx to antenna path (blue).

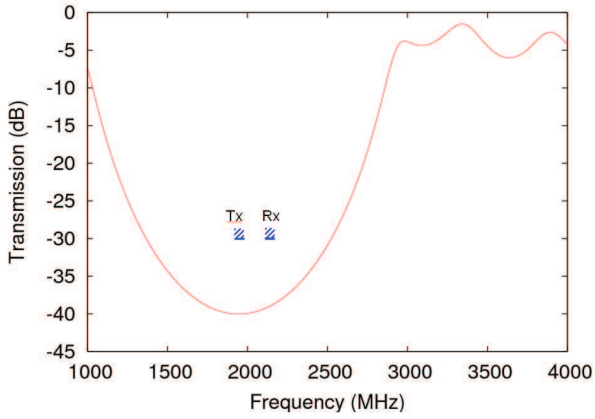


Fig. 2. Transmission coefficient of the designed SiN/SiOC Bragg mirror.

differential input and output. This model is then optimized to meet specifications for the whole duplexer module, which are summarized by the masks shown in Figure 1. The output of this optimization basically provides required resonator static capacitances and, more importantly for this study, resonance and antiresonance frequencies.

From this basis, a more precise Mason model of resonators is used. First, a Bragg mirror is designed to cover the two filter bands. This mirror is made of SiN and SiOC, which exhibit an acoustic impedance ratio larger than 6 and thus can provide a mirror offering a transmission lower than -30 dB over a bandwidth of more than 1 GHz with 3 pairs of layers, as seen in Figure 2. Such a wide bandwidth is more than sufficient to cover the two filter bands. It also ensures a rather small change in reflection coefficient between the filters: this is mandatory to ensure that both devices will benefit from a quasi-similar acoustic isolation and from small changes in phase of the reflection coefficient, what is necessary for the maximum achievable effective electromechanical coupling factor of resonators to remain similar for the two filters.

Two Molybdenum/AlN/Molybdenum piezoelectric stacks have been designed so that resonators simultaneously reach

the specified resonance and antiresonance frequencies in the series paths of the two filters. This design also takes into account an additional SiN passivation layer whose thickness is otherwise defined for reliability reasons [9]. However, we added the additional constraint to try to have as many common layers as possible, to ease subsequent processing. For the set of frequencies defined during the optimization of the filter topology, it proved impossible to achieve sufficient electromechanical coupling factor with the same piezoelectric layer thicknesses or with the same bottom or top electrodes. However, the required amount of mass loading necessary to induce frequency shifts of 69 and 56 MHz between series and parallel resonators, respectively for the Rx and Tx filter, is obtained by almost adding the same thickness for the two material stacks. This means that the same mass-loading can be applied at least to the two filters.

III. Filter fabrication

From these acoustic design considerations, one can see that the fabrication of the two filters differs only at the level of the two electrodes and at the level of the piezoelectric film. Therefore, for most of the fabrication process, we use a conventional BAW fabrication process. We start with the deposition onto a high resistivity wafer of the 7 layers forming the common Bragg mirror, alternatively SiOC (335 nm) and SiN (1200 nm). These layers are deposited by chemical vapor deposition (CVD). The whole deposition process has been optimized to minimize stress wile providing a flat and smooth surface on which to grow the piezoelectric stacks.

A Molybdenum film is deposited on top of these Bragg mirrors by sputtering the thickest of the two bottom electrodes, i.e. 400 nm of Molybdenum corresponding to the thickness required for the Tx filter. This metal film is then locally partially etched over 50 nm to reach the thickness required by the Rx filter. To ensure a precise control of this thickness difference, we use very slow etching process in an ion beam tool. Then, the bottom electrodes for the two filters are simultaneously patterned using standard photolithography and a fluorine-based dry etching step.

A first piezoelectric AlN film close to 1300 nm is deposited by reactive pulsed DC sputtering, followed by the deposition of the first top molybdenum electrode (140 nm) in the same cluster tool (a Sigma Fxp sputtering system from Aviza Technologies) to avoid any vacuum break. These layers form the piezoelectric stack for the Tx filter. Large patterns are then defined on the top metallization to define areas where Tx filters will reside, using standard photolithography and dry etching. The molybdenum film then serves as a metal mask for the chemical wet etching of AlN. We checked that this chemical etching leaves a flat surface without any AlN residues which could locally degrade the growth of the second piezoelectric stack.

This second piezoelectric stack is deposited the same way as the previous one on top of the existing materials. It consists in an AlN film of 1150 nm and a Mo film of 120 nm. The top Mo film is patterned to define a second metal mask,

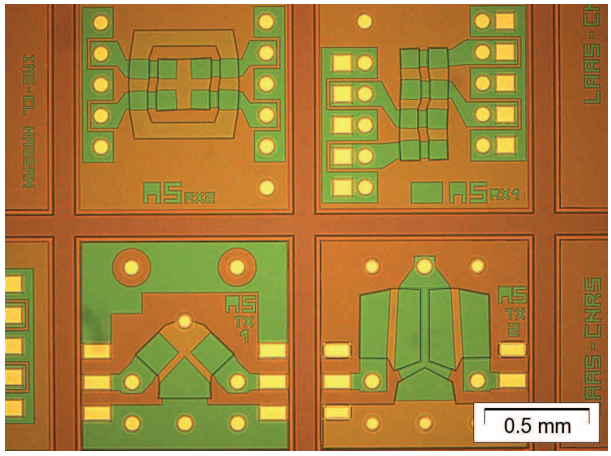


Fig. 3. Optical photograph of Rx and Tx filters after fabrication.

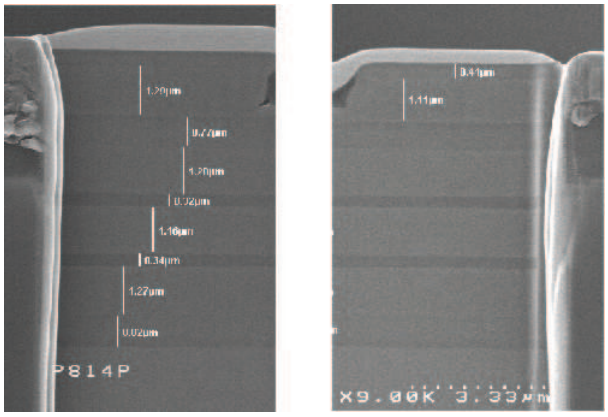


Fig. 4. SEM cross sections of Tx (left) and Rx (right) islands.

complementary of the previous one, which defines the areas where the Rx filters will reside, and the second AlN layer is etched again. At this stage, the first Mo metal mask protects the first AlN film, acting as an etch stop for the second AlN, so that two kinds of islands with varying piezoelectric stacks are defined. Microscope inspection revealed that no remains of the second AlN film were noticed over the top electrode of Tx islands after this etching step, so that the fabrication process could be continued safely. We also assessed the crystalline quality of AlN films by rocking curve measurements. No noticeable change in AlN texture was noticed between Rx and Tx islands and between reference wafers which have only seen the deposition of one piezoelectric stack.

Frequency separation between series and parallel resonators for the two filters is obtained by partially etching the two top molybdenum layers over the locations where series resonators will reside. As for the bottom electrode, we partially etch these layers using slow ion beam etching for a precise thickness control and a high on-wafer uniformity. The top electrodes are then patterned and vias are etched in the AlN islands using photolithography and another wet etching to reach the bottom electrodes. Devices are then passivated using an SiN film covering the whole surface, except where contacts towards

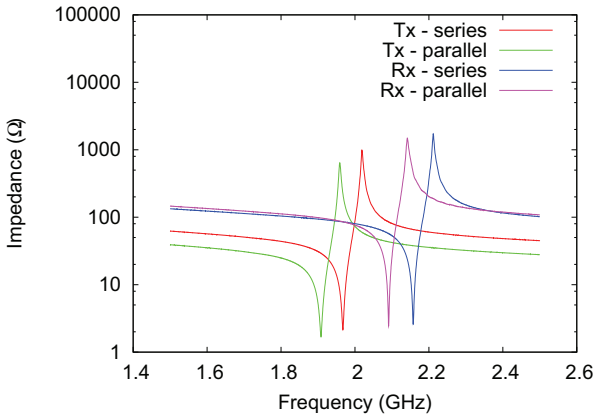


Fig. 5. Electrical response of Tx and Rx resonators measured on the same wafer die.

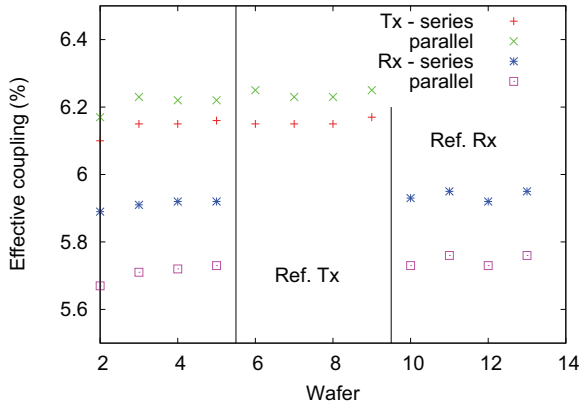


Fig. 6. Comparison of average electromechanical coupling factors measured between wafers having seen the fabrication of the two filters and reference filters.

top and bottom electrodes are opened by fluorine based dry-etching. Finally, a thick aluminum layer is deposited and patterned to cover vias and to form contact pads for RF probing.

Figure 3 is an optical photograph showing filters after fabrication. Each filter resides on an 1×1 mm^2 die having its own piezoelectric stack. Vertical and horizontal lines separate islands corresponding to the Rx or the Tx material stack. Due to the relatively small change in AlN thickness between the two types of filter, no noticeable colour change can be observed. Figure 4 shows cross sections of Tx and Rx islands obtained on the same wafer at a small distance from each other.

IV. Electrical Characterizations

After fabrication, a set of 167 test resonators regularly distributed accross each wafer were tested using RF probes connected to a vector network analyzer and calibrated using the conventional SOLT method. For each of them, we extracted resonance and antiresonance frequencies, as well as quality and electromechanical coupling factors. Figure 5 shows sample impedance responses of resonators for the Rx and Tx filters measured on the same wafer die. Figure 6 gives the values of electromechanical coupling factor averaged over each wafer (i.e. over 167 samples) compared to reference wafers having only seen the fabrication of one filter. As can be seen, changes in resonator properties remain within measurement uncertainties, showing that the multiple frequency process has not significantly impacted resonator performances. Measurements of two filters located in the same 2 mm^2 die are shown in Figure 7: these two filters exhibit excellent performances: respectively 1 and 2.3 dB insertion losses for the Transmit and Receive filters. It is to be noted that the somehow reduced rejection of the Transmit filter will be greately improved when mounted inside the duplexer chip.

V. Conclusion

We have simultaneously fabricated Rx and Tx filters for the WCDMA standard based on SMR. The fabrication process

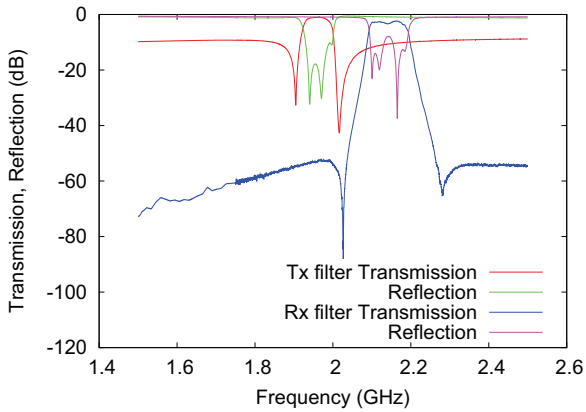


Fig. 7. Transmission and reflection measurement for adjacent Tx and Rx filters.

employed makes use of the deposition and patterning of two piezoelectric films which are then patterned thanks to the very large etching selectivity between AlN and the Molybdenum electrodes used. Electrical measurements reveal that no change in performances is induced in individual resonators or in filters compared to a standard single-frequency BAW process, proving the interest of this method, and moreover does not put strict constraints in terms of acoustic design of resonators. Moreover, this approach can be easily generalized to alternative technologies using other electrode materials or other piezoelectric films. It could also be generalized to a larger number of filters, being only limited by the ability to design a common Bragg mirror suitable for all filters.

Acknowledgment

The authors acknowledge financial support from the European Commission in the FP6 IST 2004-027003 MOBILIS project. They want to thank the whole LETI BAW team for fruitful discussions, for their support and for their deep involvement in device fabrication and characterization. They also thank all partners of the MOBILIS project for their contributions.

REFERENCES

[1] R. Aigner, "SAW and BAW technologies for RF filter applications: a review of the relative strengths and weaknesses," in *2008 IEEE International Ultrasonics Symposium Proceedings*, 2008, pp. 582–589.

[2] K. M. Lakin, K. T. McCarron, J. Belsick, and R. Rose, "Filter banks implemented with integrated thin film resonators," in *Proceedings of the 2000 IEEE International Ultrasonics Symposium*, 2000, pp. 851–854.

[3] K. Nam, Y. Park, B. Ha, C. Kim, J. Shin, S. Yun, J. Pak, G. Park, and I. Song, "Monolithic 1-chip FBAR duplexer for W-CDMA handsets," *Sensors and Actuators A*, vol. 143, pp. 162–168, 2008.

[4] J. Ella and H. Pohjonen, "BAW filters having different center frequencies on a single substrate and a method for providing the same," United States Patent US 6,518,860 B2.

[5] F. Martin, P. Muralt, M. Cantoni, and M. Dubois, "Re-growth of c-axis oriented AlN thin films," in *Proceedings of the 2004 IEEE International Ultrasonics Symposium*, 2004, pp. 169–172.

[6] E. Iborra, M. Clement, and J. Olivares, "Device for filtering bulk acoustic waves," Patent application WO2010/116011 A1, Oct. 2010.

[7] S. Giraud, S. Bila, M. Chatras, D. Cros, and M. Aubourg, "Bulk acoustic wave filter synthesis and optimization for UMTS applications," in *Proceedings of the 39th European Microwave Conference*, 2009, pp. 456–459.

[8] M. Clement, E. Iborra, J. Olivares, N. Rimmer, S. Giraud, S. Bila, and A. Reinhardt, "DCS Tx filters using AlN resonators with iridium electrodes," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 57, no. 3, pp. 518–523, March 2010.

[9] R.-Y. Fillit, B. Ivira, J. Boussey, R. Fortunier, and P. Ancey, "Structural and thermal investigation for FBAR reliability in wireless applications," in *Proceedings of the 43rd Annual International Reliability Physics Symposium*, 2005, pp. 342–346.

Resonant Subterahertz Coherent Acoustic Waves Excitation by Josephson Junction

Natalia I. Polzikova, Georgy D. Mansfeld, Yury S. Tokpanov, and Valery P. Koshelets
Kotel'nikov Institute of Radio-Engineering and Electronics, Russian Academy of Science
Moscow, Russia
polz@mail.cplire.ru

***Abstract*—A superfine resonant structure with a voltage spacing of about 19 nV (corresponding frequency 9 MHz) and an extremely low differential resistance has been observed in the Josephson Flux Flow Oscillator (FFO) IVCs. We have analyzed coherent phonon radiation and detection due to the interaction of Josephson's electromagnetic oscillations with mechanical displacement field caused by piezoelectric effect or electrostriction effect. In contrast to the other works we consider sound generation in the Josephson tunneling structure, deposited on the rather thick substrate. This layered structure plays the role of high overtone composite resonator for bulk acoustic waves propagating normally to the layers.**

I. INTRODUCTION

The spectrum of acoustic phonons covers frequencies up to several terahertz. So the excitation of coherent acoustic waves (AW) with subterahertz frequencies is a problem of great importance. As attenuation of AW is reduced at liquid helium temperatures, in 1970s and later the superconducting tunnel junctions have been used to generate and also to detect both incoherent and coherent AW above 100 GHz [1-4]. The possibility of AW generation in high T_c Josephson junctions has been considered in [5, 6]. Recently, phonon radiation in the Josephson junctions as a reason for decoherence of superconducting quantum bits is also discussed [7]. Most of the papers mentioned above deal with interaction of AW with electromagnetic fields in the junction via piezoelectric coupling.

In the present paper the possibility of AW generation due to electromagnetic wave, accompanying Josephson vortex motion, is considered. The Josephson flux flow oscillator (FFO) based on unidirectional flow of vortices in a long tunnel junction is shown in Fig.1. Usually in these structures with low dumping the step-like current-voltage characteristics, so called Fiske steps (FS), are occurred. The position of the steps depends upon Josephson junction resonant frequencies defined by junction length. We'll consider AW generation in the tunneling structure, deposited on the relatively thick flat substrate with double side polished parallel surfaces. This layered structure plays the role of high overtone composite resonator for bulk AW propagating normally to the layers. The resonant generation and detection of AW are displayed on I-V curve (IVC) as a series of additional steps on the top of the ordinary FS series. The experiments carried out on Nb-AlO_x-Nb junction on Si substrate revealed the fine structure of IVC with steps separation and low differential resistance, which is pointing out AW generation [8].

II. THEORY

A. Dynamics of long Josephson junction

Consider two superconductors biased at the constant voltage and spaced by dielectric layer of thickness *d*. External magnetic field lies in the interface along the x axis (Fig. 1). This structure is deposited on a substrate of thickness *D*. The current flowing in the dielectric interlayer will be in the form $j_y = j_{s,y} + j_{n,y} = j_1 \sin\varphi(z,t) + \sigma_T e_y^0$. Here $j_{s,y}$ and $j_{n,y}$ are, respectively, Josephson tunneling current of Cooper pairs with amplitude j_1 and normal electrons current with tunnel conductivity σ_T, $\varphi(z,t)$ is the difference of SC wave functions.

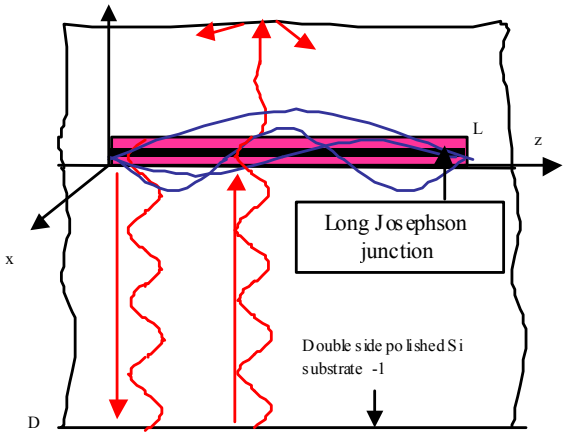


Figure 1. Schematics of the structure: 1- the Si substrate, 2 -long Josephson junction Nb-AlO_x-Nb, the black area – the AlO_x insulator film.

In linear approximation upon field amplitudes one can write down the expression for phase difference $\varphi = \varphi_0 + \varphi_1$, where $\varphi_0 = \omega t - qz$. Hereafter upper subscript (0) denotes the values integrated on dielectric layer thickness. The Josephson relations $\omega = \partial\varphi_0 / \partial t = 2eV_0 / \hbar$, $q = \partial\varphi_0 / \partial z = 2e\bar{d}H_0 / \hbar c$ connect the frequency ω and the wave vector q with constant voltage V_0 and magnetic field H_0. Here $\bar{d} = d + 2\lambda_L$ is an effective size of magnetic field localization, λ_L is a London penetration depth. The phase addition φ_1 is connected with ac voltage on the junction $v_1 = e_y^0 d$ as $\partial\varphi_1 / \partial t = (2e/\hbar)v_1$ [9].

The interaction of Josephson's electromagnetic oscillations in the junction with mechanical displacement field $\vec{u}$ caused by piezoelectric effect or electrostriction effect produces a polarization density $\vec{p}$ and contributes to an electrical displacement $d_y^0(z) = \varepsilon_d e_y^0(z) + 4\pi p_{(U)y}^0(z)$. By standard integration of Maxwell equations the generalized Josephson relationship between phase gradient and polarization can be deduced in the form

$$\frac{\partial^2\varphi_1}{\partial z^2} - \frac{1}{\tilde{c}^2}\left(\frac{\partial^2\varphi_1}{\partial t^2} + \Gamma\frac{\partial\varphi_1}{\partial t} + \frac{8\pi e d}{\hbar\varepsilon_d}\frac{\partial p_{(U)y}}{\partial t}\right) = \frac{1}{\lambda_J^2}\sin\varphi_0, \quad (1)$$

with a Swihart electromagnetic wave velocity $\tilde{c} = c\sqrt{d/(\varepsilon_d \bar{d})}$, a Josephson penetration depth $\lambda_J^2 = \hbar c^2 / (8\pi e \bar{d} j_1)$, a Maxwell relaxation frequency Γ. For typical Nb- AlO_x-Nb junction with $2\lambda_L \sim 170$ nm, d ~ 1-2 nm, $\varepsilon_d \sim 10$ the estimation gives a wave slowing-down $\tilde{c}/c \approx 0.03$. In spite of this, electromagnetic wave velocity in the Josephson junction is much greater than AW velocity.

B. Equations for elastic displacements

Effective piezoelectric constants $\tilde{\beta}_{i,kl} = \beta_{i,kl} + a_{imkl}E_m / 4\pi$ describe both normal piezzoeffect and the induced one by the electrostriction. In linear approximation the electric field in dielectric layer is uniform $E_m = V_0 / d$. Piezoelectric interaction contributions to stress tensor and electric polarization are $\Delta T_{ij} = \tilde{\beta}_{l,ij} e_l$ and $p_i = -4\pi\tilde{\beta}_{ikl}u_{kl}$ respectively. While the piezoelectric effect is absent in bulk dielectrics with inversion symmetry, it is not possible to eliminate this coupling at the superconductor-insulator boundary where this symmetry is always broken. Assuming that electromechanical coupling tensor has 4mm symmetry with axis lying in the interface, one can obtain the desired polarization as $p_2^0 = -4\pi\beta_{24}(\partial u_2^0 / \partial z)$.

After integration of the equation of motion for elastic medium along dielectric thickness we get

$$\rho\frac{\partial^2 u_2^0}{\partial t^2} = \frac{T_2(d) - T_2(0)}{d} + C_{44}\frac{\partial^2 u_2^0}{\partial z^2} + \beta_{24}\frac{\partial e_2^0}{\partial z} \quad (2)$$

The boundary conditions for elastic displacements and stresses at the dielectric layer surfaces are $T_2(0) = T_2^{(1)}(0)$, $T_2(d) = T_2^{(2)}(d)$, $u_2(0) = u_2^{(1)}(0)$, $u_2(d) = u_2^{(2)}(d)$, where the quantities with superscripts 0, (1), (2) refer to dielectric layer and surrounding media. If we assume that elastic displacements slightly vary along the thickness of dielectric, then we can easily obtain the following relationships

$$u_2(0) = u_2^0 + \frac{d}{2C_{11}}Z_1\frac{\partial u_2^{(1)}(0)}{\partial t}, \qquad u_2(d) = u_2^0 - \frac{d}{2C_{11}}Z_2\frac{\partial u_2^{(1)}(d)}{\partial t}.$$

Here $Z_{1,2}$ are acoustic impedances of the surfaces $y = 0, d$ respectively. For perfectly reflecting lower surface of the substrate ($y = -D$) $Z_1 \approx iZ_0 \tan kD$, where $k = \omega / v_l$, v_l - longitudinal AW velocity in the substrate, $Z_0 = v_l\rho_s$ - acoustic impedance of the substrate. If the upper surface of the structure (Fig. 1) is not perfectly reflecting, the impedance Z_2 contributes to acoustic losses.

C. Dispersion equation and current-voltage characteristics

For harmonic dependence of all variables on time and coordinates $\exp i(k_y y + qz - \omega t)$ one can get the dispersion relation $\omega^2 = q^2\tilde{c}^2(\varepsilon/\varepsilon^*)$ for coupled electroacoustic oscillations in the form

$$\omega^2 = q^2\tilde{c}^2\left(1 + \frac{4\pi\tilde{\beta}_{24}^{\,2}}{\varepsilon\rho\tilde{c}^2}\frac{1}{1 + i(Z_2' - Z_1')/(\omega d\rho) - q^2(v_\perp/\tilde{c})^2}\right) \quad (3)$$

with $Z_{1,2}' = Z_{1,2}/(1 \pm id\omega Z_{1,2}/C_{11})$, $v_\perp = \sqrt{C_{44}/\rho}$. The last term in the denominator of (3) may be ignored owing to inequality $v_\perp << \tilde{c}$.

According to standard procedure [9] one can find the uniform component of Josephson current. For the case of long junction the following expression can be obtained

$$j_0(V_0, H_0) = -\tilde{j}_0\Omega_n^2\,\mathrm{Im}[\omega^2 + i\bar{\Gamma}\omega - q^2\tilde{c}^2(\varepsilon/\varepsilon^*)]^{-1} \quad (4)$$

where $\tilde{j}_0 = j_1\tilde{c}^2/(2\lambda_J^2\Omega_n^2)$, $\Omega_n = \pi n v_l / D$ - the frequency of n-th mode of high overtone resonator. The value $\bar{\Gamma} = \Gamma_E + \Gamma_{AW}$ describes electromagnetic losses and nonresonance AW radiation. For the frequencies from 100 to 700 GHz mode numbers of excited AW fall in the range of (10^3 – 10^4). As shown in Fig. 2, there are several bumps of current at AW resonance frequencies $\omega = \Omega_{n+m}$ against a background of usual Lorentz form of one of the FS, corresponding to the voltages $V_{n+m} = \Omega_{n+m}\hbar/(2e)$.

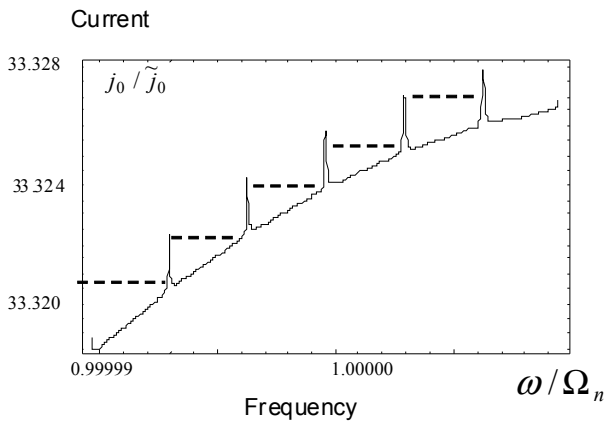


Figure 2. Calculated dependence of relative current $j_0 / \tilde{j}_0$ versus relative frequency ω / Ω_n or relative voltage $\hbar\omega/(2e\Omega_n)$ at the electromechanical coupling constant $\delta = 4\pi\tilde{\beta}_{24}^{\ 2}/(\varepsilon\rho\tilde{c}^{\,2})$=2.510^{-7}, $\overline{\Gamma} = 3\cdot 10^{-2}\Omega_n$, $\mathrm{Im}\,k / \mathrm{Re}\,k = 10^{-6}$.

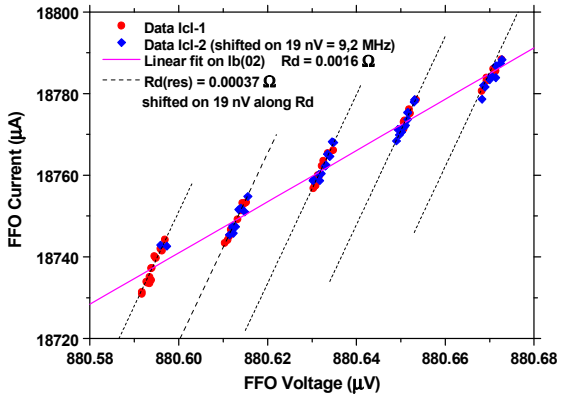


Figure 3. Reconstructed IVC of the FFO. Data marked "$I_{CL-1,-2}$" are measured at a slightly different currents I_{CL} that flow under the FFO and induce magnetic field.

III. SUPERFINE RESONANT STRUCTURE

Flux flow oscillator on the base of Nb-AlO_x-Nb superconducting tunnel junction was fabricated on double side polished Si substrate (350 μm). The thicknesses of the junction layers were correspondingly 1-2 nm for AlO_x, 200 nm and 500 nm for base and top Nb electrodes. By measuring the frequency of the FFO radiation emission [10], its IV-curve can be reconstructed with accuracy better than 1 nV [8]. From Fig. 3 one can see that the FFO IVC consists of a set of separate steps rather than being a continuous curve. The differential resistance on these steps is extremely low, R_d^B(res) = 0.00037 Ω. It is important to note that this value is considerably lower than the average value on the FS (R_d^B = 0.0016 Ω).

The voltage spacing of 19 nV, corresponding to a separation of 9 MHz, coincides with difference between the acoustic modes frequencies of the resonator structure consisting of substrate plate and the covering FFO layers

$$v_l /(2D) = (\Omega_{n+1} - \Omega_n)/(2\pi).$$

The value of voltage (frequency) spacing and separate steps with low differential resistance show a good agreement with theoretical treatment.

IV. CONCLUSION

The superfine resonant structure of the FFO IVC with step separation that coincide with acoustic modes frequency spacing may be attributed to acoustic wave generation by the FFO and propagation in thick Si substrate as in acoustic resonator. So the generation and detection of coherent acoustic waves with frequencies up to 700 GHz due to piezoelectric effect in dielectric interface in superconductor junction may be achieved.

REFERENCES

[1] W.Eisenmenger and A.H.Dayem, "Quantum generation and detection of incoherent phonons in superconductors", Phys. Rev. Lett., vol 18, pp.125-127, January1967.

[2] H.Kinder, "Spectroscopy with phonons on Al_2O_3: V^{3+} using the phonon bremsstralung of a superconducting tunneling junction", Phys. Rev. Lett., vol 28, pp.1564-1567, June 1972.

[3] P.Berberich, R.Buemann, and H.Kinder, "Monochromatic phonon generation by Josephson effect", Phys. Rev. Lett., vol 49, pp.1500-1503, November 1982.

[4] E.H.Jacobsen, "Generation and detection of ultrasonic waves beyond 100GHz", J.de Phys. Colloq., vol 33/C6, pp.C6-25--C6-31, November 1972.

[5] T. Hristov and A. Groshev, " Coupled Josephson and piezo oscillators. The effects of capacitance, noise and quality" Physica C, vol 175, pp. 600-602, May 1991.

[6] A. Groshev and T. Hristov, "Shapiro steps without RF source by coupling of Josephson and piezooscillators", Physica C, vol 160, pp. 317-319, September 1989.

[7] L. B. Ioffe, V. B. Geshkenbein, Ch. Helm, and G. Blatter, "Decoherence in superconducting quantum bits by phonon radiation", Phys. Rev. Lett, vol 93. pp. 057001-057004, July 2004.

[8] V.P. Koshelets, A.B.Ermakov, S.V.Shitov, P.V.Dmitriev, L.V.Fillipenko, A.M.Baryshev, et al. "Superfine resonant structure on IVC of long Josephson junctions and its influence on flux flow oscillator linewidth", IEEE Trans. on Appl. Supercond., vol 11, pp. 1211-1214, January 2001.

[9] A.Barone, G.Paternò, Physics and Applications of Josephson Effect, New York: John Wiley & Sons, Inc., 1982.

[10] V.P. Koshelets, S.V. Shitov, L.V. Filippenko, V.L. Vaks, J. Mygind, A.B. Baryshev, W. Luinge, N. Whyborn, " Phase Locking of 270-440 GHz Josephson Flux Flow Oscillator", Rev. of Sci. Instr., vol 71, pp. 289-293, January 2000.

Thin film quartz layer reported on silicon

B. Imbert, A. Reinhardt, T. Ricart, C. Billard, E. Defaÿ, H. Virieux, T. Jouanneau, F. Pierre, V. Delaye, P. Gergaud, E. Augendre, T. Signamarcheix, C. Deguet
CEA – LETI – MINATEC Campus
17 rue des Martyrs, 38054, Grenoble, France
imbert.bruno@cea.fr

S. Ballandras
UMR CNRS-UFC-ENSMM-UTBM
FEMTO-ST
Besançon, France

***Abstract*— For the first time, in this work, we report on the transfer of single crystal quartz layers onto silicon wafer through wafer bonding and grinding. A Film Bulk Acoustic Resonator (FBAR) made of a free standing quartz membrane vibrating on the thickness-shear mode has been achieved using this approach. Electrical measurements have been achieved for various temperature conditions. The obtained resonators exhibit fundamental resonance frequencies close to 200 MHz for the fundamental mode. The electromechanical coupling factor is about 0.4 % and the temperature coefficient of frequency is near +26 ppm/°C. These first electrical results highlight the quality piezoelectric response of the reported layer conforms to the bulk material properties.**

I. INTRODUCTION

Quartz has been among the first piezoelectric materials exploited in electronic devices such as resonators and filters [1]. Although it suffers from relatively low piezoelectric properties, yielding radiofrequency (RF) filter manufacturers to use other piezoelectric substrates such as Lithium Niobate or Lithium Tantalate for Surface Acoustic Wave (SAW) devices or Aluminum Nitride (AlN) for Bulk Acoustic Wave (BAW) devices, its very low viscoelastic losses and the existence of crystalline orientations offering thermal compensation effects made it a key element in time references. In current industry, BAW or SAW quartz resonators are fabricated from single crystal wafers and thus suffer from two major drawbacks: first, their operation frequency is limited by the ability to thin these wafers or to locally etch them to reduce their thickness; secondly, and most importantly, they are reputed to be incompatible with CMOS circuits and require lying on a specific die.

In this paper, we propose a first step to overcome these limitations by transferring a single crystal quartz film onto a silicon wafer and to use it for the fabrication of a Film Bulk Acoustic Resonator (FBAR) made of a free standing quartz membrane vibrating on its fundamental thickness-shear mode. In section II, we describe the fabrication of such resonators, using a direct bonding and thinning processes. In section III, we then describe the electrical characterization of these devices and discuss their actual properties.

II. EXPERIMENTAL PROCESS

The manufacturing of Quartz-on-Silicon stacks highlights some physical property differences. The main problem is linked to the Temperature Coefficient of Expansion (TCE) difference between quartz (13.2 ppm/°C) and silicon (2.3 ppm/°C). This induces very high mechanical constraints during thermal steps of the process.

The fabrication process used is summarized in Figure 1. In principle, it is similar to an approach previously employed to transfer thin films of lithium niobate onto silicon substrates [2]. We start with an AT-cut quartz wafer and a silicon wafer which are cleaned to remove any particle without degrading the initial surface quality so as to leave these substrates compatible with the direct bonding technology employed.

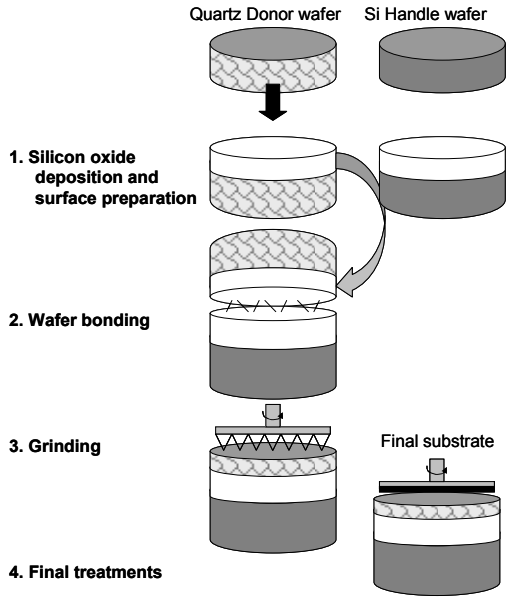


Figure 1. Thin single crystal Quartz transferred layer technique.

A PECVD silicon oxide layer is then deposited onto both wafers. The top surfaces are then prepared by combination of Chemical Mechanical Polishing (CMP) and cleaning, and bonded at room temperature thanks to hydrophilic direct

bonding. The silicon oxide / silicon oxide bonding interface has been observed by Transmission Electron Microscopy (TEM) which reveals a sharp and uniform interface (Figure 2).

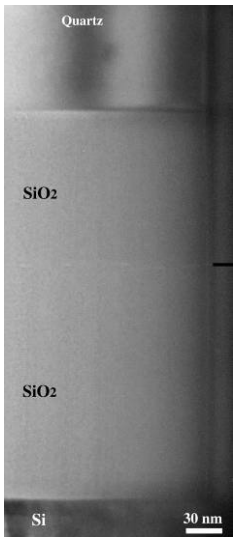


Figure 2. Silicon oxide interface TEM cross section of the quartz/silicon substrate.

Finally the quartz layer bonded onto the silicon substrate is thinned down to a thickness between 10 µm and 15 µm. The backside of the bonded quartz wafer is then grinded and finally polished, with a careful control of process parameters. Due to the non-uniformity of the CMP process, the quartz wafer cannot be thinned below 5 µm using this fabrication process.

The crystal quality of the thinned quartz film has been characterized to provide some insight into the influence of the grinding process onto the transferred film. For this, we measured both thinned and bulk quartz substrates using Rutherford Backscattering Spectrometry Channelling (RBS-c). Results shown in Figure 3 reveal that the used thinning process caused the increase of quartz crystal disorder.

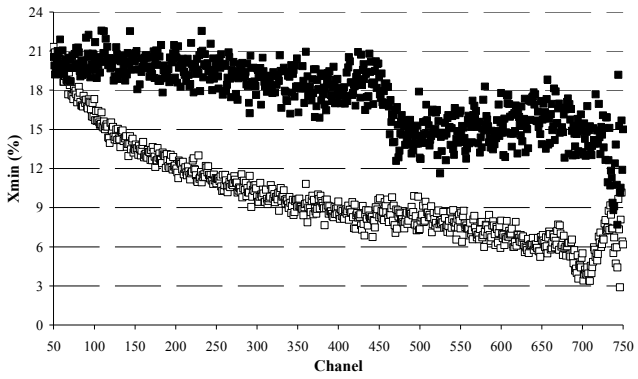


Figure 3. Crystal order variations obtained by RBS-c for bulk substrate (empty symbol) and thinned quartz layer (filled symbol).

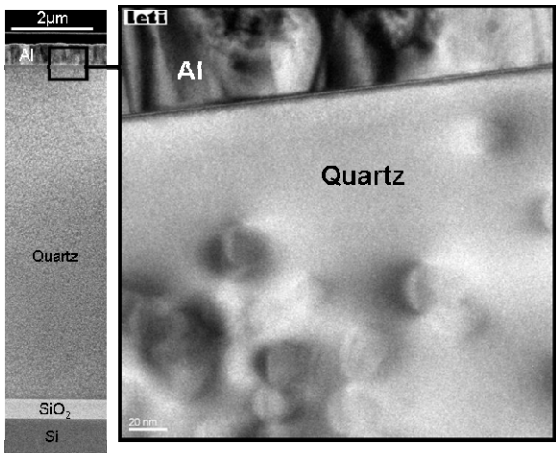


Figure 4. TEM cross section: punctual defects can be observed in the transferred quartz layer after the thinning process.

TEM observations revealed defects in the whole transferred layer, and not only at the surface of the film, as shown in Figure 4. Such defects, linked to dislocation formations, have been already observed in specific process conditions of pressure and temperature [3]. However pole figures measurements by X-ray diffraction (Figure 5) reveal same crystal orientation between initial bulk quartz wafer and quartz transferred layer.

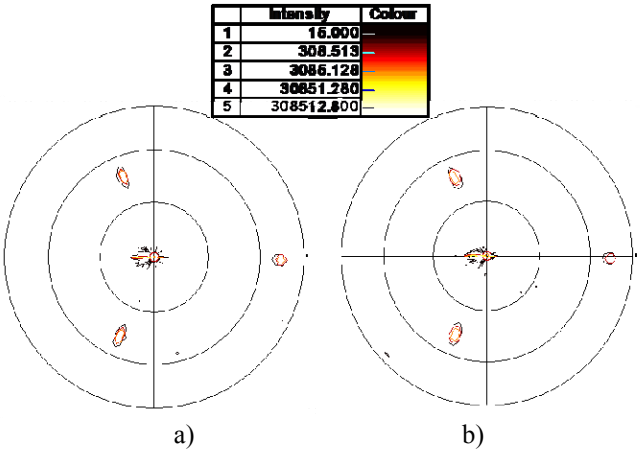


Figure 5. (101) pole figures measurements by –X-ray diffraction on bulk (a) and transferred (b) AT-cut quartz.

After the transfer of the quartz film onto silicon, we deposited top aluminum electrodes by evaporation and patterned them using standard photolithography and wet etching. The quartz film was released by a backside Deep Reactive Ion Etching (DRIE) of the Si substrate using a conventional Bosch process with the silicon oxide bonding layer acting as an etch stop. The backside of the wafer was then metalized by evaporation to allow metal to fill the bottom of the membranes and thus to provide a bottom electrode, as shown in Figure 6(a).

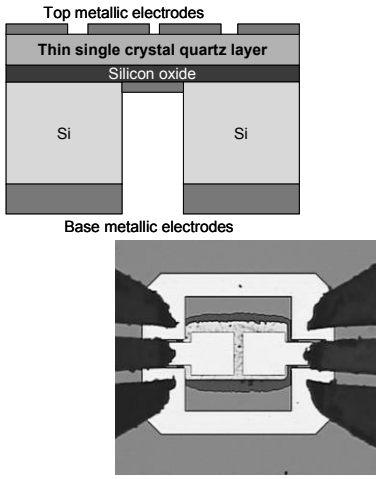


Figure 6. (a) Schematic cross section view of the final thin film quartz resonator. (b) Electrical tip on top metallic electrode of the measured device.

III. RESONATOR CHARACTERIZATION

The fabricated resonators are shown on the photograph in Figure 6(b). They consist in two resonators connected in series through the metallization of the backside of the quartz membrane. A ground ring is also located along with the two top electrodes.

These resonators were measured using two coplanar GSG probes. First, a resonator without the two square pads located on the top electrode seen in Figure 6(b) was measured, to determine the parasitic capacitance brought by the contact pads and the ground ring. This measurement was then subtracted from the measurement of the full device to suppress these parasitic influences. Although we measured this resonator as a two port device, we numerically shorted one of the two ports to retrieve the electrical behavior of a single resonator.

The resonance frequencies of the fundamental and of the third order resonance were found respectively at 156 and 464 MHz (Figure 7). Their respective electromechanical coupling factors, derived from the relative frequency spacing between resonance and antiresonance, were estimated at 0.4 and 0.03 % respectively. Quality factors are respectively of 2000 and 2500.

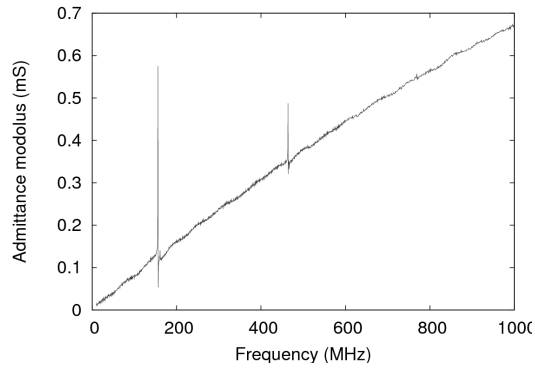


Figure 7. Electrical response of a thin film quartz resonator.

Figure 8 shows the comparison between measurements and simulations of the electrical response of a resonator based on the material stack performed using the scattering matrix method [4]. As it can be seen, an excellent agreement is obtained for the two frequencies. However, the theoretically expected electromechanical coupling factor is close to 0.75 %, being larger than what was measured. Also, the quality factor is slightly overestimated, which is attributed to lack of homogeneity of the bottom electrode layer deposited inside the membrane.

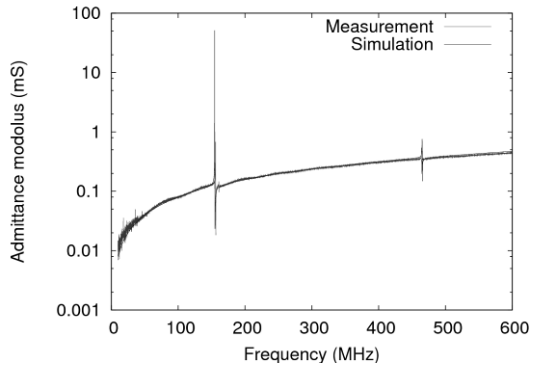


Figure 8. Comparison between the theoretical (light grey) and measured (dark grey) electrical response of a resonator.

For further investigations, we measured resonators for temperatures ranging from -25 to +125°C, by steps of 50°C. These measurements were carried out in an oven, with a narrow frequency span around the fundamental and third overtone so as to allow precise frequency steps of respectively 200 and 400 Hz. Figure 9 shows the relative frequency drifts as a function of the oven temperature: the two measured resonances exhibit almost similar linear temperature dependence (as expected). An average TCF of +26 ppm/°C was obtained. This result yields two main observations: first, its magnitude is rather large, being far from what was expected from the nearly temperature-compensated AT-cut of quartz, even if in our case metallizations are relatively thick with respect to the acoustic wavelength. Secondly, frequency shifts with temperature are positive, and can therefore not be attributed to the electrode materials or to any possible silicon remains, which would all ensure negative variations. Indeed, the theoretical temperature coefficient of frequency lies in the –3 ppm/°C, due to the silicon oxide bonding layer partially compensating metallizations. Only the silica layer may be responsible for such TCF result, but its thickness (200 nm) tends to eliminate this explanation.

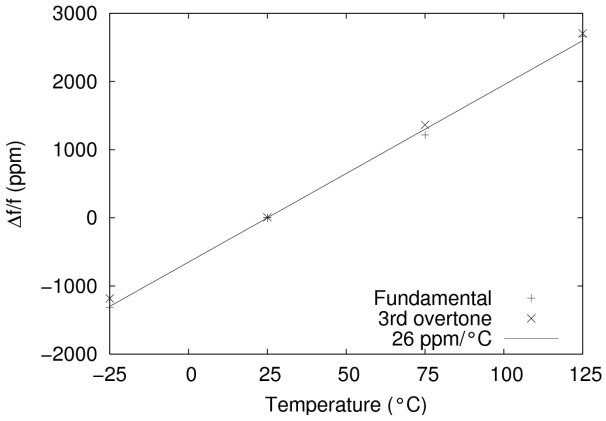


Figure 9. Relative frequency drifts vs temperature for the fundamental resonance and the third overtone.

A possible explanation for both large positive TCF and the decrease in effective electromechanical coupling factor could be the quartz amorphization during grinding. To assess the impact of the generation of an amorphous quartz layer, we modeled the resonators with an added silicon oxide layer which partially replaces the quartz film. Its impact on coupling factor and TCF is shown in Figure 10. This reveals that an equivalent layer of 4 µm of silicon dioxide could reasonably explain at the same time an effective coupling factor of 0.4 % and a TCF of +26 ppm/°C. Another explanation, implying less radical changes, could be the formation of twinned crystals in the material during grinding [5], leading to changes in the crystalline orientation of part of the material, which would manifest as changes in acoustic properties such as electromechanical coupling factor and temperature dependence, and could therefore lead to similar effects. Such a phenomenon would however have been revealed by X-Ray pole figures (Figure 5). Instead, pole figures performed on samples after bonding revealed only the existence of the crystalline orientation corresponding to the AT-cut without any other noticeable parasitic orientation.

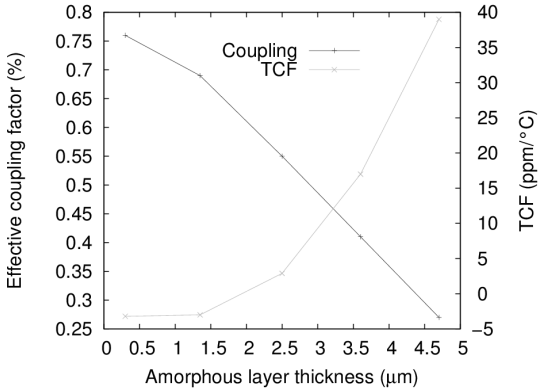


Figure 10. Impact of an amorphous (silica) layer on the performances of a quartz thin film resonator.

IV. CONCLUSION

This work reports on the transfer of a thin quartz film on a silicon substrate through wafer bonding and grinding. This was used to fabricate quartz crystal-based FBAR resonators, which were measured and characterized under various temperature conditions. The electromechanical coupling factor of 0.4 % and the temperature dependence of frequency of +26 ppm/°C, respectively lower and larger than theoretical predictions, suggest that the grinding step had somehow an impact on the piezoelectric material: possibly leading to the formation of an amorphous layer below its top surface. Future work will investigate these effects more into details and will focus on suppressing them to leave a fully defect-free quartz film.

ACKNOWLEDGMENT

This work has been performed with the help of the "Plateforme Technologique Amont de Grenoble".

REFERENCES

[1] R. G. Kinsman, "A history of crystal filters" in Proceedings of the 1999 IEEE International Frequency Control Symposium, pp. 563-570, 1999.

[2] C. Deguet, M. Pijolat, N. Blanc, B. Imbert, S. Loubriat, E. Defay, L. Clavelier, J.S. Moulet, B. Ghyselen, F. Letertre, S. Ballandras, Meet. Abstr. - Electrochem. Soc. 1002, pp. 1730, 2010.

[3] R. A. Murray, "Movement of dislocations in quartz", Forty-Fifth Annual Symposium On Frequency Control, pp.2-8, 1991.

[4] A. Reinhardt, V. Laude, R. Lanz, P. Muralt, M. Solal, S. Ballandras and W. Steichen, "Design of coupled resonator filters using admittance and scattering matrices", in Proceedings of the 2003 IEEE International Ultrasonics Symposium, pp. 1428-1431, 2003.

[5] T. Uno, "Temperature Compensation of Thickness Shear Mode Resonators Formed on Artificial Twinned Quartz Plates", IEEE International Frequency Control Symposium, pp.521-536, 1996.

Characterization of Amorphous Tantalum Oxide for Insulating Acoustic Mirrors

J. Capilla, J. Olivares, M. Clement, J. Sangrador and E. Iborra
Grupo de Microsistemas y Materiales Electrónicos.
Universidad Politécnica de Madrid
Madrid, Spain
eiborra@etsit.upm.es

A. Devos
IEMN Dpt ISEN, UMR CNRS 8520
Lille, France
Arnaud.devos@isen.fr

Abstract— **This work describes the assessment of the acoustic properties of sputtered tantalum oxide films intended as high impedance films for the acoustic isolation of bulk acoustic wave devices operating in the GHz frequency range. The films are grown by sputtering a metallic tantalum target under different oxygen and argon gas mixtures, total pressures, pulsed DC powers and substrate bias. The structural properties of the films are assessed through infrared absorption spectroscopy and X-ray diffraction measurements. Their acoustic impedance is obtained after estimating the mass density by X-ray reflectometry measurements and the longitudinal acoustic velocity by analyzing the longitudinal λ/2 resonance induced in a tantalum oxide film inserted between an acoustic reflector and an AlN-based resonator. A second measurement of the sound velocity is achieved through picosecond acoustic spectroscopy.**

I. INTRODUCTION

Bulk acoustic wave resonator technology is a recent solution for low cost, high performance bandpass filters for digital wireless communications, such as wireless networking, cellular phones, or global positioning systems. These filters are composed of several resonators electrically connected to obtain the desired filter characteristics. To obtain well-performing devices, a good acoustic isolation of the resonators is required to avoid energy losses through the substrate that reduce their quality factor. There are two typical methods to achieve a good acoustic isolation. The first one consists in creating an air gap between the resonator and the substrate; as the acoustic radiation to the air is minimal, energy losses can only take place through the supports. These structures, called film bulk acoustic resonators (FBAR), have some drawbacks, as a complex fabrication technology or a too high thermal isolation that limits the power handling. The alternative method is the use of acoustic reflectors between the resonators and the substrate; the resulting structures are called solidly mounted resonators (SMR) [1]. Bragg mirrors are typical acoustic reflectors formed by alternating a quarter wavelength-thick layers (λ/4-layers) of low and high acoustic impedance [2]. The acoustic isolation achieved (reflection coefficient) increases with the number of layers and the mismatch between the acoustic impedances of two consecutive layers. Considering that the acoustic impedance of a given material is the product of the longitudinal sound velocity and the mass density, these two material properties have to be perfectly known in order to set the thickness of the films to λ/4 to get the targeted resonant frequency and to fix the number of layers required for a specified mirror reflectance. Table I summarizes the experimental values of such material properties for the most commonly used acoustic mirrors constituents.

TABLE I. ACOUSTIC IMPEDANCE OF METALLIC AND INSULATING MATERIALS

Material	Mass density ($kg \cdot m^{-3}$)	Sound speed ($m \cdot s^{-1}$)	Acoustic impedance ($kg \cdot m^{-2} \cdot s^{-1}) \times 10^{-3}$
SiOC	1500	2400	360
Al	2700	5100	1122
SiO_2	2200	6000	1320
Si_3N_4	2200	9900	2178
AlN	3300	11000	3630
Ta_2O_5	7000 – 8200	4000-5000	2800-4100
Mo	10300	6190	6376
W	19290	5180	9992
Ir	22500	5300	11920

Silicon dioxide (SiO_2) layers are commonly used as low acoustic impedance material, owing to their easy processing and full compatibility with standard silicon technologies [3]. Other low impedance materials are aluminum (Al) [4] and silicon oxicarbide (SiOC) [5], although they are less frequently used due to the metallic character of Al and the elevated acoustic losses of SiOC. On the other hand, materials with high acoustic impedance are typically high density metals exhibiting, additionally, a moderate sound velocity, such as iridium (Ir) [6], molybdenum (Mo) [7], or tungsten (W) [8]. If these high impedance layers are substituted by insulating films, the fabrication process of BAW filters can be

978-1-61284-110-6/11/$26.00 ©2011 IEEE

significantly simplified, as there is no need to define the continuous metallic layers that short circuit the different resonators built on top of the same acoustic mirrors. Achieving insulating Bragg mirrors requires, thus, high acoustic impedance insulating layers, which are unfortunately not easily available. So far, insulating acoustic mirrors have been obtained by combining SiO_2 or SiOC layers with aluminum nitride (AlN) [9], silicon nitride (Si_3N_4) [5] or tantalum oxide (Ta_2O_5) [10, 11], although the acoustic properties of this last have not been extensively characterized. However, none of these three materials exhibits really high acoustic impedances, in comparison with that offered by high density metallic layers. The considerably large acoustic velocity of AlN combined with a low mass density (see table I) provides a moderate acoustic impedance. On the other hand, TaO_x films offer similar acoustic impedance than AlN by virtue of its high mass density, although the sound velocity appears to be considerably smaller.

In this work we have characterized the acoustic properties of TaO_x films in order to investigate whether these layers, in combination with SiO_2 films, may provide better acoustic isolation than SiO_2/AlN-based Bragg mirrors. The structure, mass density and sound velocity of sputtered TaO_x films have been assessed by different techniques, which include the fabrication and characterization of specific devices for the evaluation of the sound velocity.

II. EXPERIMENTAL

A. *Sputtering Process*

TaO_x films were deposited in a Leybold Z-550 system by pulsed-DC magnetron sputtering of a 150 mm-wide Ta target in Ar/O_2 admixtures. A pulsed-DC power of 50 kHz and duty cycle 75% was applied to the target. More than 150 films were deposited under different conditions by varying the total pulsed-DC power between 500 W and 1200 W, the total pressure in the chamber between 1.5 mTorr and 3.3 mTorr and the percentage of O_2 in the gas between 30% and 100%. An RF source was used to polarize the substrate between -50 V (with no RF power applied to it) and -150 V. The substrates were not intentionally heated during deposition.

The sputter process of a Ta metallic target in Ar/O_2 atmospheres shows a transitory behavior owing to the oxidation of the target. When bombarding a clean metallic surface at constant power with a gas containing a certain amount of oxygen, the time to stabilize the plasma impedance can reach several hours. In order to guarantee that the target had reached a stationary state before starting deposition, the target was first preconditioned at the deposition power in pure oxygen atmosphere; the voltage of the cathode attained a constant value after around 15 min, when the plasma impedance reached the steady state. The selected gas mixture was then fed into the chamber at the chosen pressure and the voltage of the cathode controlled until the new steady state was attained (15 to 30 min). At this moment the shutter was opened for film deposition during the desired time. Subsequent depositions under identical conditions required only a few minutes for stabilization. Every time the deposition conditions were modified, the complete conditioning process was restarted. The TaO_x films were deposited on (100) silicon wafers for IR transmission measurements and on acoustic reflectors (described below) built on top of silicon substrates for the assessment of the sound velocity through the frequency response of test solidly mounted resonators (SMR).

B. *Film characterization*

The structure and morphology of the TaO_x were assessed by infrared transmission measurements carried out with a Fourier transform infrared (FTIR) Nicolet 5-PC spectrophotometer, which provided the chemical bonding configuration of the films. We used non-polarized light at normal incidence over the 400 cm^{-1} to 4000 cm^{-1} range with a spectral resolution of 4 cm^{-1}. The absorption bands corresponding to the Si substrate were eliminated by subtracting the absorbance spectrum measured in bare Si. X-ray diffraction patterns were measured in conventional Bragg-Brentano geometry in a Supratech XPert MRD diffractometer between $2\theta = 10º$ and $2\theta = 80º$.

The assessment of the acoustic properties of TaO_x films requires very precise measurements of their mass density, acoustic velocity and thickness. The thickness of the TaO_x films and that of the other films involved in test devices were measured with a Veeco Dektak 150 profilometer. The density of the TaO_x layers was assessed by X-ray reflectometry (XRR) measurements using the same diffractometer operated at grazing incidence between 0.05º and 2.5º. Finally, the longitudinal sound velocity was assessed using two different techniques. First, the sound speed was measured in different test structures by the picosecond ultrasonic technique [12]. A second assessment was deduced from the frequency response of the electrical impedance of specific SMR test devices containing the TaO_x layers under study, measured with an Agilent PNA N5230A network analyzer. The TaO_x films, with thickness ranging between 150 nm and 1000 nm, were inserted between a well characterized acoustic reflector composed of four alternated porous-SiO_2 and Mo films and a Mo/AlN/Mo resonator, as shown in Fig. 1. The excitation of the resonator induced a $\lambda/2$ resonance in the TaO_x film. The fitting of the experimental data with Mason's model allowed deriving the longitudinal acoustic velocity, provided that the material properties and the thickness of all the layers involved in the device were known.

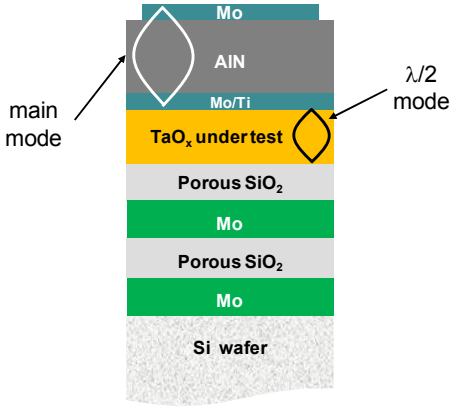


Figure 1. Sketch of the test devices used for assessing the sound velocity in sputtered TaO_x films. The main resonant mode in the piezoelectric stack is represented in white, and the $\lambda/2$ mode in the TaO_x film in black.

III. RESULTS AND DISCUSSION

A. *Structure and Composition of Sputtered Tantalum Oxides*

Fig. 2 shows the IR transmission spectra of three TaO_x of different thickness. The bands at 950 cm^{-1} and 650 cm^{-1} have been associated to the Ta-O-Ta stretching mode and the O≡Ta bond, respectively, of amorphous TaO_x films [13]. The main absorption band at 650 cm^{-1} has been reported to shift to lower wavenumbers (down to 510 cm^{-1}) in polycrystalline films. In our spectra this band appears as a shoulder in the 530 cm^{-1} region. The spectra suggest that our TaO_x films are mainly amorphous, but contain a certain amount of microcrystals. Considering the amplitude ratio of the bands at 530 cm^{-1} and 650 cm^{-1} as a qualitative estimation of the amount of microcrystals embedded in the amorphous matrix, a slight increase is observed with the thickness of the films. However, in films of similar thickness, we have not observed any clear dependence between the amount of microcrystals and the deposition conditions.

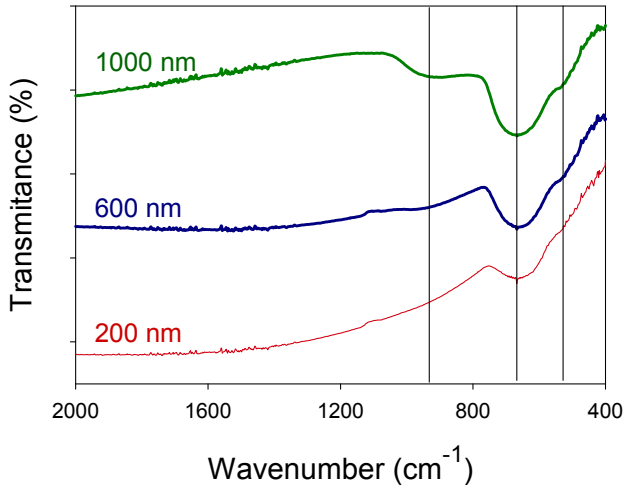


Figure 2. Infrared transmision spectra between 2000 cm^{-1} and 400 cm^{-1} of three TaO_x films of different thickness.

X-ray diffraction measurements confirm this behavior. Fig. 3 shows the XRD patterns of the three TaO_x films of different thickness analyzed by FTIR. It is important to note that, among all the films deposited, only a few exhibited clear XRD peaks corresponding to the orthorhombic β-phase of Ta_2O_5, whereas most of the films deposited in similar conditions were clearly amorphous. Peaks at 2θ=26º, 52º, 48º, 56º and 62º corresponding to the (210), (200), (3121), (200), and (201) directions of the Ta_2O_5 orthorhombic β-phase appear in the patterns corresponding to the thicker films, whereas no traces of these peaks are visible in the thinner one, which indicates its amorphous character. Additional peaks corresponding to AlN and Mo are also observed, as the films are characterized after depositing the piezoelectric stack for electric characterization. The data suggest again that thicker films tend to contain some nanocrystals whose grain size, derived from the width of the TaO_x peaks, is around 10 nm. However, we have not found any clear trend between the microcrystalline character of the samples (or any other film property) and the deposition conditions.

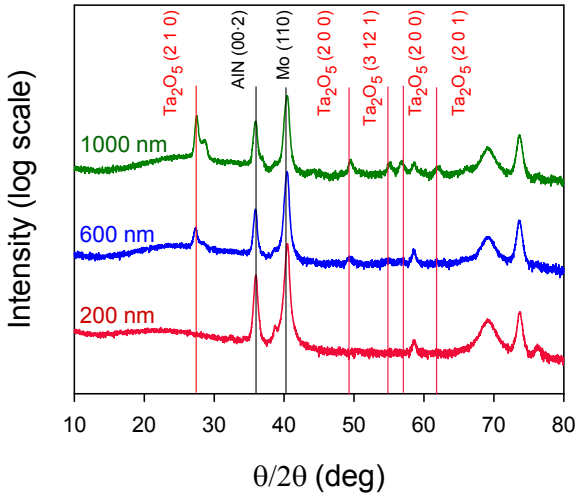


Figure 3. XRD patterns of TaO_x films of different thickness.

B. *Density assessment*

The mass density ρ of the TaO_x films was measured by XRR, which provides a direct assessment of ρ independently of the thickness. X-rays impinging on smooth TaO_x surfaces at grazing incidence below a critical angle θ_c are totally reflected. When the incident angle exceeds θ_c, the X-rays penetrate into the film and the intensity of the reflected beam drops sharply. According to [14], the ρ can be obtained through the measured θ_c. The XRR patterns of the densest and less dense TaO_x films are shown in Fig. 4; the patterns of all the remaining films lie between these two traces. The experimental data were fitted using the commercial software RCREFSimW from IHP [15] which provided an accurate value of ρ. The values obtained were 8100 kg/m^3 and 7300 kg/m^3, respectively, revealing that even the denser films do not reach the nominal mass density of Ta_2O_5 (8200 kg/m^3). These low values of ρ could be associated either to a lack of stoichiometry (films containing oxygen in excess) or more likely to a non-packed structure associated to the amorphous structure of most of the films. The high value of the refractive index (2.26) measured at a wavelength of 400 nm confirms this hypothesis.

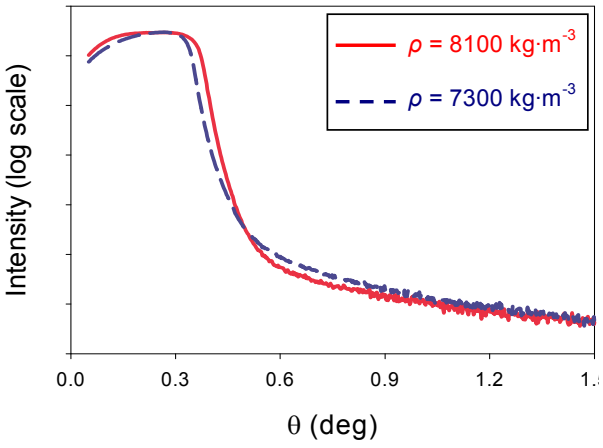


Figure 4. X-ray reflection curves of two 200 nm-thick TaOx films grown on acoustic reflectors. The red solid line corresponds to the film of highest mass density and the blue dashed line to the film of lowest mass density.

C. Sound Velocity Assessment

1) Assessment through test SMRs: The first way to assess the longitudinal sound velocity in the TaO_x films consisted in the measurement of the λ/2 resonant mode induced in these layers when they were placed in intimate contact with a piezoelectric resonator, as shown in Fig. 1. Taking into account that the wavelength of the stationary wave of this mode is ideally twice the thickness of the layer, the sound velocity v_s can be estimated, in a first approach, from the resonant frequency f_0 of the $\lambda/2_{TaOx}$ and the thickness t of the layer through:

$$v_S = f_0 \times 2 \cdot t \qquad (1)$$

However, as the resonant frequency also depends on the weight of the different layers involved in the test SMR, reliable values of the longitudinal sound velocity were only achieved if the whole experimental frequency response of the test resonator was fitted with Mason's physical model [16].

The variation with frequency of the modulus of the electrical impedance of two test SMR containing TaO_x layers of different thickness is shown in Fig. 5. Two resonances appear in each spectrum at different frequencies. The low frequency modes correspond to the main resonance in the AlN layer whereas the high frequency modes are associated with the λ/2 mode in the TaO_x film. Fig. 5 shows that upon decreasing the thickness of the TaO_x film, the two main resonances shift to higher frequencies, because the weight of this layer decreases and influences the travel of the wave in the whole structure.

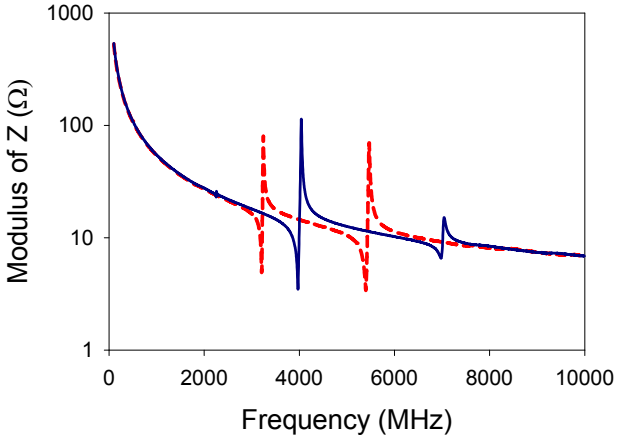


Figure 5. Modulus of the impedance versus frequency of two identical test SMRs containing 180 nm-thick (blue solid line) and 400 nm-thick (red dashed line) TaO_x films.

Fitting the experimental data to Mason's model requires precise values of the material properties of all layers involved in the test structures and of all the dimensions of the devices (thickness of layers and areas of the resonators). The area of the top electrode, which defined the areas of the SMRs, was measured carefully with an optical microscope to avoid errors due to the undercutting during the electrode definition. The properties of the different materials used for fitting the experimental data are listed in Table II. Data of Mo and AlN were taken from the literature and those of porous SiO_2 had been carefully characterized in a previous work [17]. Besides, all these layers have been routinely tested in other devices that do not contain TaO_x films.

TABLE II. MATERIAL CONSTANTS FOR SMR MATERIALS

Material	v_S (m/s)	ρ (kg/m^{-3})
Mo	5500	10500
SiO_2	5350	1900
AlN	11000	3300

The fitting of the experimental data of Fig. 5 allowed deducing the value of the longitudinal sound velocity of the different TaO_x layers, which are presented in Fig. 6. The dispersion observed is very likely due to the structural variations of the TaO_x films revealed by XRD and FTIR.

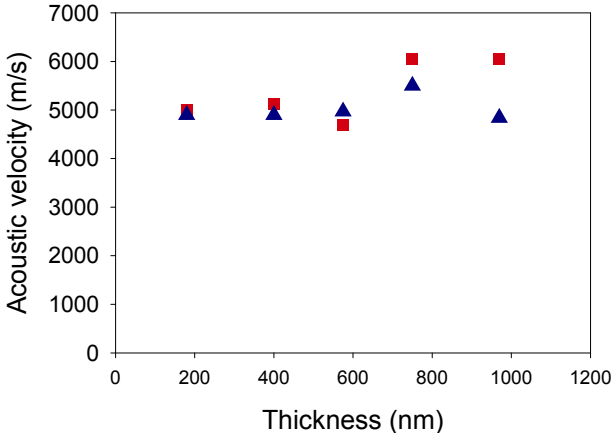


Figure 6. Sound velocity of a set of TaO_x of different thicknesses grown under identical conditions. BAW devices measurements (▲) and picosecond ultrasonics measurements (■).

2) Assessment through picosecond measurements: An alternative way of assessing the sound velocity is by the picosecond ultrasonic technique [18]. In this technique, a pump laser pulse heats the surface of the metallic layer producing a local dilatation. This generates a very high frequency longitudinal strain wave that propagates through the film to the surface at the longitudinal sound velocity. Two different methods were used to detect acoustic waves.

In the first method we detect the arrival of the acoustic pulse in the substrate. Assuming the thickness of the layer is known, the longitudinal acoustic velocity is obtained by dividing the thickness d by the time of transit τ of the echo:

$$v_s = \frac{d}{\tau} \qquad (2)$$

In the second method and as the strain wave propagates through the transparent film, it causes local variations of the refractive index. These changes can be detected by a second light beam, called the beam probe. The beam probe reflected at the surface of the films interferes with the beam reflected at the front of the traveling strain wave, causing the intensity of

the signal in the detector to oscillate with time. The period T of the so-called Brillouin oscillations is directly related to the longitudinal sound velocity through the equation,

$$v_s = \frac{\lambda}{2nT\cos\theta} \quad (3)$$

where λ is the probe wavelength, n the refractive index of the dielectric layer at λ and θ the incident angle [19,20].

The measurements were carried out on test structures composed of TaO_x layers of different thicknesses deposited on top of bare silicon substrates. A thin Al layer was deposited on the TaO_x film to allow the pumping pulse to be absorbed. The picosecond acoustic response of a representative sample is presented in Fig. 7, showing the first echo of the acoustic pulse at 133 ps and the period of the Brillouin oscillations of 20.3 ps. The value of the refractive index (2.26) at the laser probe wavelength (400 nm) was measured by ellipsometry. The acoustic velocity of this particular sample calculated by the two methods is 4500 m/s.

The sound velocities obtained by this technique are depicted in Fig. 6 together with those obtained by the BAW resonator fitting method. The results suggest that the sound velocity tends to increase with the thickness of the films, which would be in agreement with the nanocrystallites observed in the thicker films by XRD. They also reveal some discrepancies between the values achieved with the two techniques, especially for the thicker films. This can be due to the fact that the two techniques use different layered structures; whereas the TaO_x films assessed through the BAW test devices are grown on SiO_2 layers, those assessed by the picosecond measurements are grown on crystalline Si substrates. The structural differences of the underlying substrates may have affected the crystalline properties of the TaO_x films, and hence their sound velocities. Other picosecond measurements performed in TaO_x films deposited on molybdenum films yield sound velocities slightly greater than those achieved on Si substrates, which confirms the influence of the substrate on the TaO_x deposition.

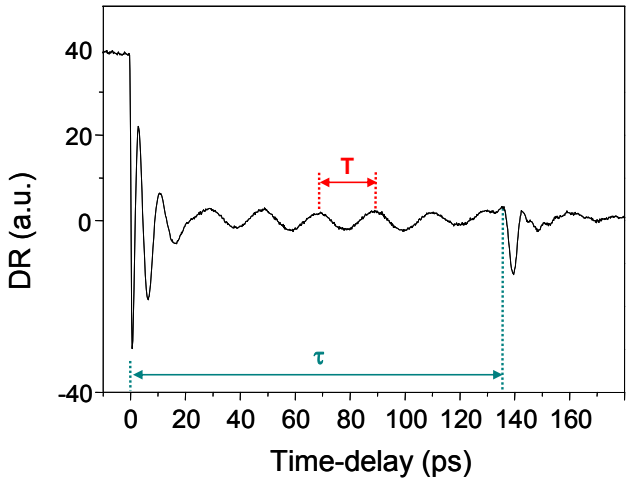


Figure 7. Picosecond acoustic response of a Al/TaO_x/Si structure showing the Brillouin oscillations of period T and the time of transit of the echo τ.

In order to investigate whether these TaO_x films may contribute to optimize the performance of insulating acoustic reflectors, we have simulated the acoustic transmittance of mirrors composed of nine alternated layers of porous SiO_2 and TaO_x films and compared it with that of conventional AlN-based mirrors. The simulations, shown in Fig. 7, suggest that even for the lowest values of the mass density and acoustic velocity, TaO_x-based mirrors exhibit lower transmission coefficients in a wider frequency band that AlN-based reflectors. Additionally, the dispersion of the values of the mass density and acoustic velocity does not affect significantly the performance of the reflectors in the middle of the band.

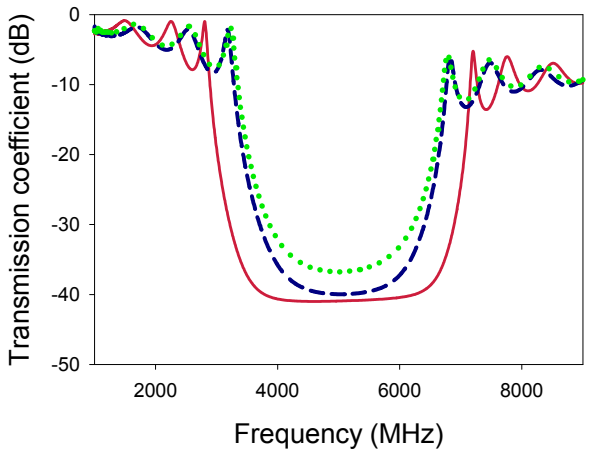


Figure 8. Simulated transmision coefficient of acoustic reflectors composed of nine porous SiO_2/TaO_x layers using a) the greatest values of ρ and v_s (red solid line) and b) the smallest values of ρ and v_S (blue dashed line) compared to that of conventional SiO_2/AlN reflectors (green dotted line).

Preliminary results on the performance of AlN-based SMRs using insulating TaO_x/SiO_2 acoustic reflectors yield quality factors up to 500 at 8 GHz.

IV. Conclusions

In this paper we assessed the acoustic properties of TaO_x films deposited at room temperature by reactive pulsed-DC sputtering. The sound velocity derived from picosecond measurements and the frequency response of BAW test devices slightly depends on the thickness of the films, which has been associated to their different structure. Thinner films appear to be amorphous whereas films 600 nm-thick and above tend to exhibit some nanocrystallites. The typical sound speed in films with thickness below 400 nm is around 5000 m/s. The mass density of the TaO_x films was assessed by X-ray reflectometry; the average value is 7600 kg/m^3, around 93% of the nominal density of Ta_2O_5 films. The typical acoustic impedance of the films is 38×10^6 kg·m^{-2}·s^{-1}, which is greater than that of AlN. Combined with a low acoustic impedance material, such as porous SiO_2, TaO_x films allow to fabricate all dielectric acoustic reflectors with good performances for GHz filtering applications.

ACKNOWLEDGMENT

This work has been supported by the Ministerio de Ciencia e Innovación of Spain through project MAT2010-18933, by the Comunidad de Madrid and the UPM through the IV PRICIT programme and by the European Union through the European Regional Development Fund (FEDER).

REFERENCES

[1] R. Ruby, "Review and Comparison of Bulk Acoustic Wave FBAR, SMR Technology", in *Proc. 2007 IEEE International Ultrasonics Symposium*, New York 2007, 1029-10407.

[2] K.M. Lakin, K.T. McCarron, R.E. Rose, "Solidly Mounted Resonators and Filters", in *Proc. 1995 IEEE International Ultrasonics Symposium*, Seattle 1995, pp. 905-908.

[3] S. Marksteiner, J. Kaitila, G.G. Fattinger, and R. Aigner, "Optimization of Acoustic Mirrors for Solidly Mounted BAW Resonators", *in Proc. 2005 IEEE International Ultrasonics Symposium*, Rotterdam 2005, vol. 1, pp 329-332.

[4] J. Enlund, D. Martin, V. Yantchev and I. Katardjiev, "Solidly mounted thin film electro-acoustic resonator utilizing a conductive Bragg reflector", *Sensors and Actuators A 141* (2008) pp. 598–602.

[5] A. Reinhardt, N. Buffet, A.Shirakawa, J. B. David, G. Parat, M. Aid, S. Joblot, P. Ancey, "Simulation of BAW resonators frequency adjustment", in *Proc. 2007 IEEE International Ultrasonics Symposium*, New York 2007, pp. 1444-1447.

[6] J. Olivares, E. Wegmann, M. Clement, J. Capilla, E. Iborra, and J. Sangrador, "Wide Bandwidth Bragg Mirrors for Multi-band Filter Chips", in *Proc. 2009 IEEE International Ultrasonics Symposium*, Rome 2009, pp. 2119-2122.

[7] C.-J. Chung, Y.-C. Chen, C.-C. Cheng, K.-S. Kao, "Fabrication and frequency response of solidly mounted resonators with 1/4λ mode configuration", Thin Solid Films 516 (2008) pp. 5277–5281

[8] R. Aigner, J. Kaitila, J. Ella, L. Elbrecht, W. Nessler, M. Handtmann, T.-R. Herzog, and S. Marksteiner, "Bulk-Acoustic-Wave Filters: Performance Optimization and Volume Manufacturing", in *Proc. 2003 IEEE International MTT-S Digest*, Philadelphia 2003, pp. 2001-2004.

[9] S.H. Kim, J.H. Kim, J.K. Lee, S.H. Lee, K. H. Yoon, "Bragg reflector thin film resonator using aluminium nitride deposited by rf sputtering", in *Proc. 2000 IEEE Microwave Conference Asia-Pacific*, Sidney 2000, pp. 1535–1538.

[10] J-W. Lobeek, R. Strijbos, A. Jansman, N. X. Li, A. B. Smolders, and N. Puisford "High-Q BAW Resonator on Pt/Ta_2O_5/SiO_2-based Reflector Stack", in Proc. 2007 *IEEE International MTT-S Digest*, Hawaii 2007, pp. 2047- 2050.

[11] H.P.Loebl, C.Metzmacher, R.F.Milsom, R.Mauczok, W.Brand , P.Lok, A.Tuinhout, F. Vanhelmont, "Narrow Band Bulk Acoustic Wave Filters", in *Proc. 2004 IEEE International Ultrasonics Symposium*, Montréal 2004, pp. 411-415.

[12] P. Emery, G. Caruyer, R. Vélard, N. Casanova, P. Ancey, and A. Devos, "Picosecond Ultrasonics: An original tool for physical characterization of Bragg reflectors in bulk acoustic wave resonators" in *Proc. 2005 IEEE International Ultrasonics Symposium*, Rotterdam 2005, pp. 906-909.

[13] A.P. Huang, S.L. Xu, M.K. Zhu, B. Wang, H. Yan, "Crystallization control of sputtered Ta_2O_5 thin films by substrate bias", *Appl. Phys. Lett*. vol 83, nº 16, pp. 3278-3280 (2003).

[14] A. C. Ferrari, A. Libassi, B. K. Tanner, V. Stolojan, J. Yuan, L. M. Brown, S. E. Rodil, B. Kleinsorge, J. Robertson "Density, sp3 fraction, and cross-sectional structure of amorphous carbon films determined by x-ray reflectivity and electron energy-loss spectroscopy", *Phys. Rev. B 62*, pp. 11089-11103

[15] P. Zaumseil, Reflectivity software RCRefSimW, Version 1.09 Innovations for high performance Microelectronics. Frankfurt Germany. (http://www.ihp-microelectronics.com)

[16] J.F. Rosenbaum, *Bulk Acoustic Wave Theory and Devices*, Artech House, Boston 1988.

[17] J. Olivares, E. Wegmann, J. Capilla, E. Iborra, M. Clement, L. Vergara, and R. Aigner, "Sputtered SiO_2 as Low Acoustic Impedance Material for Bragg Mirror Fabrication in BAW Resonators", *IEEE Trans. Ultrason. Ferr. Freq. Control.,* vol. 57 nº 1, January 2010, pp. 23-29.

[18] P. Emery, A. Devos, P Ancey "Picosecond Ultrasonics: the preferred tool for BAW characterization" *Proc. 2008 IEEE International Ultrasonics Symposium*, Beijin 2008, pp. 2205-2208.

[19] A. Devos and R. Côte, "Strong oscillations detected by picosecond ultrasonics in silicon: Evidence for an electronic-structure effect", *Phys. Rev. B 70*, 125208 (2004).

[20] A. Devos and R. Côte, "Refractive index, sound velocity and thickness of thin transparent films from multiple angles picosecond ultrasonics", *Rev. Sci. Instrum. 76*, 053906 (2005).

Introducing porous silicon as a getter using the self aligned maskless process to enhance the quality factor of packaged MEMS resonators

Wajihuddin Mohammad, Chester Wilson and Ville Kaajakari*
Institute for Micromanufacturing (IfM), Louisiana Tech University
Ruston, LA 71270
Email: wmo003@latech.edu

***Abstract*—Vacuum encapsulated MEMS resonators are used in frequency references and gyroscopes. We present the use of porous silicon as a getter material for MEMS devices. Two types of devices were fabricated using the electrochemical etching and compared for quality factor. One type was with a cavity in the substrate of an SOI die, which helps in reducing the parasitic capacitance and the air damping. The other was with cavity and porous silicon (getter) in the substrate. The measured resonant profiles show that the device with the getter material has 2x higher quality factor than the one without the getter.**

I. INTRODUCTION

Encapsulation of MEMS devices protects structures from external environment and maintains low package pressures. As the typical mass of the resonating structure is in the order of 10^{-13} to 10^{-9} kg, any micro/nano dust particle could significantly increase the mass and cause a major drift in the frequency [1]. Several encapsulation techniques have been discussed in [2], such as anodic [3], fusion, eutectic, solder, adhesive and glass frit. Commercially available Bosch airbag and yaw rate sensors are one example of mass produced wafer bonded devices [4]. Hermetic encapsulation is needed in different devices such as accelerometers [5], RF switches for controlled air damping, sensitive gears and gyroscopes for moisture control of drag and friction, IR MEMS for vacuum insulation [6]. The main challenge in MEMS encapsulation is achieving the low pressure.

Getters are widely used to lower the effect of out gassing and maintain the required vacuum. The beneficial effects of getter material for MEMS resonators are well documented [6, 7]. Marco *et al* [6] presented a method to pattern the getter film named as High Porosity Thick Film (HPTF). The conventional method of using the getter films in ceramic and metallic packages uses the mixture of Ti and Zr getter alloy deposited on the metallic substrates. The getter has high porosity that absorbs H_2O, O_2, CO, CO_2, H_2 and N_2. The drawbacks of this approach are that *1.* An external layer needs to be introduced *2.* The process needs a mask alignment for pattering.

*Currently at VTI Technologies, Pasadena, CA.

We have presented a self aligned, mask-less electro-chemical etching process to etch cavities in the substrate of commercially available SOI wafers [8]. It was shown that with the introduction of cavity in the substrate of SOI, the quality factor increases by two times as compared to the quality factor of a device with no cavity. Using the same technology we present here the introduction of porous silicon (PS) in the substrate of SOI wafers, after the cavity is introduced, so that the porous silicon can be used as a getter material.

II. POROUS SILICON AS GETTER MATERIAL

Porous silicon has large internal surface area (in our case 200 m^2/cm^3). This surface is highly reactive silicon that can absorb gases. For example, the porous silicon layer is easily oxidized to capture oxygen gas as shown in Figure 1. The advantages of porous silicon getter along with cavity are: *1.* The cavity and porous layer reduces the parasitic capacitance and air damping that helps resonators perform better *2.* The porous layer needs no mask and is self aligned *3.* The porous getter material needs no external heat to be activated.

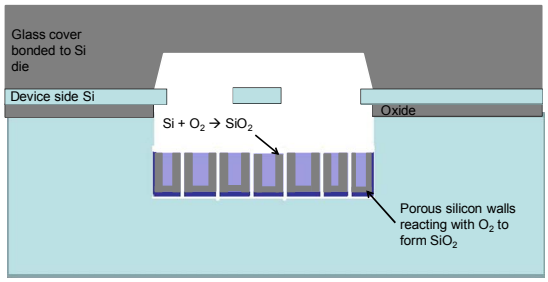


Figure 1. Schematic of the bonded SOI die that shows the reaction of silicon with oxygen resulting in silicon dioxide

III. EXPERIMENTAL SETUP AND RESULTS

Two types of encapsulated devices were fabricated for the comparison of non-getter and getter devices. First is with only a cavity in the substrate, and the second has a thick layer of porous silicon (PS) introduced in the substrate that acts as a getter. Our self-aligned process is shown in Figure 2. The cavity is formed by electrochemical etching of the substrate through the openings in the structural wafer. First, the buried oxide layer is removed with the concentrated HF solution. Next, a 10-µm thick PS layer is formed by electrochemical etching in diluted HF solution (10% HF, current density of 3.0 mA/cm^2 for 20minutes). This electrochemical etching works on the areas where the oxide has been removed. The sacrificial porous silicon layer is removed by a KOH etch. A 40 second KOH etch at room temperature is sufficient for removing the porous silicon layer but crystalline silicon is etched by less than 5 nm. Figure 2C shows the schematic of the die with only cavity in the substrate. Figure 3 shows the SEM image of the die with only cavity. To introduce the porous silicon in the substrate, after the KOH etch step, another cycle of electrochemical etching with the same process parameters mentioned above is performed. Figure 2 shows the schematic of the device with cavity and porous silicon getter in it. Figure 4 shows the SEM of the die with cavity and porous silicon. The dies were anodic bonded with glass for comparison of packaging pressure. Figure 5 shows the SEM of a bonded die that shows the glass and silicon bonded sample. A cavity of 100-150µm is introduced in the capping die. The cavity in the glass cap increases the volume in the encapsulated MEMS die and hence decreases the air damping on the resonator. Figure 6 shows the tuning fork resonator used to test the use of porous silicon as a getter.

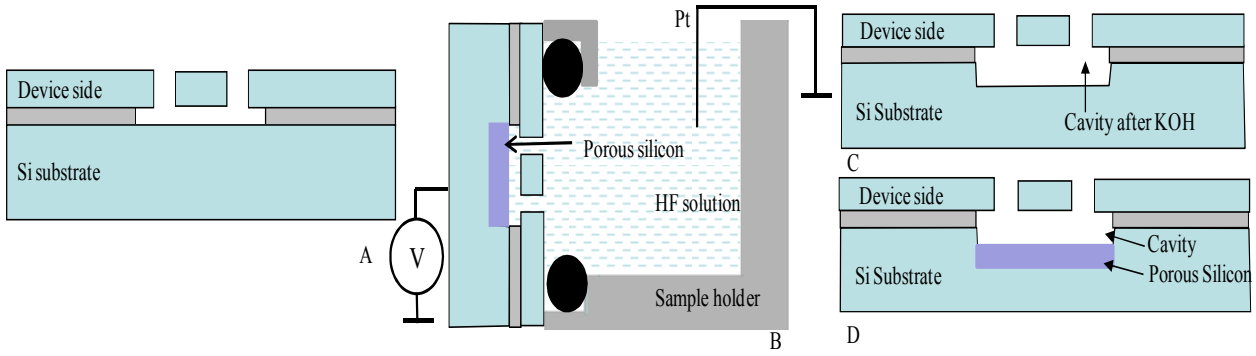


Figure 2. A) Original SOI die with etched buried oxide layer B) release, sacrificial porous silicon for cavity and porous silicon as getter C) SOI die with sacrificial porous silicon etched for only cavity D) SOI die with cavity and porous silicon as a getter

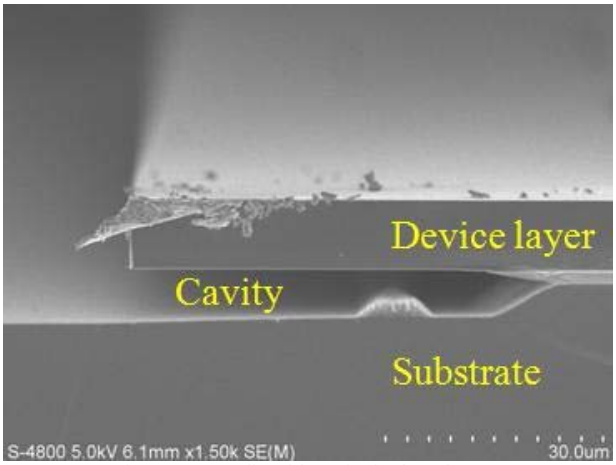


Figure 3. SOI die with cavity in the substrate

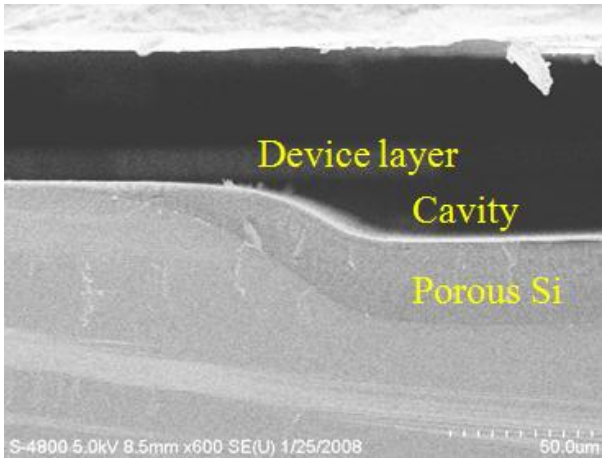


Figure 4. SOI die with cavity and porous silicon in the substrate

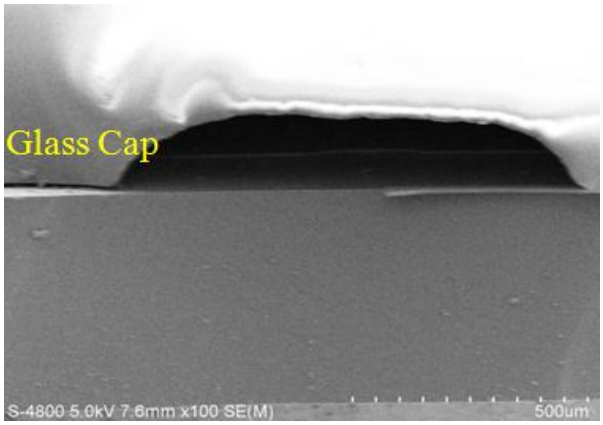


Figure 5. Anodic bonded SOI and glass die with cavity in the glass for extra volume in sample.

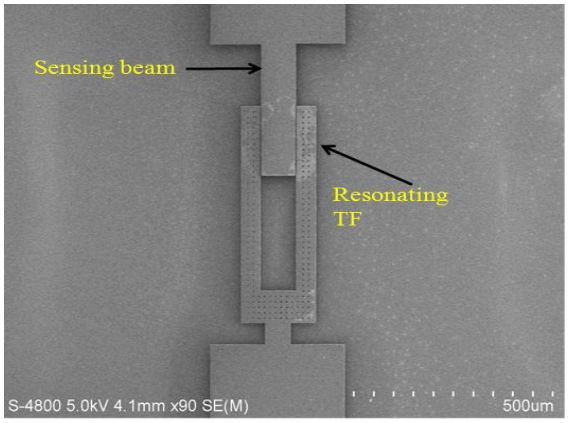


Figure 6. The tuning fork resonator used to test the use of porous silicon.

IV. RESULTS OBTAINED

The encapsulated devices were then tested for the resonant profiles with a network analyzer. Figure 7 shows the transmission response of resonators with the resonance frequency of f_o =243.54Khz. The device with only cavity had a quality factor of 8,000 and device with cavity and getter had a quality factor of 14,000. Hence the device with cavity and getter had a quality factor approximately two times higher than the device with only the cavity. Multiple samples were fabricated to confirm the performance of the porous silicon getter.

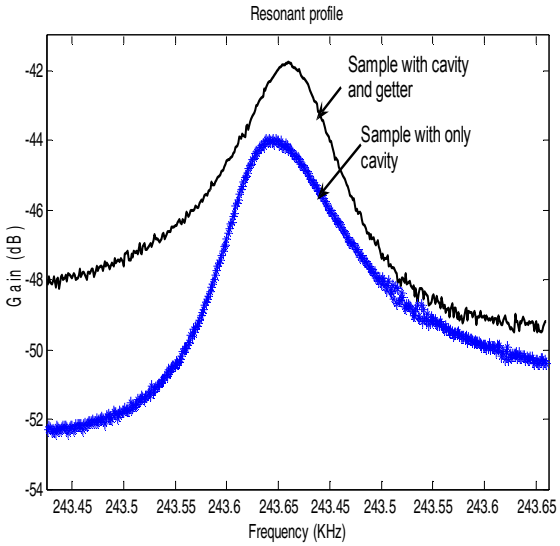


Figure 7. Resonant profiles of the devices with cavity and porous silicon getter, and with only cavity in the substrate.

Table I summarizes the quality factors obtained from these samples. It could be seen that the quality factor of the devices with getter was always better than the devices without getter.

TABLE I. SUMMARY OF QUALITY FACTORS OBTAINED FROM DIFFERENT DEVICES

	Q of devices with only cavity	**Q of devices with cavity and porous silicon**
Batch 1	7000	13520
Batch 2	5873	16083
Batch 3	2410	6701

V. PRESSURE IN THE BONDED SAMPLES

To estimate the pressure inside the bonded MEMS device, an uncapped die was tested in vacuum at different pressures. The pressure in the cavity was back calculated to be 5 Torr for the sample with only cavity and 1 Torr for the sample with cavity and getter. As the process parameters remained same in bonding both type of devices, the low pressure in the sample with the cavity and the getter is due to the getter in the device.

VI. THE NUMBER OF GAS MOLECULES ABSORBED BY PS

The SOI die size is 8mm×8mm and the area where porous silicon can be formed is 3.3 $\times 10^{-7}$ m^2. The volume of the porous silicon (V_{PoSi}) in the die is 3.3×10^{-11} m^3. The BET (Brunauer, Emmett and Teller) [9] surface area (A_{BET}) for the electrochemical etching parameters used is 200 m^2/cm^3 [10]. The surface area of the porous silicon (A_{PoSi}) is

$$A_{PoSi} = A_{BET} \times V_{PoSi} = 6.6\ m^2. \quad (1)$$

The number of active sites (N_{AS}) present on the surface of porous silicon that can react to the gases present are

$$N_{AS} = A_{PoSi} / CA_{Si} = 22\times10^{18}. \quad (2)$$

In Equation (2), the CA_{Si} represent the area of the crystalline Si. The number of moles of gas molecules (n)

present in the bonded devices can be calculated using the ideal gas law equation

$$n = \frac{PV}{RT}. \quad (3)$$

Table II summarizes the values of pressure (P), volume (V), gas constant (R), and temperature (T) used to calculate the number of moles of gas present in the sample with only cavity and in a sample with cavity and getter.

TABLE II. VALUES USED FOR COMPUTING THE NUMBER OF MOLES OF GAS MOLECULES

	Only cavity	**Cavity and getter**
P	5 Torr = 666 Pa	1 Torr = 133 Pa
V	33×10^{-12} m^3	33×10^{-12} m^3
R	8.314JK^{-1} mol^{-1}	8.314JK^{-1} mol^{-1}
T	300 K	300 K
n	8.8×10^{-11}	1.8×10^{-11}

The difference between the number of moles of gas present in the sample with only cavity ($n1$) and the number of moles of gas present in the sample with cavity and getter ($n2$) gives us the number of moles of gas molecules absorbed by the porous silicon getter which is

$$\Delta n = n1 - n2 = 7.0\times10^{-11} \text{ mol.} \quad (4)$$

Hence the number of molecules of gas molecules absorbed is estimated to be 4.3×10^{-13}.

VII. DISCUSSION

The benefits of cavity in the substrate of the MEMS devices are:

1. The parasitic capacitance to the substrate is reduced. The parasitic capacitance C_P is related to the device-substrate gap d by

$$C_P = \frac{\varepsilon_o A_{el}}{d}, \quad (5)$$

Where A_{el} is the area and ε_0 is the permittivity of free space. Hence, with the increase in the depth of the cavity, the parasitic capacitance C_P shown in Figure 7 is decreased. According to Eq. (2), the increase in distance of movable structure to the substrate from 2 to 10 μm reduces the parasitic capacitance by a factor of five.

2. The air damping is reduced. The main damping mechanisms are the squeeze film damping between the moving resonator and the fixed electrode caused by the air flowing from the gap to empty space above the device and the shear damping between the device and the substrate. By providing additional space under the device, the travel length for the air in between the electrodes is effectively reduced by half with air escaping to above and below the structure. In theory, this reduces the air damping by a factor of six. If shear damping (Couette damping) dominates, the quality factor is related to the distance between the movable structure and substrate [11, 12] by

$$Q = \frac{d\sqrt{mk}}{\mu A}, \quad (6)$$

where d is the distance between the devise and the substrate, m is the resonator mass, k is the spring constant, μ is the viscosity, A is the device area. According to (6), we expect 5x improvements in the quality factor.

3. The cavity under the resonator essentially eliminates the stiction problem where the device gets stuck to the substrate. We find that with the 10-μm cavity, even a millimeter long and slender structures never get stuck to the substrate.

The benefit from porous silicon getter:

The highly reactive porous silicon can readily react to the oxygen gas and form the oxide layer that can trap other gas molecules and hence maintain low pressure in the cavity. This benefit is realized without additional materials or masks in the process.

VIII. REFERENCES

[1] Y. T. Cheng, W. T. Hsu, K. Najafi, C. T. C. Nguyen, and L. Lin, "Vacuum packaging technology using localized aluminum/silicon-to-glass bonding," *J. Microelectromech. Syst ,* vol. 11, pp. 556-565, 2002.

[2] P. Monajemi, P. J. Joseph, P. A. Kohl, and F. Ayazi, "A low cost wafer-level MEMS packaging technology," in *MEMS'05,* pp. 634-637, 2005.

[3] V. Kaajakari, J. Kiihamäki, A. Oja, H. Seppä, S. Pietikäinen, V. Kokkala, and H. Kuisma, "Stability of wafer level vacuum encapsulated single-crystal silicon resonators," in *TRANSDUCERS '05*, Seoul, Korea, pp. 916-919, 2005.

[4] M. Lutz, W. Golderer, J. Gerstenmeier, J. Marek, B. Maihoefer, S. Mahler, H. Muenzel, and U. Bischof, "Precision yaw rate sensor in silicon micromachining," in *TRANDUCERS'97*, Chicago, IL, USA, pp. 847-850, 1997.

[5] B. L. Lee, C. H. Oh, S. Lee, Y. S. Oh, and K. J. Chun, "Vacuum packaged differential resonant accelerometer using gap sensitive electrostatic stiffness changing effect," in *MEMS'00,* Miyazaki, Japan, pp. 352-357*,* 2000.

[6] M. Moraja and M. Amiotti, "Getters films at wafer level for wafer to wafer bonded MEMS," in *Proceedings of the Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS* , Cannes-Mandelieu, France, pp 346-349, 2003.

[7] D. R. Sparks, S. Massoud-Ansari, and N. Najafi, "Chip-Level Vacuum Packaging of Micromachines Using Nanogetters," *IEEE Trans. Advanced Packaging,* vol. 26, pp. 277-282, 2003.

[8] W. Mohammad and V. Kaajakari, "Self-aligned maskless process for etching cavities in SOI wafers to enhance the quality factor of MEMS resonators," in *Reliability, Packaging, Testing, and Characterization of MEMS/MOEMS and Nanodevices IX*, San Francisco, California, USA, pp. 75920U-7, 2010.

[9] S. Brunauer, P.H. Emmet and E. Teller, " Adsorption of gases in multimolecular layer", *J. American Chemical Society*, vol. 60, pp. 309-19, 1938.

[10] R. Herino, G. Bomchil, K. Barla, C. Bertrand, and J. L. Ginoux, "Porosity and pore size distributions of porous silicon layers," *J. Electrochemical Society,* vol. 134, pp. 1994-2000, 1987.

[11] T. R. Hsu, *MEMS Packaging*. London: INSPEC, 2004.

[12] A. Berny, "Substrate Effects in Squeeze Film Damping of Lateral Parallel-Plate Sensing MEMS Structures," in *http://robotics.eecs.berkeley.edu/~pister/245/project/Berny.pdf.*

A Micro-Resonator for Fundamental Physics Experiments and its Possible Interest for Time and Frequency Applications

Le Traon O., Bahriz M., Ducloux O., Masson S., Janiaud D.
Physics Department,
ONERA, the French Aerospace Lab
Châtillon, France
olivier.le_traon@onera.fr

Kuhn A., Briant T., Cohadon P-F, Heidmann A.
Laboratoire Kastler Brossel, Ecole Normale Supérieure,
Paris, France
Michel C. , Pinard L., Raphaele Flaminio R.
Laboratoire des Matériaux Avancés, CNRS,
Villeurbanne, France

***Abstract*— a dedicated micro-resonator in a length-extension mode (LEM) has been developing for fundamental physics experiments, aiming to detect the quantum ground state of a resonator. The experiment is based on a high frequency, high quality factor and low mass micro-resonator, implemented in a high finesse Fabry-Perot cavity at cryogenic temperature. Very interesting preliminary measurements have been obtained on a 3.9 MHz, 25 µg resonator, given a 1.95 million quality factor, at room temperature and under a 5 10^{-2} mbar vacuum. By its planar configuration compatible with collective etching process, this resonator should also find interesting applications in the field of ultra stable oscillator (USO) and future quartz-MEMS USO.**

I. INTRODUCTION

Highlight the quantum ground state (QGS) of a micro-resonator is a quest involving many fundamental physics laboratories [1-6]. Such quantum fluctuations correspond to ridiculously small position fluctuations for usual macroscopic objects and two very stringent conditions must be fulfilled in order to reach and detect the mechanical QGS.

Firstly, the thermal energy ½ $k_B T$ at the cooled temperature T (T= temperature in Kelvin and k_B = Boltzmann constant) has to be small compared to the fundamental quantum energy ½ $\hbar\Omega_m$ of the harmonic oscillator (Ω_m = its mechanical resonance angular frequency of the resonateur, $\hbar$ = reduced Planck constant). Typical temperatures range of 50 mK for a 1-GHz resonator must be met which is a readily accessible temperature by standard dilution cryogenic techniques. However, temperature down to 50 µK is required for a resonator oscillating at 1 MHz. In that case, other cooling technique has to be implanted like the recently-demonstrated laser cooling based on the optomechanical coupling between the resonator and a laser beam in a detuned Fabry-Perot cavity [7-9].

Secondly, very small position fluctuation $\Delta x_0 = (\hbar/2M\Omega_m)^{1/2}$ of the QGS has to be detected, depending on the mass resonator M, its quality factor Q and its frequency Ω_m: For a 3-MHz oscillator with a 25-µg mass and a 3.10^6 quality factor (corresponding to a 1-Hz mechanical resonance bandwidth), this condition requires measurement sensitivity better than 10^{-34} m^2/Hz.

To fulfill these both conditions, the experiment is based on a high finesse Fabry-Perot cavity (figure 1) at cryogenic temperature, able to detect extremely small position fluctuation, better than 10^{-36} m^2/Hz, thanks to a high cavity finesse of 10 000, a power source of 1 mW and a laser wavelength λ of 1064 nm.

.

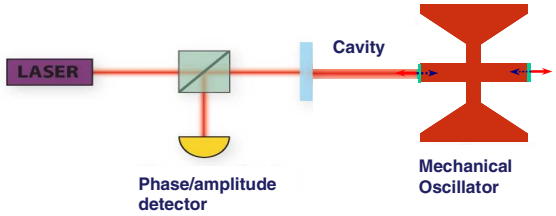


Figure 1. Experiment to detect the quantum ground state of a resonator : The resonator takes place of a moving mirror of a high finesse Fabry-Perot cavity at cryogenic temperature.

In this experiment, the resonator takes place of the moving mirror of the cavity and we have choosen a length-extension mode (LEM) in order to maximize the optomechanical coupling (the position fluctuations are indeed directly coupled with the cavtiy length).

An other important advantage of a length-extension resonator concerns the mirror-carrying surface at the top of the resonator. This area corresponds to a quasi null strain zone, lowering the coating losses of the highly reflective mirror deposited on it [10]. The intrinsic quality factor of the resonator is thus preserved .

978-1-61284-110-6/11/$26.00 ©2011 IEEE

A few MHz resonator with a few 10 µg mass and a quality factor in the order of 10^6 at room temperature is to be achieved in order to be able to highligth the QGS. A high initial quality factor is needed because of the laser cooling technique used in this experiment which reduces the quality factor of the resonator in same factor than the temperature: lowering the dilution cryogenic temperature of 50 mK to 50 µK reduces the quality factor by a factor 1000 and requires more acurate detection.

Quartz crystal has been chosen for the realisation of the resonator, because of its well known high intrinsic quality even at cryogenic temperature [11]. Quality factor above 300 million has also been recently reported [12] for a quartz macro-resonator in a length-extension mode at 15 MHz.

This paper describes the design, the optimization, the manufacturing and the first evalution of these dedicated length extension resonator. First reflexions of its possible application in time and frequency domains is also discussed.

II. RESONATOR CONCEPT

The design of a high Q resonator is not so easy and care must be taken in order to confine the vibrating energy into the resonator. The quality factor Q of a resonator is indeed defined as the ratio between the stored vibrating energy (W) and the total dissipated energy (ΔW) per cycle of vibration (times 2π). Just consider a 10^{-6} energy loss of the resonator energy, for example in its clamping, and further consider that all of this energy loss is dissipated, you finally obtain a resulting Q factor limited to a maximum of 10^6.

The initial concept of length-extension resonator was based on pillar, clamped at its center by a thin membrane. In that case, reducing the energy loss out of the resonator is a real challenge because of the well-know Poisson effect (figure 2.), relating the fact that a pillar submitted to an extension strain induces a transverse strain in the ratio of the Poisson coefficient ν (generally of 0.1 to 0.5, ν = 0.5 corresponding to incompressible material).

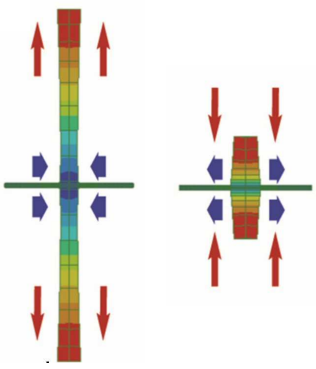


Figure 2. Resonator in length-extension mode: the Poisson effect induces a strong lateral displacement

An original concept has been proposed [13] in order to solve this problem. It is based on a complementary ring around the pillar and an external membrane for the clamping of the resontor (figure 3.).

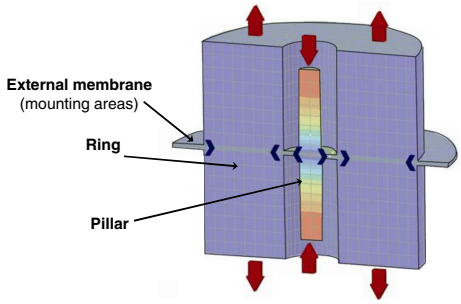


Figure 3. Proposed concept: a ring around the pillar with an out of phase length-extension vibration compared to the pillar.

Thanks to particular tuning conditions of the pillar and the ring frequencies, it is possible to avoid the lateral displacement of the external membrane, allowing a mounting of the resonator without energy loss. This effect is illustrated on figure 4. showing the resonator vibration displacement in free-free boundary conditions (the external membrane is not clamped).

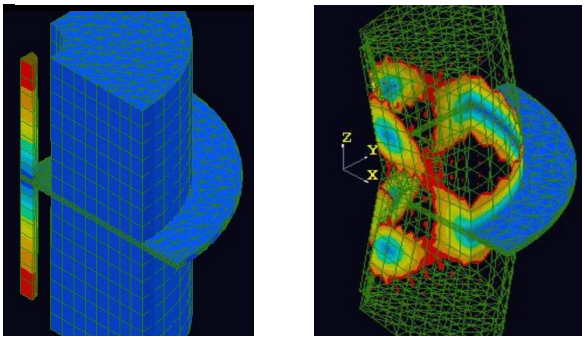


Figure 4. Resonator in free-free boundary conditions
-Left: pillar vibration
-Right: observation scale magnified by 400. No displacement of the external membrane.

On the left, the length-extension pillar vibration is shown (only one quarter of the structure is represented) and the right picture gives the same picture with an observation scale magnified by a factor 400. No vibration displacement is induced to the external membrane, thanks to a compensation of the pillar Poisson effect by an out of phase length-extension of the ring (coupled with a ring bending).

III. QUARTZ RESONATOR OPTIMIZATION

A dedicated optimization has been undertaken, taken into account the specificity of quartz crystal and the chosen

micromachining by wet chemical etching for the manufacturing of the quartz resonator. A crystallographic orientation of the pillar along the Z axis has been chosen, as well as a pillar equilateral cross-section with respect to the quartz trigonal symmetry: a nominally symmetric structure with regard to its median plane is thus possible, which is a sine qua non condition to obtain a good balancing of the length extension mode (manufacturing by wet chemical etching).

Figure 5. gives the planar shape of the quartz resonator (left picture) and its 3D Finite Element Model (FEM), including the etching facets created by the chemical etching.

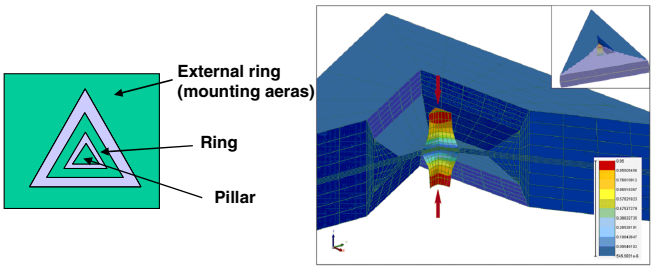


Figure 5. Quartz resonator configuration :
-Left: triangular shapes taken into account the trigonal symetry of quartz (Z cut quartz wafer).
- Right: finite element model taking into account the etching patterns obtained by chemical etching of quartz.

The typical dimensions of the pillar are a triangular cross-section side of 200 µm and a total length of 1 mm. The external and internal are 50 µm thick, given a theoretical resonance frequency and an effective mass around 3.5 MHz and 25 µg respectively.

Finite Element simulations (SAMCEF code) have been done in order to optimize the resonator dimensions and study the impact of some dissymmetry effects as a lateral misalignment (figure 6.) or a dissymmetry with respect to the median plan (figure 7.). The optimization has been focused on the dynamic behaviour of the resonator to avoid any interference with other vibration modes and to reduce the energy loss out of the resonator.

By defining the $Q_{support}$ as the ratio of the total stored energy in the resonator by total energy out of the resonator (the clamping for example), the last optimization of the quartz structure gives a $Q_{support} > 10^6$, including a lateral misalignment and a height dissymmetry of 2 µm, corresponding to our actual technology limitation. This $Q_{support}$ is thus compatible with an intrinsic quality factor in the range 1 – 10 million, depending of the effective support energy dissipation.

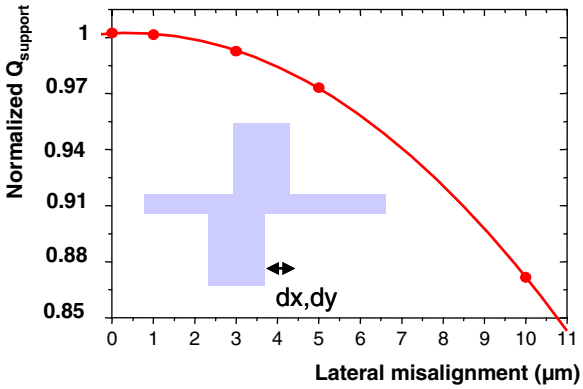


Figure 6. Effect of a lateral misalignment on the $Q_{support}$ for the optimiezd quartz strcuture.

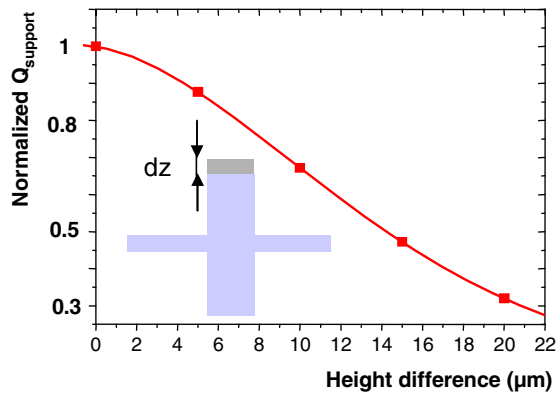


Figure 7. Effect of a dissymmety with respect to the median plan on the $Q_{support}$ for the optimizd quartz structure.

IV. QUARTZ RESONATOR MANUFACTURING & CARACTERIZATION

Only one step of chemical etching is needed to obtain the quartz LEM resonator (figure 8.). This very simple process is very similar to those developed for other advanced quartz MEMS devices [14-15].

A dedicated NH_4F/HF mixture is used for the chemical etching of quartz, allowing a low roughness of the etched surface defining the internal and external thin membranes.

Critical steps concern photoresist deposition on the 3D structure obtained after chemical etching, the photolithographic process defining the front and back mirrors, and most of all the removing of the photoresist after the coating of the multilayer high reflective SiO_2/Ta_2O_5 mirrors (because of the high temperature reached by the photoresist during the ion beam sputtering). The photoresist will be soon replaced by a metallic mask at this process step.

Figure 9. shows a 1.5" square quartz wafer including 9 LEM resonators obtained by collective chemical etching.

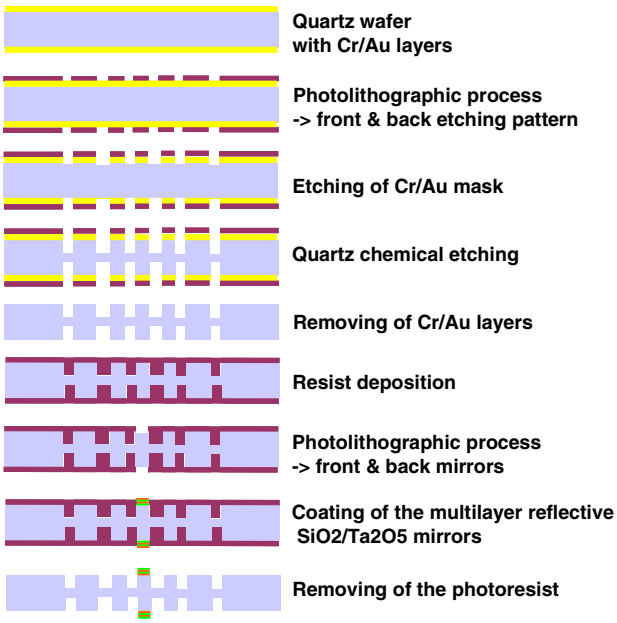


Figure 8. manufacturig process of the planar-planar LEM resonator: only one step of chemical etching is needed to obtain the quartz structure.

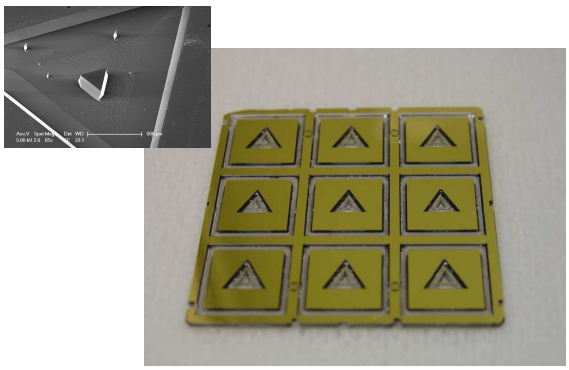


Figure 9. 1,5" square quartz wafer including 9 LEM resonators obtainend collectivelly by wet quartz chemical etching. Top & left, detail of the micro-pillar.

Charaterization of the second generation of LEM resonators (at this time, still without the hight refelctive mirrors), have been made thanks to a dedicted test-bench based on a Michelson-interferometer configuration, able to work under vacuum [16]. Particularly intersting results have been obtained since a quality factor of 1.95 million has been measured at room temperature (by a ring-down method), for a LEM resonateur frequency of 3.87 MHz and a vacuum of 0.2 mbar vacuum (figure 10.). These results are in improvement by a factor 5 in comparaison with previous results [16], mainly due to the new optimization of the structure and the improvement of the manufacturing quality.

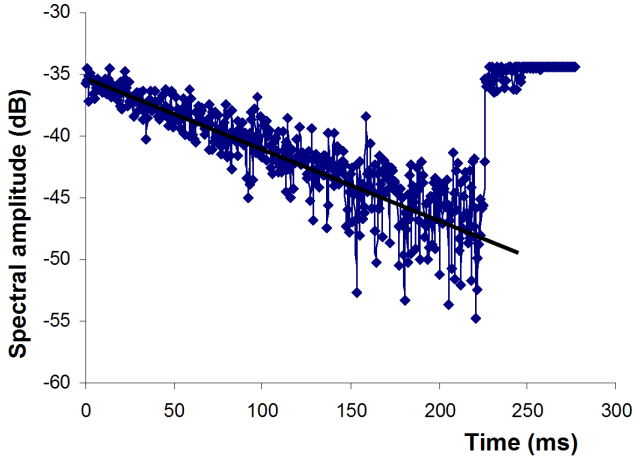


Figure 10. Measurement of the last generation of quartz LEM resonator:s a quality factor of 1.95 million at a frequency of 3.87 MHz has been obtained at room temperature and under a vacuum of 0,2 mbar.

V. CONCLUSION & PERSPECTIVES

A new concept of a length-extension mode (LEM) quartz resonator has been proposed and developed for the observation of its quantum ground state. Very interesting results have been obtained on the last generation of resonators, given a quality factor near 2 million at a frequency around 4 MHz and a very small effective mass of 25 µg: the measured Q x F product of 8.10^{12} is near the state of the art of ultrastable oscillator [17] where the best shear mode macro quartz resonators show a 3. 10^{13} Q x F at room temperature.

Next steps will be the coating of a highly reflective mirror on the top of the resonator, and its implementation in a dedicated Fabry-Perot cavity. Thanks to the high mechanical quality factor of the resonator and thus to the efficient cooling mechanisms that can be used, low temperature operation of such an optomechanical device would open the way to the observation of the quantum ground state of a micro-mechanical resonator.

By its planar configuration compatible with collective etching process, this resonator should also find interesting opportunities in the time and frequency domain, in particular for a new generation of miniature ultrastable oscillator (USO) compatible with collective manufacturing (Quartz MEMS USO). Work has been undertaking in this way to first identify a temperature compensated cut suited to our resonator design and its actuation/detection by piezoelectric effect, as it has been found for other length-extension design resonators [18]. Collective micromachining of quartz wafer based on recent quartz Deep Reactive Ion Etching techniques will be also developed, as well as a quartz wafer level packaging compatible with the vacuum requirement.

ACKNOWLEDGMENT

This work was done thanks to the financial support of the "Agence Nationale de la Recherche" (ANR) France, Programme blanc, N° ANR-07-BLAN-2060.

REFERENCES

[1] O. Arcizet *et al*., "High-sensitivity optical monitoring of a micromechanical resonator with a quantum-limited optomechanical sensor", Phys. Rev. Lett., vol. 97, p. 133601, 2006

[2] T. Corbitt, D. Ottaway, E. Innerhofer, J. Pelc and N. Mavala, "Measurement of radiation-pressure-induced optomechanical dynamics in a suspended Fabry-Perot cavity," Phys. Rev. A 74, 021802 (2006)

[3] S. Gröblacher *et al*., "Demonstration of an ultracold microoptomechanical oscillator in a cryogenic cavity, Nature Physics, vol. 5, p. 485, 2009

[4] A. Schliesser, O. Arcizet, R. Rivière, G. Anetsberger, T.J. Kippenberg, "Resolved-sideband cooling and position measurement of a micromechanical oscillator close to the Heisenberg uncertainty limit", Nature Physics, vol. 5, p. 509, 2009

[5] T. Rocheleau *et al*., "Preparation and detection of a mechanical resonator near the ground state of motion", Nature, vol. 463, p. 72, 2010

[6] A.D. O'Connell *et al*., "Quantum ground state and single-phonon control of a mechanical resonator", Nature, vol. 464, p. 697, 2010

[7] P.F. Cohadon, A. Heidmann, M. Pinard, "Cooling of a mirror by radiation pressure", Phys. Rev. Lett., vol. 83, p. 3174, 1999

[8] O. Arcizet, P.-F. Cohadon, T. Briant, M. Pinard, A. Heidmann, "Radiation pressure cooling and instability of a micromirror in a detuned high-finesse optical cavity", Nature, vol. 444, p. 71, 2006

[9] S. Gigan *et al*., "Self-cooling of a micromirror by radiation pressure", Nature, vol. 444, p. 67, 2006

[10] O. Le Traon, , S. Masson, C. Chartier, D. Janiaud, "LGS and GaPO4 piezoelectric crystals: new results", Solid State Sciences, vol. 12, Issue 3, March 2010,p. 318-324.

[11] B. Komiyama, "Quartz Crystal Oscillator at Cryogenic Temperature", Frequency Control Symposium, 1981, p. 339

[12] S. Galliou, J. Imbaud, M. Goryachev, R. Bourquin, P. Abbe, "Losses in high quality quartz crystal resonators at cryogenic temperatures", APL vol. 98 Issue 9 p. 091911-091911-3, 2011

[13] M. Bahriz, O. Le Traon, "Module de découplage mécanique d'un résonateur à grand coefficient de qualité" French patent n° 10/02829 July 2010

[14] O. Le Traon, D. Janiaud, B. Lecorre, S. Muller, M. Pernice, J.-Y. Tridera "A New Quartz Monolithic Differential Vibrating Beam Accelerometer" ,Proceedings of the2006 IEEE/ION Position, Location And Navigation, San Diego, April 2006

[15] O. Le Traon, D. Janiaud, S. Masson, S. Muller, M. Pernice, " The DIVA Accelerometer and VIG Gyro : Two Quartz Inertial MEMS for Guidance and Navigation Applications" ION NTM, January 2007

[16] M. Bahriz, O. Ducloux, S. Masson, D. Janiaud, O. Le Traon, A. Kuhn, C. Molinelli, T. Briant, P.-F. Cohadon, A. Heidmann, C. Michel, L. Pinard, R. Flaminio, "A Micromechanical Resonator to Reach the Quantum Regime", IEEE Sensors 2010

[17] R. Besson *et al*., "BVA resonators and oscillators: a review. Relation with space requirements and quartz material characterization," Frequency Control Symposium, 1995, pp 590-599, San Francisco

[18] Yokoyama Y., Kawashima H., Kanie H., Design and Fabrication of a Length-Extensional Mode Rectangular X-Cut Quartz Resonator with Zero Temperature Coefficient, IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol. 53, no. 5, 2006

Characteristics of AlN Lamb Wave Resonators with Various Bottom Electrode Configurations

Chih-Ming Lin[1], Ventsislav Yantchev[2], Yung-Yu Chen[3], Valery V. Felmetsger[4], and Albert P. Pisano[1]
[1]Department of Mechanical Engineering, University of California, Berkeley, California, USA
[2]Department of Solid State Electronics, Uppsala University, Uppsala, Sweden
[3]Department of Mechanical Engineering, Tatung University, Taipei, Taiwan
[4]PVD Product Group, OEM Group Incorporated, Gilbert, Arizona, USA

***Abstract*—The characteristics of aluminum nitride (AlN) Lamb wave resonators utilizing the lowest symmetric (S_0) mode with grounded, floating, and open bottom surface configurations are theoretically and experimentally investigated. The Lamb wave resonator without the bottom electrode exhibits a quality factor (Q) as high as 2,573 but a low effective coupling of 0.18% at 949.7 MHz. On the contrary, the Lamb wave resonator with a floating bottom electrode shows an effective coupling of 1.05% but a low Q of 850 at 850.3 MHz because the imperfect interface between the AlN plate and bottom electrode usually degrades the Q. Limited by the larger static capacitance, in contrast to the floating bottom electrode, the Lamb wave resonator with a grounded bottom electrode shows a smaller effective coupling of 0.78% and a low Q of 800 at 850.5 MHz. These results suggest that the resonator with a floating bottom surface is suitable for filter applications, whereas that with an open bottom surface is preferred for sensor and oscillator applications.**

Keywords–Lamb Wave Resonator, Aluminum Nitride, Quality Factor, Electromechanical Coupling, Floating Bottom Electrode

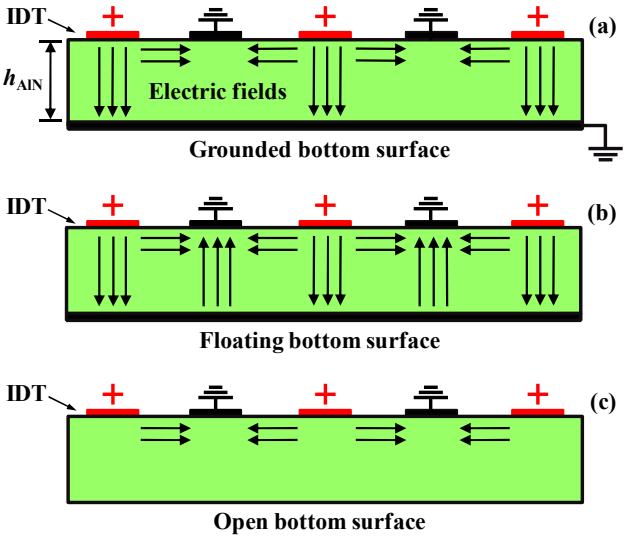


Figure 1. Illustration of the cross-section of AlN Lamb wave resonators with grounded, floating, and open bottom surface configurations.

I. INTRODUCTION

Recently, there are a large number of long-term evolution (LTE) frequency bands that are being designated as the possibilities for the third generation (3G) and the fourth generation (4G) wireless mobile systems. For the frequency division duplex (FDD) and time division duplex (TDD) technologies, most of the LTE frequency bands are allocated within the ultra-high-frequency (UHF) region. Several LTE frequency bands have designated with the narrow bandwidths between 0.75% and 2% [1]. In recent years, a new class of aluminum nitride (AlN) Lamb wave resonators utilizing the lowest symmetric (S_0) mode is being widely explored [2]–[7]. The AlN Lamb wave resonator usually exhibits a high phase velocity around 10,000 m/s and a moderate electromechanical coupling, very suitable for the high-frequency and narrowband wireless transmission systems.

During the past six years, two distinct designs of S_0 mode Lamb wave resonators are simultaneously developing. One is based on the Lamb wave excitations from the interdigital transducer (IDT) and the reflections from the periodic gratings on the AlN membrane (the grating-type reflector) [2], [3]; instead, the other one is based on the Lamb wave excitations from the IDT and the reflections from the suspended free edges of the AlN membrane (the edge-type reflector) [4]–[7]. Furthermore, the robust temperature compensation for the two types of S_0 mode Lamb wave resonators has been theoretically analyzed and experimentally demonstrated [8]–[10]. These aforementioned results reveal the potential of Lamb wave resonators for the future narrowband wireless transmission systems because this technology can reliably provide a wide frequency range from hundreds MHz to several GHz on the single silicon chip and simultaneously offer the high quality factor (Q) as well as the low motional resistance (R_m) in air.

In this work, the S_0 mode Lamb wave resonators utilizing the edge-type reflector with the various bottom surface configurations are studied. As shown in Fig. 1, the electric field distributions in the AlN membrane would dominate the performances of the resonators with grounded, floating, and open bottom surface configurations. Based on the same IDT

This work was supported by the DARPA Chip-Scale Mechanical Spectrum Analyzers (CSSA) award (Grant No. NGC-8140000613) under the DARPA Microsystems Technology Office (MTO).

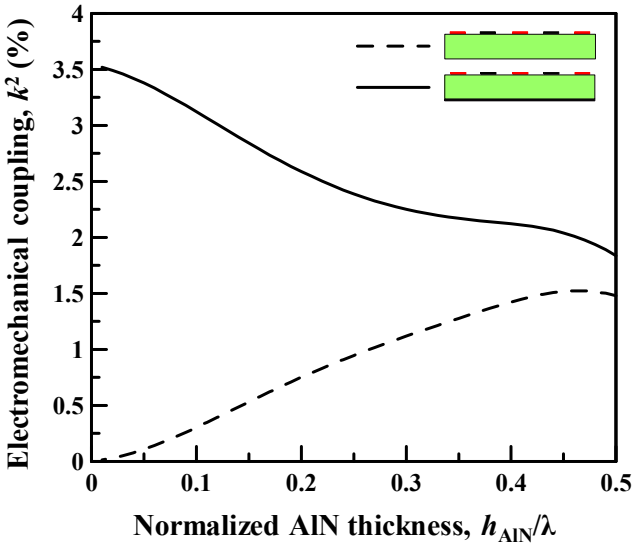


Figure 2. The dispersive electromechanical coupling coefficients of the lowest symmetric (S_0) mode Lamb waves in the AlN membrane with the non-metallized and metallized bottom surfaces.

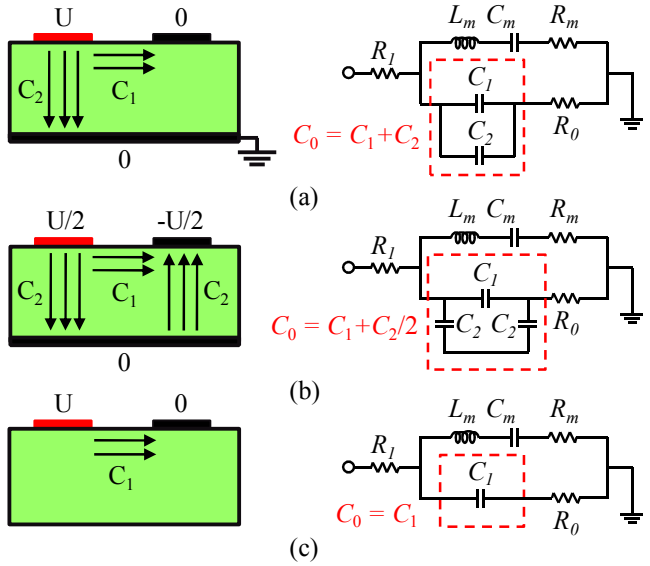


Figure 3. A simple physical model for the capacitive feedthrough to ground in the one-port Lamb wave resonator and the equivalent C_0 in the modified Butterworth-Van dyke (MBVD) models for the electrically (a) grounded, (b) floating, and (c) open bottom surfaces.

design, the AlN Lamb wave resonator with a floating bottom electrode provides a larger effective coupling (k^2_{eff}) of 1.05%, presenting a 1.35× increase in k^2_{eff} over that with a grounded bottom electrode because the floating bottom electrode makes a smaller static capacitance in the membrane. However, the electrode-to-resonator interface stress causes the degradation in Q. In contrast to the metallized bottom surface, the Lamb wave resonator with the non-metallized bottom surface (i.e. the open bottom surface) provides a Q as high as 2,573 but a small k^2_{eff} of 0.18% due to the weak electric field in the 1.5-μm-thick AlN membrane.

II. Fininte Element Analysis of One-Port Lamb Wave Resonators With Various Bottom Surfaces

By using Adler's approach [11] or finite element analysis (FEA) method [12], the dispersive electromechanical coupling in the piezoelectric plate can be theoretically calculated. In the simulation, the S_0 mode Lamb waves propagate in the plane normal to the c-axis of the AlN membrane and the bottom electrode is assumed to be infinitely thin. As illustrated in Fig. 2, the dashed and solid lines represent the electromechanical coupling dispersion of the S_0 mode Lamb waves in the AlN membrane with the non-metallized and metallized bottom surfaces, respectively. In general, the metallized bottom surface enhances the electromechanical coupling coefficient because the electric field strength through the AlN thin plate is higher than that between the IDT electrodes. However, the electromechanical coupling calculated using Adler's approach is identical for the both grounded and floating bottom surfaces because the short-circuited boundary condition is considered in the two configurations. This is not true for the real Lamb wave resonator since the parallel frequency and the effective coupling are significantly influenced by the static and parasitic capacitances in the resonator. Therefore, the actual effective couplings for the grounded and floating bottom electrodes are different because of the distinct static capacitances in the two configurations.

As shown in Fig. 3, a simple physical model can be used to illustrate the effect of capacitive feedthrough to ground in the one-port Lamb wave resonator and the static capacitance (C_0) in the modified Butterworth-Van dyke (MBVD) models for the electrically grounded, floating, and open bottom surfaces. For the grounded bottom surface in Fig. 3(a), the lateral and vertical electric fields induce the in-line capacitance (C_1) and cross-field capacitance (C_2) so the C_0 is simply assumed to be the summation of C_1 and C_2. As for the floating bottom surface in Fig. 3(b), the electric potential in the bottom electrode is presumed to be zero and the electric potentials in the electrodes are equivalent to be U/2 and –U/2 because of the symmetric electric filed. As a result, the C_0 is simply assumed to be the summation of C_1 and $C_2/2$ for the floating bottom surface. As illustrated in Fig. 3(c), C_0 is simply equal to the in-line capacitance, C_1, in the open bottom surface. The excitation efficiency of the IDTs with grounded and floating bottom surfaces would be the same and larger than that with an open bottom surface because the backside metallization strongly enhances the vertical electric field in the AlN plate. In addition, C_0 in the open bottom surface is the smallest among the three configurations; C_0 in the floating bottom surface is smaller than C_0 in the grounded bottom surface, resulting in a larger k^2_{eff} for the resonator with a floating bottom electrode.

In order to verify the above physical model, as shown in Fig. 4, the two-dimensional (2D) model in COMSOL FEA software is used to simulate the frequency spectra of one-port AlN Lamb wave resonators with the grounded, floating, and open bottom surfaces. With the same 11 electrode design but different electric conditions in the bottom surfaces, the 2D COMSOL FEA model predicts that the floating bottom electrode provides a 1.8 times larger k^2_{eff} than the grounded bottom electrode on a 1.5-μm-thick AlN plate. The simulation results also confirm that the transduction efficiency of the IDTs with grounded and floating bottom electrodes would be

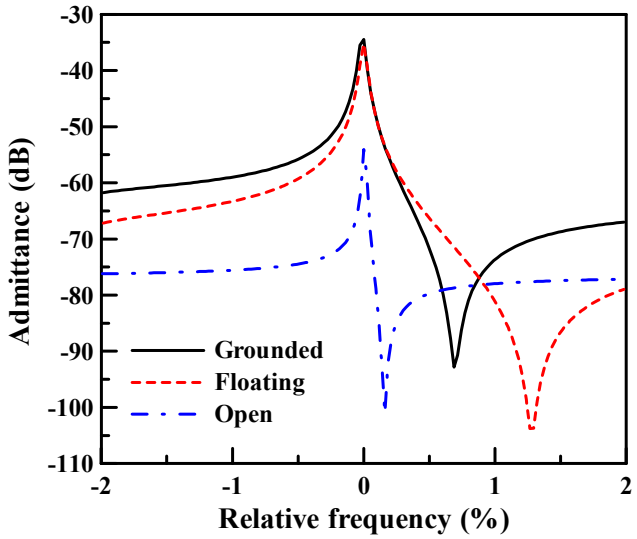


Figure 4. Simulated admittance plots of the one-port AlN Lamb wave resonators with the grounded, floating, and open bottom surfaces using the two-dimensional (2D) model in COMSOL.

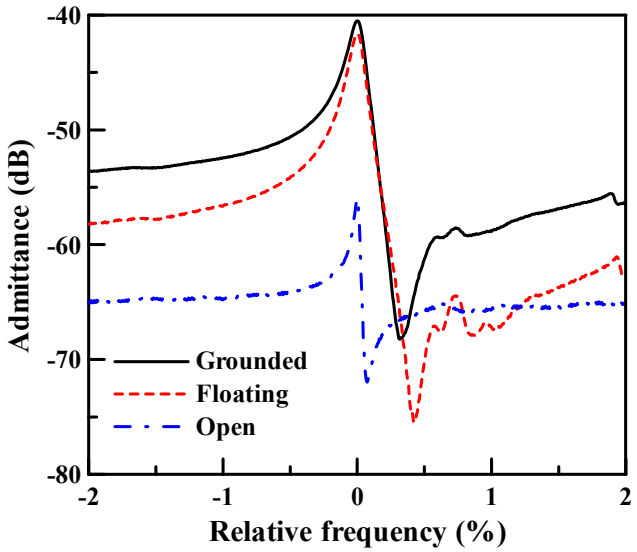


Figure 5. Measured admittance plots of the one-port AlN Lamb wave resonators with the grounded, floating, and open bottom surfaces.

TABLE I. GEOMETRIC DIMENSIONS OF LAMB WAVE RESONATORS

	Design 1	Design 2	Design 3
IDT electrodes	11	11	11
Aperture	100 μm	100 μm	100 μm
Electrode width	2.77 μm	2.77 μm	2.77 μm
Pt IDT electrode thickness	100 nm	100 nm	100 nm
Pt bottom electrode thkickness	100 nm	100 nm	-
AlN membrane thkickness	1.5 μm	1.5 μm	1.5 μm
Bottom electric condition	Grounded	Floating	Open

identical when the same even electrode pairs and acoustic loss conditions are considered in the simulation.

III. EXPERIMENTAL RESULT AND DISCUSSION

In order to study the effect of the bottom electric condition on the resonator performance, the AlN Lamb wave resonators with electrically grounded, floating, and open bottom surfaces are designed as summarized in Table I. The micro-fabrication process has been presented in [6] so it is not further considered in this work. To diminish the experimental error resulted from the fabrication process, the resonators were fabricated on the same wafer and placed in the vicinity. Lamb wave resonators were all tested in air and S_{11} parameters were extracted using an Agilent E5071B network analyzer. The measured Q was extracted from the admittance plot by dividing the resonance frequency by the 3 dB bandwidth. The k^2_{eff} of the measured devices is defined as

$$k_{eff}^2 \approx \frac{\pi^2}{4}\left(\frac{f_p - f_s}{f_p}\right) \quad (1)$$

where f_s and f_p are the series frequency and parallel frequency, respectively.

Fig. 5 presents the measured frequency spectra of one-port AlN Lamb wave resonators with the grounded, floating, and open bottom surfaces. With the same 11 IDT electrodes but different bottom electric conditions as listed in Table I, the Lamb wave resonator with the open bottom surface exhibits the smallest k^2_{eff} of 0.18% and transduction efficiency as expected. However, it exhibits a loaded Q as high as 2,573 among the three resonator designs. The transduction efficiency and effective coupling of the resonator with an open bottom surface can be enhanced by increasing the IDT electrodes or thickening the AlN membrane [13].

On the other hand, the Lamb wave resonator with a floating bottom electrode shows the highest k^2_{eff} of 1.05%, presenting a 1.35× enhancement in k^2_{eff} over that with a grounded bottom electrode, whereas the presence of additional bottom electrode decreases the Q to 850 because of the high stress levels in the interface. In addition, the extracted C_0 in the AlN plate with the floating and grounded bottom electrode are 195 fF and 355 fF, respectively. The result confirms that the floating bottom electrode efficiently reduces C_0 and boosts k^2_{eff} in the one-port AlN Lamb wave resonator. Moreover, the employment of the floating bottom electrode also makes the micro-fabrication process easier because the electrical contact via is no longer needed.

In the 2D COMSOL FEA model, the simulated C_0 ratio of the grounded bottom electrode to the floating bottom electrode is approximately equal to 2 for the same design parameters listed in Table I. However, the measured C_0 ratio is equal to 1.82 because some additional significant parasitic capacitances occur and influence the experimental C_0 ratio. In fact, these additional parasitic capacitances are dead and do not take the responsibility for the Lamb wave excitation, but only decrease the parallel frequency and consequently reduce the k^2_{eff} for a given AlN plate thickness.

In addition to k^2_{eff} and C_0, the experimental transduction efficiency of the Lamb wave resonators with grounded and floating bottom electrodes is of interest. For this purpose, the coupling-of-modes (COM) transduction coefficient, $\alpha(\omega_0)$, which is responsible for the excitation efficiency of the IDT,

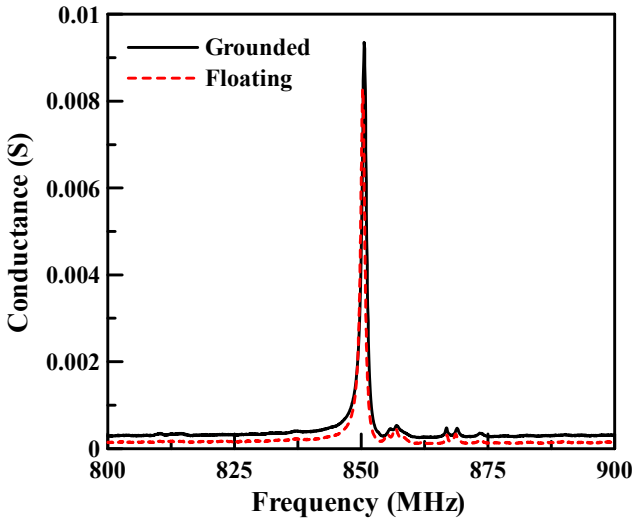


Figure 6. Measured conductance plots of the one-port AlN Lamb wave resonators with the grounded and floating bottom electrodes.

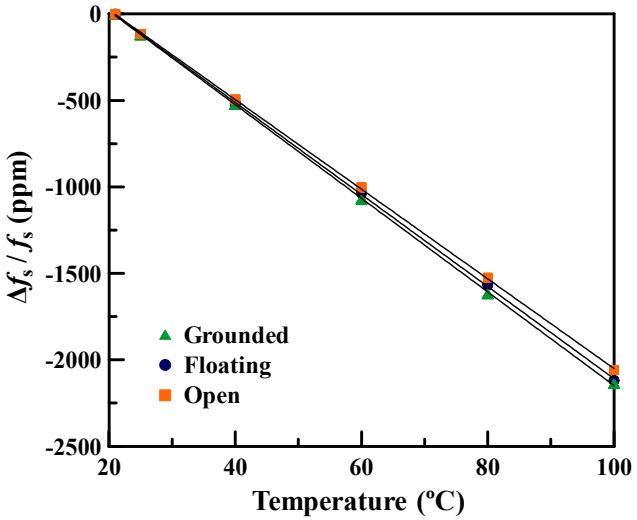


Figure 7. Measured frequency variation vs. temperature of the Lamb wave resonators with the electrically grounded, floating, and open bottom surfaces.

TABLE II. PERFORMANCE OF LAMB WAVE RESONATORS

	Grounded	Floating	Open
f_s [MHz]	850.5	850.3	949.7
Q	800	850	2573
Motional resistance, R_m [Ω]	78	90	685
Motional capacitance, C_m [fF]	2.22	1.66	0.14
Static capacitance, C_0 [fF]	355	195	96
k^2_{eff} [%]	0.78	1.05	0.18
Figure of merit, $k^2_{eff}Q$	6.2	8.9	4.6
TCF [ppm/°C]	–27.0	–26.6	–25.9

can be introduced to compare the excitation efficiency of the IDTs with grounded and floating bottom electrodes. The COM transduction coefficient $\alpha(\omega_0)$ is derived as [14]

$$\alpha(\omega_0) = \frac{Q_F(\beta)}{\Lambda_T}\sqrt{\frac{\omega_0 W \Gamma_s}{2}}, \quad (2)$$

where Λ_T is the wavelength of transduction, $Q_F(\beta)$ is the elemental charge density function, $\beta=2\pi/\Lambda_T$ is the acoustic wave number, ω_0 is the resonance angular frequency, W is the IDT aperture, and Γ_s is expressed as [14], [15]

$$\Gamma_s = \frac{k^2}{2\varepsilon_s(\infty, \beta h_{AlN})}, \quad (3)$$

where $k^2=2(v_{oc}-v_{sc})/v_{oc}$ is the intrinsic electromechanical coupling coefficient in Fig. 2, v_{oc} and v_{sc} are the S_0 mode Lamb wave velocities at the open-circuited and short-circuited surfaces, respectively, $\varepsilon_s(\infty, \beta h_{AlN})$ is the effective permittivity for infinite slowness (the effective static permittivity), and h_{AlN} is the AlN thickness. Moreover, the acoustic conductance at the resonance frequency ω_0 of a reflector-less IDT having an aperture W and N_P pairs of electrodes is defined as [15], [16]

$$G_a(\omega_0) = \omega_0 W \Gamma_s \left|Q_F(\beta)\right|^2 N_p^2. \quad (4)$$

The latter can be presented as a function of the COM transduction coefficient as:

$$G_a(\omega_0) = 2\left[\Lambda_T \alpha(\omega_0) N_p\right]^2, \quad (5)$$

revealing further the physical meaning of the COM transduction coefficient. The latter is directly related to the IDT conductance which in turn is proportional to the radiated acoustic power.

When comparing the performance of the Lamb wave resonators, one should bear in mind that (5) becomes more general. The conductance at resonance depends not only on the COM transduction and the IDT length but also on the acoustic losses since the latter determines the Q. Accordingly, only comparing the resonators with the same topology could reveal the information about the transduction efficiency. Here we compare the Lamb wave resonators of the same topology with the electrically grounded and floating bottom surfaces, respectively. As shown in Fig. 6, both the AlN Lamb wave resonators show quite similar conductance characteristics. The conductance ratio between the resonators with floating and grounded bottom electrodes is equal to 0.88, resulting in a transduction efficiency ratio of 0.94, which is very close to the simulated ratio, 0.91, obtained from the 2D COMSOL FEA model. The different IDT transduction efficiency for grounded and floating bottom electrodes is caused by the unsymmetrical electric filed in the 11 IDT electrodes, while the experimental transduction efficiency ratio is 1.01 for 12 IDT electrodes. As discussed above, when simulating the transduction efficiency ratio, all remaining contributions to the conductance should be kept the same except for the bottom electric condition. The results further confirm our preliminary observation that the excitation efficiency is the same for the IDTs with electrically grounded and floating bottom surfaces, whereas the larger static capacitance in the AlN Lamb wave resonator with a grounded bottom electrode results in the smaller k^2_{eff}.

Fig. 7 shows the plot of measured fractional frequency variation for the Lamb wave resonators listed in Table I from

room temperature to 100°C. As expected, the AlN Lamb wave resonator without the bottom electrode exhibits a smaller first-order temperature coefficient of frequency (TCF) of –25.9 ppm/°C because the platinum (Pt) bottom electrode has a large negative TCF and decreases the overall TCF accordingly. The resonators with the electrically grounded and floating bottom electrodes show the first-order TCF of –27.0 ppm/°C and –26.6 ppm/°C, respectively. Table II compares the measured performance of the AlN Lamb wave resonators with grounded, floating, and open bottom surfaces. All in all, the experimental results are in qualitative agreement with the simulated conclusions from the 2D COMSOL FEA model. It is noted that the Q and k^2_{eff} reported in this work are not optimal by far. The better resonator design, micro-fabrication process, and thin film stress control can improve the figure of merit (FOM) to a large extent.

IV. Conclusion

Based on a 1.5-μm-thick AlN plate and the same IDT design, the characteristics of one-port Lamb wave resonators with the electrically grounded, floating, and open bottom surfaces are theoretically analyzed using the FEA method and experimentally investigated. The employment of the floating bottom electrode simultaneously provides a larger effective coupling, an easier micro-fabrication process, and a higher FOM than the grounded bottom electrode without sacrificing the IDT transduction efficiency much. In particular, the Lamb wave resonator with the floating bottom electrode shows the highest FOM which is most suitable for the filter applications. The high phase velocity, moderate electromechanical coupling, weak dispersion, and ability to tune the effective coupling (i.e. the bandwidth) by changing the bottom electric condition also shows the potential of AlN Lamb wave resonators for the future high-frequency and narrowband wireless transmission systems. On the contrary, the resonator with the open bottom surface configuration usually provides a higher Q which is most suitable for the oscillator and sensor applications.

References

[1] http://www.radio-electronics.com/info/cellulartelecomms/lte-long-term-evolution/lte-frequency-spectrum.php

[2] J. Bjurström, I. Katardjiev, and V. Yantchev, "Lateral-field-excited thin-film Lamb wave resonator," in *Appl. Phys. Lett.*, vol. 86, 154103 Apr. 2005.

[3] V. Yantchev, and I. Katardjiev, "Micromachined thin film plate acoustic resonators utilizing the lowest order symmetric Lamb wave mode," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 54, pp. 87-95, Jan. 2007.

[4] A. Volatier, G. Caruyer, D. P. Tanon, P. Ancey, E. Defay, and B. Dubus, "UHF/VHF resonators using Lamb waves co-integrated with bulk acoustic wave resonators," in *Proc. IEEE Ultrasonics Symp*., 2005, pp. 902–905.

[5] H. Matsumoto, A. Isobe and K. Asai, "Multilayer film piezoelectric Lamb wave resonator for several GHz applications," in *Proc. IEEE Intl. Freq. Contr. Symp.*, 2006, pp. 797-800.

[6] G. Piazza, P. J. Stephanou, and A. P. Pisano, "One and two port piezoelectric higher order contour-mode MEMS resonators for mechanical signal processing," *Solid-State Electron*., vol. 51, pp. 1596-1608, Nov. 2007.

[7] M. Rinaldi, C. Zuniga, C. Zuo, and G. Piazza, "Super-high-frequency two-port AlN contour-mode resonators for RF applications," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol 57, pp. 38-45, Jan. 2010.

[8] G. Wingqvist, L. Arapan, V. Yantchev, and I. Katardjiev, "A micro-machined thermally compensated thin film Lamb wave resonator for frequency control and sensing applications," *J. Micromech. Microeng.*, vol. 19, 035018, Mar. 2009.

[9] C.-M. Lin, T.-T. Yen, Y.-J. Lai, V. V. Felmetsger, M. A. Hopcroft, J. H. Kuypers, and A. P. Pisano, "Temperature-compensated aluminum nitride Lamb wave resonators," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 57, pp. 524-532, Mar. 2010.

[10] C.-M. Lin, T.-T. Yen, V. V. Felmetsger, M. A. Hopcroft, J. H. Kuypers, and A. P. Pisano, "Thermally compensated aluminum nitride Lamb wave resonators for high temperature applications," *Appl. Phys. Lett.*, vol. 97, 083501, Aug. 2010.

[11] E. L. Adler, "Electromechanical coupling to Lamb and shear-horizontal modes in piezoelectric plates" *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 36, pp. 223-230, Mar. 1989.

[12] F. Di Pietrantonio, M. Benetti, D. Cannatà, R. Beccherelli, and E. Verona, "Guided Lamb wave electroacoustic devices on micromachined AlN/Al plates," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 57, pp. 1175-1182, May 2010.

[13] C. Zuo, J. Van der Spiegel, and G. Piazza "1.05-GHz CMOS oscillator based on lateral-field-excited piezoelectric AlN contour-mode MEMS resonators," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 57, pp. 82-87, Jan. 2010.

[14] B. P. Abbott, C.-S. Hartmann, and D. C. Malocha, "Transduction magnitude and phase for COM modeling of SAW devices," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 39, pp. 54-60, Jan. 1992.

[15] V. Yantchev, "Coupling-of-modes analysis of thin film plate acoustic wave resonators utilizing the S0 Lamb mode," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 57, pp. 801-807, Apr. 2010.

[16] D. P. Morgan, *Surface-wave devices for signal processing*, 1st ed., Amsterdam, New York: Elsevier, 1985.

Multi-Mode Multi-Band Tunable Active Inductor Oscillators

Ulrich L. Rohde[1,2], *Fellow IEEE*
[1]Univ. of Cottbus, BTU Cottbus 03046, Germany
[2]Synergy Microwave Corp., NJ 07504, USA

Ajay K. Poddar[1,2], *Senior Member IEEE*
[1]Synergy Microwave Corp., NJ 07504, USA
[2]Technical University Munich, Germany

Abstract—**A novel multi-band active inductor (AI) voltage controlled oscillator (VCO) is discussed and implemented using SiGe heterojunction bipolar transistors (HBTs). A design strategy centered on improving the locking range and broadband tunability of the transistorized tuned-inductor is being discussed and validated through an example of triple-band (0.9-1.8 GHz, 1.8-3.6 GHz, and 3.6-6.9 GHz) VCOs. Detailed analysis has been performed to minimize the phase noise by incorporating dynamic self-injection mechanism, and study the method of suppression of undesired modes in active inductor VCOs for current and later generation communication systems**.

I. Introduction

Recent trend pushes towards multi-standard terminals where tunable oscillator is a key element to cover different wireless standards and applications such as cellular phones, bluetooth, and wireless local area network. The important requirements of multi-band oscillator are low cost, low phase noise, low power consumption, and should be amenable for integration in small chip area [1]-[3]. Varactor-tuned multi-band VCOs solution using switched inductors, switched capacitors, and switched resonators can cover different wireless standards operated at different frequency bands but not preferable due to die-size, cost, tuning-range, switching noise, and speed [4].

Tunable active inductor (TAI) technology offers advantages of wide tuning-range in small chip area and enabling as a cost-effective alternative but at the cost of lower dynamic range and poor phase noise performances [5]. The broadband tuning at high frequencies become very difficult with varactor-tuned VCOs topologies due to package parasitic associated with the tuning diodes, therefore, even with the degraded dynamic range and poor phase noise still it is a win-win situation to use active inductor topology for monolithic integrated multi-band VCOs solution at high frequencies [6].

An in-depth research of different active inductor circuit topologies are needed in order to lower the phase noise, power consumption, and improve the dynamic range prior to a large scale deployment of inductor-tuned oscillators where varactor-tuned oscillators are dominating currently. In this paper, detailed analysis has been performed to minimize the phase noise by incorporating dynamic self-injection mechanism using group-delay and phase-shifter network for controlling the large-signal drive-level and also study the method of suppression of undesired modes in active inductor-tuned VCOs.

A design strategy centered on improving the locking range and broadband tunability of the active inductor VCO is being discussed and validated through an example of triple-band (0.9-1.8 GHz, 1.8-3.6 GHz, and 3.6-6.9 GHz) VCOs for current and later generation communication systems.

II. Oscillator Inductive Dynamics

One of the key inductive dynamics in communication systems is the tunable LC oscillators using passive spiral inductors. With monolithic microwave integrated circuit (MMIC) technology, inductances are traditionally realized using spiral or stacked spiral inductors [1].

Figs. 1 and 2 show the simplified representation and characteristics of the spiral inductor. As shown in Fig.2a, the values of the inductance is proportional to the number of the turns/layers resulting silicon area required for routing is large [5]. Moreover, with the increase in the number of turns/layers, self-resonant frequency (SRF) decreases, and limits the applications at high frequencies (Fig. 2b). In this paper, we explore the possibility to replace the spiral inductor in multi-band oscillator with a tunable active inductor which is amenable for integrated circuit realization requiring only a fraction of the die-area and with improved noise dynamics at microwave frequencies.

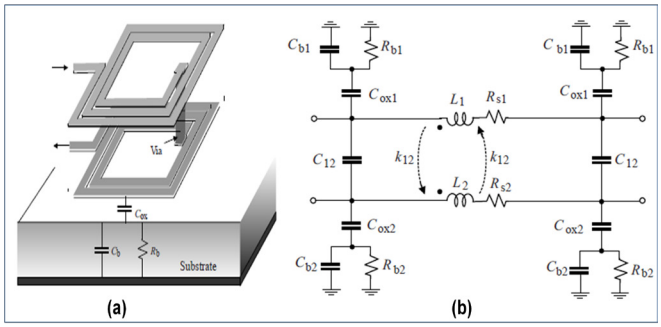


Fig. 1: (a) a typical sketch of spiral inductor, and (b) lumped model

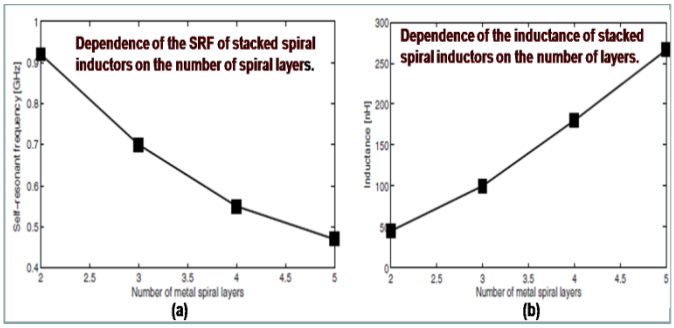


Fig. 2 Plot shows the dependence of the (a) Inductance, and (b) SRF.

Tunable multi-band oscillators can be realized using different technologies depending on the cost, performance, and applications [6]-[12]. Table 1 describes the different approach for the multi-band VCOs with pros. and cons. features for a given figure-of-merit (FOM). The performance comparison criterion (Q-factor, Die-area, Power-consumption, Linearity, Noise, Electro-Magnetic Interference, and Floor-planning) establishes important design guide lines for a selection of active inductor topologies for a given application and FOM. Table 2 summarizes the performance criteria and comparative analysis of the passive and active inductor for multi-band integrable tunable oscillator applications.

Table 1 Different approaches: Realization of Multi-Band VCOs

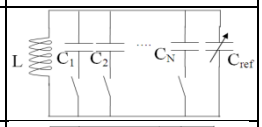
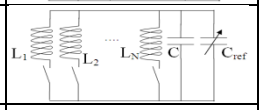
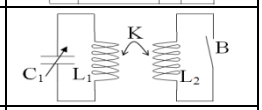
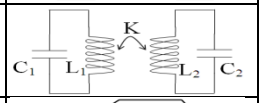
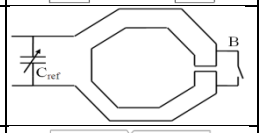
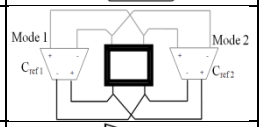
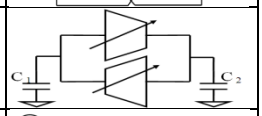
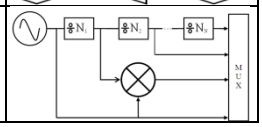
Topology	Features: Pros. (+) and Cons.(-)	Topology
Switched capacitor LC resonator	+Reduced K_{VCO} Limited C_{Max}/C_{Min} -Trade-off in switch size	
Switched inductor LC resonator	+Reduced K_{VCO} -Large area -Trade-off in switch size	
Switched coupled inductor LC resonator	+Smaller area -Trade-off in switch size (switch loss vs. Parasitic Capacitance)	
High-K transformer LC resonator	+high quality factor +Good Phase noise -Only one resonance mode	
Self-switching inductor LC resonator	+Smaller Area -Trade-off in switch size (switch loss vs. Parasitic Capacitance)	
Reusing internal turn of an inductor	+smaller area +No switch in tank for changing the mode	
Tunable active inductor (TAI)	+Smaller area -poor phase noise -Large power consumption	
Dividers, Mixers, and Multiplexers	+Small area -Large power consumption -spurious tone	

Table 2 Comparative analysis: Passive and Active Inductors

Performances	Passive Inductor (Spiral)	Active Inductor (Simulated: Active Device
Q-factor	Low Q: Q-factor can be improved by differential method but added cost	High Q: active inductor offers higher Q than the passive spiral inductor
Tunability	Fixed/Limited	Large tuning range
Die-Area	Large die-area	Small die-area
Power-Consumption	Zero	Significant: consumes power for generating active inductance, resulting to high Q that may offsets the power consumptions
Linearity	Good Linearity	Poor Linearity: driven under large-signal condition, causing shift in operating point, distortions, and impedance fluctuations
Noise	Superior: good phase noise performance	Poor : poor phase noise performance
EMI	Significant: Due to EM coupling in spiral inductors	EMI insensitive
Floor-Planning	Poor: large die- area makes difficult floor-planning	Not required

In this paper, design techniques towards minimizing the phase noise of the active inductor VCOs are discussed while maintaining a small real estate area and low power consumptions.

A. Active Inductor: State-of-The-Art

Active inductors are commonly implemented using transistors and associated circuitry (inductor-less passive components). Active devices (Transistors: Bipolars/FETs) exhibit 0^0 or 180^0 phase difference between output currents and input voltages/currents, therefore, at least one capacitor is required for the realization of an active inductor (since its V-I has 90^0 phase difference). By appropriate conversion, the port voltage can be imparted to lead the port current by 90^0, which is the intrinsic property of an inductor. This conversion is typically done using Gyrator (two transconductors connected back-to-back resulting inductive impedance) (Fig. 3).

The input admittance looking into port 2 of the gyrator is

$$Y = \frac{I_{in}}{V_2} = (sC_2 + G_{02}) + \left[\frac{1}{s\left(\frac{C_1}{G_{m1}G_{m2}}\right)+\left(\frac{G_{01}}{G_{m1}G_{m2}}\right)}\right] \quad (1)$$

From (1), equivalent RLC networks (Fig.3) is given by

$$R_p = \frac{1}{G_{02}},\ C_p = C_2, R_s = \left(\frac{G_{01}}{G_{m1}G_{m2}}\right), L_{eq} = \left(\frac{C_1}{G_{m1}G_{m2}}\right) \quad (2)$$

When either the input or the output impedances of the transconductors of Gyrator are finite, the synthesized inductors are no longer lossless and exhibit inductive property only in a specific frequency range. The performance of active inductor can be characterized by maximum usable frequency f_0 and Q-factor, can be characterized by

$$f_0 = \frac{1}{2\pi\sqrt{(L_{eq}C_p)}} = \frac{1}{2\pi}\sqrt{\frac{G_{m1}G_{m2}}{C_1C_2}} = \sqrt{f_1 f_2} \rightarrow f_{1,2} = \frac{G_{m1,2}}{C_{1,2}} \quad (3)$$

Where f_0 is the self-resonant frequency of the gyrator and $f_{1,2}$ is the cut-off frequencies of the transconductors.

The Q-factor of a conventional passive LC resonator at a given frequency is independent of the current swing passes through the resonator and can be described by

$$Q(\omega) = \frac{\omega_0}{2}\left[\frac{\partial\phi(\omega)}{\partial\omega}\right]_{\omega\to\omega_0} \quad (4)$$

where $\varphi(\omega)$ is the phase of the LC resonator impedance.

Unlike passive LC resonator, the inductance of the active inductors varies with the current/voltage passes through the inductors. Therefore, definition of Q-factor given in (4) does not hold well for active inductors, an alternative definition that accounts for the swing of the voltage/current can be given by

$$\overline{Q_a(\omega)}_{\omega\to\omega_0} = \left[\frac{\omega}{2(I_{max}-I_{min})}\int_{I_{min}}^{I_{max}} Q_a(\omega.i)\,di\right]_{\omega\to\omega_0} \quad (5)$$

where I_{min} and I_{max} are the minimum and maximum currents of the transconductors of active inductors, and $Q_a(\omega,i)$ is the instantaneous quality factor at frequency ω and channel current i provides an effective mean to quantify the Q-factor of active inductor when operated in a large-signal conditions, especially in tunable LC resonator oscillators.

B. Active Inductor: Realizations and Constraints

The realization of active inductor is mostly influenced by the fabrication technology and cost factor for available devices characteristics and FOM. For example, in CMOS technology, coupling capacitor is usually not available, hence, transistors may be connected directly for satisfying both DC biasing and RF signal flow at the same time. In RF/mixed-signal CMOS technology and other RF processes: GaAs and SiGe BiCMOS; MiM and poly-poly capacitors are generally available for signal coupling, and there is more design freedom [12]. Fig. 4 shows the typical implementation of an active inductor (AI) using SiGe HBTs in common emitter cascode topology.

We report SiGe HBTs AI in cascode topology (Fig.4). As shown in Fig. 4, transistors Q_1 (common emitter) & Q_2 (common base) are connected in cascode to provide high impedance. The O/P is feedback to the base of Q_1 through CB transistor Q_3 & Q_5, Q_5 is used as a voltage variable resistor. The base of Q_3 is grounded through R_3 that controls the –ve resistance of the AI, and Q_4 is used as a current source.

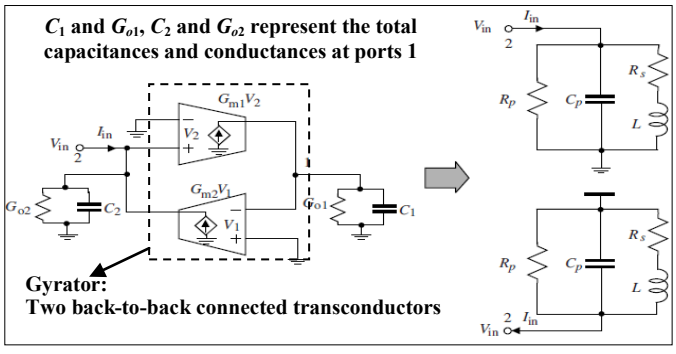


Fig.3 Typical realization of active inductor topology using a Gyrator

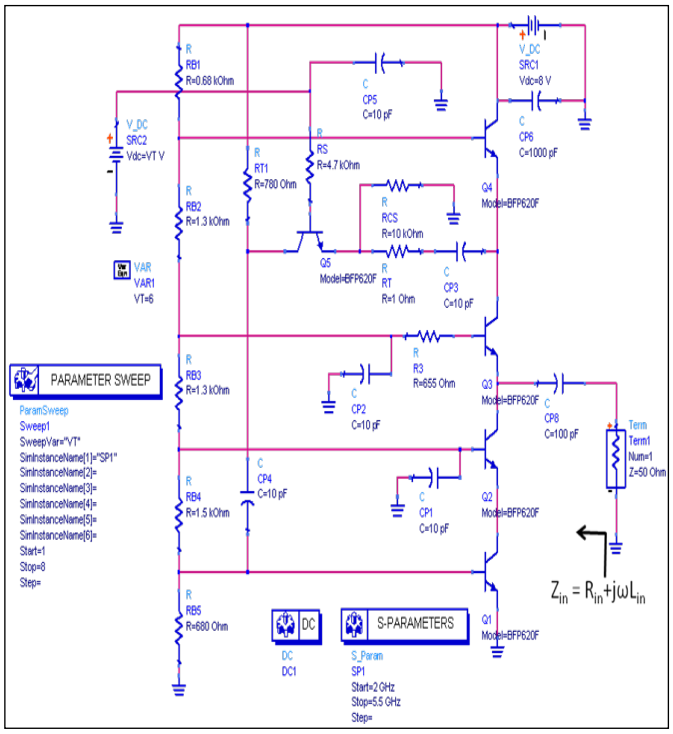


Fig.4 A typical schematic of active inductor using cascode topology

Fig. 5 shows the equivalent circuit model of active inductor (Fig. 4), for simplification of the analysis, all the transistor parameters are assumed to be identical ($C_{be1} \approx C_{be2} \approx C_{be3} = C_{be}$ and $g_{m1} \approx g_{m2} \approx g_{m3} = g_m$). The transistors Q_1 to Q_4 are biased at (2V, 6 mA), inductive property can be given by

$$L_{in} = \frac{1}{\omega}\left\{\frac{\omega C_{be}R_3(g_m+\varphi\gamma)-(\omega C_{be}-\varphi\delta)}{(g_m+\varphi\gamma)^2+(\omega C_{be}-\varphi\delta)^2}\right\} \quad (6)$$

$$R_{in} = \frac{(g_m+\varphi\gamma)+\omega C_{be}R_3(\omega C_{be}-\varphi\delta))}{(g_m+\varphi\gamma)^2+(\omega C_{be}-\varphi\delta)^2} \quad (7)$$

where

$$\gamma = g_m - (\omega C_{be})^2(R_{cs}+R_T),\ \delta = \omega[g_m C_{be}(R_{cs}+R_T)+C_{be}],$$
$$\varphi = \frac{{g_m}^3 R_{CS}}{\gamma^2+\delta^2}$$

Fig. 6 shows the typical variation of inductance and series loss resistance with the applied tuning voltage V_T at 8 GHz.

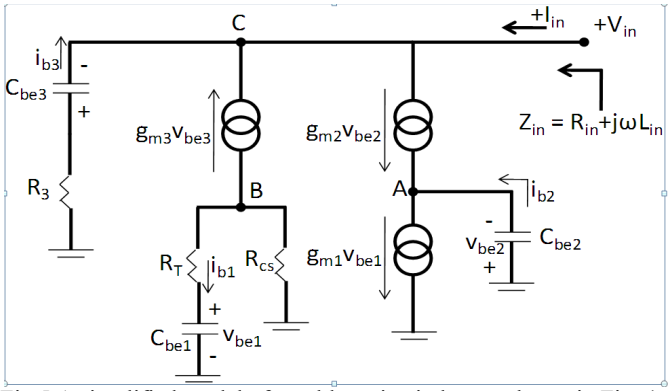


Fig.5 A simplified model of tunable active inductor, shown in Fig. 4

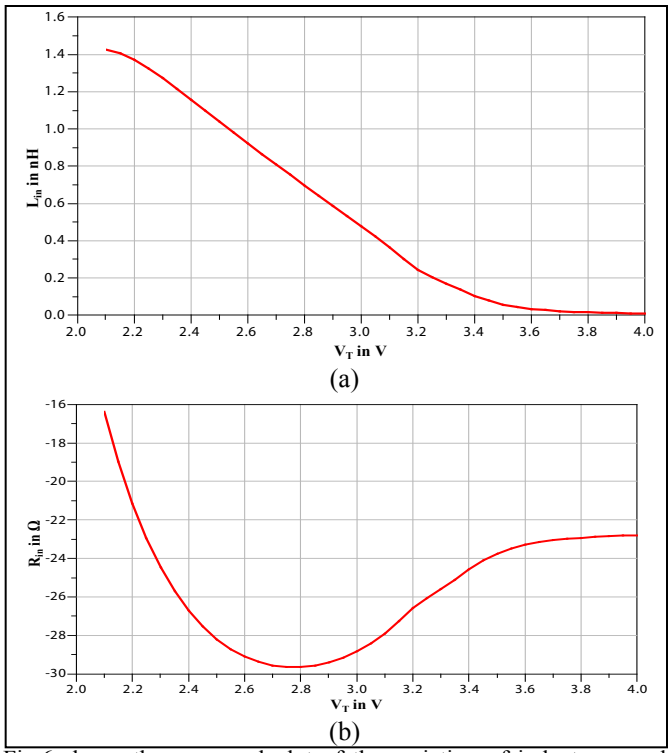


Fig.6 shows the measured plot of the variation of inductance and series resistance at 8 GHz for the tunable active inductor shown in Fig. 4. : (a) variation of inductance with respect to tuning voltage, and (b) shows the measured plot of the variation series loss resistance with respect to tuning voltage

III. Multi-Mode Multi-Band TAIO

Fig. 7 shows the equivalent representation of tunable active inductor oscillator (TAI) circuit. The steady-state oscillation condition for series feedback active inductor oscillator as shown in Fig. 7 is given by

$$Z_{osc}(I_L,\omega)+Z_L(\omega)=0 \Rightarrow Z_{osc}(I_L,\omega)=R_{osc}(I_L,\omega)+jX_{osc}(I_L,\omega) \quad (8)$$

$$R_{osc} = \left[\frac{\frac{1}{R_2}}{\left(\left(\frac{1}{\omega L_2}-\omega C_2\right)^2+\left(\frac{1}{R_2}\right)^2\right)}\right] - \left[\frac{-gm\left(\left(1-\omega^2 C_{b'e}L_1\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right)\left(\left(\frac{1}{\omega L_2}-\omega C_2\right)^2+\left(\frac{1}{R_2}\right)^2\right)-\frac{\omega^2 C_{b'e}C_{ce}}{R_2}\left(2\left(1-\omega^2 C_{b'e}L_1\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right)}{\omega C_{ce}\left(\left(\frac{1}{\omega L_2}-\omega C_2\right)^2+\left(\frac{1}{R_2}\right)^2\right)\left(\left(\left(1-\omega^2 C_{b'e}L_1\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right)^2+\left(\frac{1-\omega^2 C_{b'e}L_1}{R_2}\right)^2\right)}\right] \quad (9)$$

$$X_{osc} = \left[\frac{\left[\frac{-1}{\omega C_{ce}}\left(\frac{1}{\omega L_2}-\omega C_2\right)^2+\left(\frac{1}{\omega L_2}-\omega C_2\right)-\frac{1}{\omega C_{ce}}\left(\frac{1}{R_2}\right)^2\right]\left[\left(\left(1-\omega^2 C_{b'e}L_1\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right)^2+\left(\frac{1-\omega^2 C_{b'e}L_1}{R_2}\right)^2\right]}{\left[\left(\frac{1}{\omega L_2}-\omega C_2\right)^2+\left(\frac{1}{R_2}\right)^2\right]\left[\left(\left(1-\omega^2 C_{b'e}L_1\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right)^2+\left(\frac{1-\omega^2 C_{b'e}L_1}{R_2}\right)^2\right]}\right] + \left[\frac{\omega C_{b'e}\left(\frac{1}{\omega L_2}-\omega C_2\right)\left[\left(1-\omega^2 C_{b'e}L_1\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right]+\frac{\left(1-\omega^2 C_{b'e}L_1\right)}{\omega R_2 C_{ce}}g_m\left[\left(\frac{1}{\omega L_2}-\omega C_2\right)^2-\left(\frac{1}{R_2}\right)^2\right]-\frac{\left(1-\omega^2 C_{b'e}L_1\right)}{R_2}\frac{\omega C_{b'e}}{R_2}}{\left[\left(\frac{1}{\omega L_2}-\omega C_2\right)^2+\left(\frac{1}{R_2}\right)^2\right]\left[\left(\left(1-\omega^2 C_{b'e}L\right)\left(\frac{1}{\omega L_2}-\omega C_2\right)-\omega C_{b'e}\right)^2+\left(\frac{\left(1-\omega^2 C_{b'e}L_1\right)}{R_2}\right)^2\right]}\right] \quad (10)$$

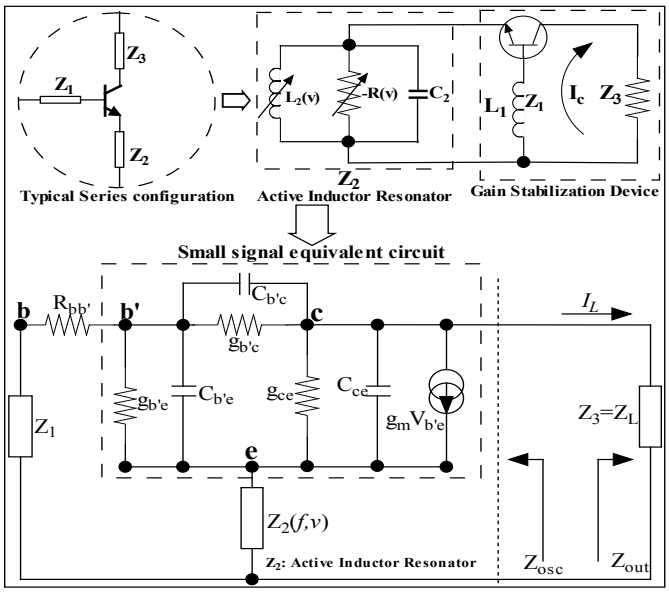


Fig. 7. A typical series feedback active inductor oscillator topology

For sustained oscillation $X_{osc} = 0$, (10) reduces to possible modes of oscillations:

$$\omega_1=\left[\frac{1}{L_2C_2}\right]^{1/2},\quad \omega_2=\left[\frac{1}{L_1C_{b'e}}\right]^{1/2},\quad \omega_3=\frac{1}{C_{b'e}C_{ce}}\left[2C_{b'e}C_{ce}\left(\frac{g_m}{R_2}\right)-\left(\frac{g_m}{R_2}\right)^2-8C_{b'e}^2\left(\frac{1}{R_2}\right)^2\right]^{1/2} \quad (11)$$

For dominant mode oscillation condition, negative dynamic resistance associated with R_{osc} can be described by solving (7) and (9), and the value of R_{osc} can be optimized in such a way that only desired fundamental mode (ω_1) of oscillation sustains the growth under the constraints of not enough gain supported by R_{osc} for other degenerate modes (ω_2 and ω_3). Fig 8 shows the typical simplified schematic of TAI oscillator of n^{th} order that can be reconfigurable with respect to multi-mode and multi-band wireless systems. As shown in Fig 8, the topology of multi-mode multi-band active inductor VCO is sensitive to mode-jumping and jitter and exhibits inferior phase noise performance in comparison to commercially available spiral inductor LC resonator based oscillator topology for a given DC power consumptions and figure-of-merit (FOM).

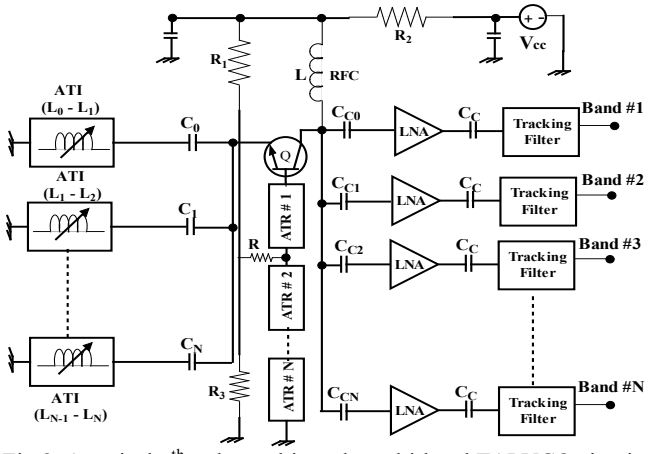


Fig.8. A typical n[th] order multi-mode multi-band TAI VCO circuits

In addition to this 1/*f* noise associated with active that forms active inductor under large signal drive level can become the bottleneck and restricts the applications. The novel and cost-effective approach is to incorporate self-injection-locking mechanism for improving the noise dynamics of inductor-tuned oscillator circuit for a given FOM. This can be realized by incorporating phase shifter network in the feedback path that improves the 1/*f* noise, including linearization of the large signal drive-level characteristics of the realized inductor as shown Fig. 4.

The goal for improving 1/*f* noise characteristics is paramount for enabling seamless transition from varactor-tuned LC resonator oscillator topology to inductor-tuned oscillator wireless applications. Nevertheless, the advantages of integration in small chip area (typically less than 10 % area of equivalent spiral inductor based RFIC VCO) make inductor-tuned VCO technology as a cost-effective alternative for RFIC/MMIC VCOs solutions for current and later generation of wireless systems [1]-[12].

IV. Examples And Validation

Fig.9 shows the typical schematic of self-injection-locked inductor-tuned Colpitts Oscillator realized by incorporating phase shifter network in the feedback path that improves the 1/*f* noise, including linearization of the large signal drive-level characteristics of the simulated realized inductor.

Fig.10 shows the CAD simulated phase noise plots with and without injection-locked inductor-tuned oscillators for giving brief insights about the improvement in phase noise performance. It can be seen from the phase noise plots (Fig. 10) that the phase noise of the inductor-tuned oscillator without injection locking is approximately -41dBc/Hz @ 10 kHz offset from 1.9 GHz the carrier frequency. When the inductor-tuned oscillator is locked, the phase noise drops to - 58dBc/Hz @ 10 kHz offset from the carrier frequency.

As shown in Fig. 9, self-injection locking mechanism improves phase noise typically 17 dB @ 100 kHz and 10 dB @ 1 MHz offset from carrier frequency 1.9 GHz. This observation is justified as injection-locking is more effective at lower offset (less than 100 kHz) and starts to degrade when the offset frequency exceeds 1 MHz due to broadband thermal noise trades with locking range for a given strength of injection level and phase shifts.

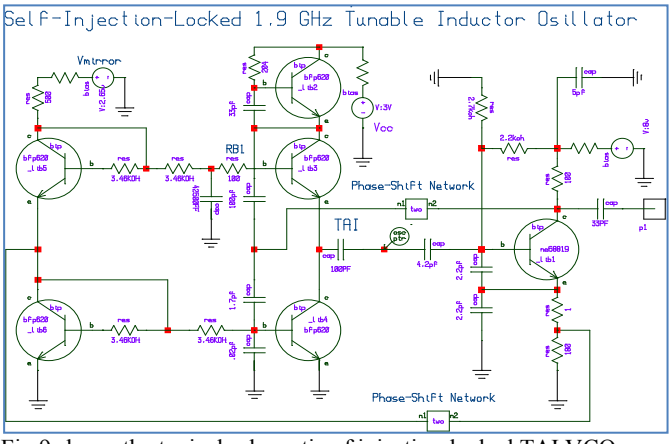


Fig.9 shows the typical schematic of injection-locked TAI VCO

Fig.11 shows the measured phase noise plots of injection-locked inductor-tuned oscillator circuit as shown in Fig. 9. It can be seen from the phase noise plot shown in Fig. 11 that the measured plot shows superior performance and vary more than 10 dB as compared to CAD simulated phase noise plot (Fig. 10), reason being unknown package parasitic and convergence problem (CAD tools) in presence of injection-locking signal which is fed through delay line phase shifter network into realized inductor circuit.

For validation of this approach, triple-band (0.9-1.8/1.8-3.6/3.6-6.9 GHz) TAIO is fabricated using SiGe HBTs on PCB substrate with ε_r of 3.38 and a thickness of 22mils (microstripline), however, this topology is amenable for integration in chip form.

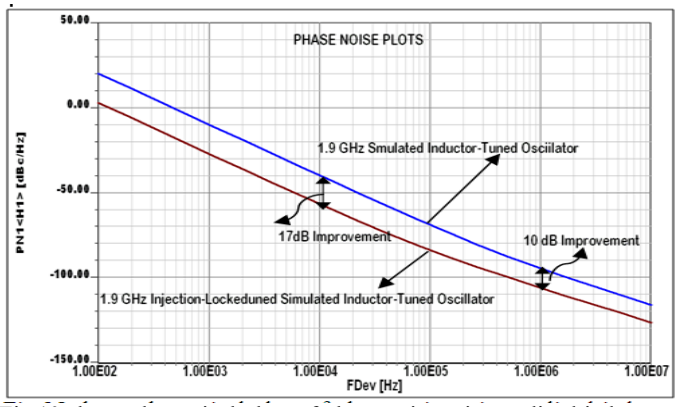


Fig.10 shows the typical plots of phase noise using realized inductor and injection-locked realized inductor.

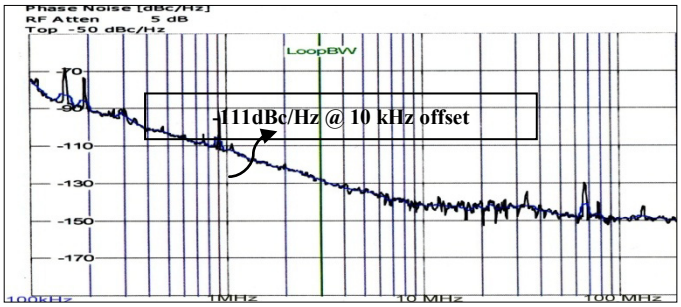


Fig.11 shows the measured phase noise plots of injection-locked inductor-tuned oscillator circuit as shown in Fig. 9.

The broadband injection-locking scheme shown in Fig. 12 can be realized by phase shifter network in conjunction with varying the realized inductance for optimum drive-level and dynamic range. The locking range of the injection-locked inductor-tuned oscillator can be described in term of maximum change in frequency for a given operating conditions. The frequency of TAIO (Fig. 12) is tuned by varying the inductance (Fig.4) as

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \rightarrow f_0 + \Delta f = \frac{1}{\sqrt{(L+\Delta L)C}} \approx f_0\left[1 - \frac{1}{2}\frac{|\Delta L|}{L}\right]_{\Delta L \ll L} \quad (12)$$

$$|\Delta f| \cong f_0\left[\frac{1}{2}\frac{\Delta L}{L}\right]_{\Delta L \ll L} \quad (13)$$

where ΔL and Δf, is the maximum change of the inductance of the inductor (shown in Fig. 4) due to injection-locking signal (realized using delay-line phase shifter as shown in Fig. 12), and maximum frequency shifts that translates to locking range of the inductor-tuned oscillator.

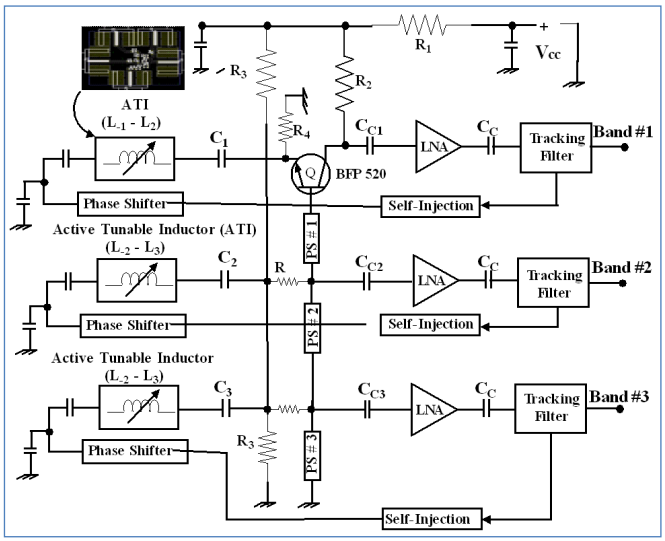


Fig.12 schematic of self-injection locked triple-band TAIO

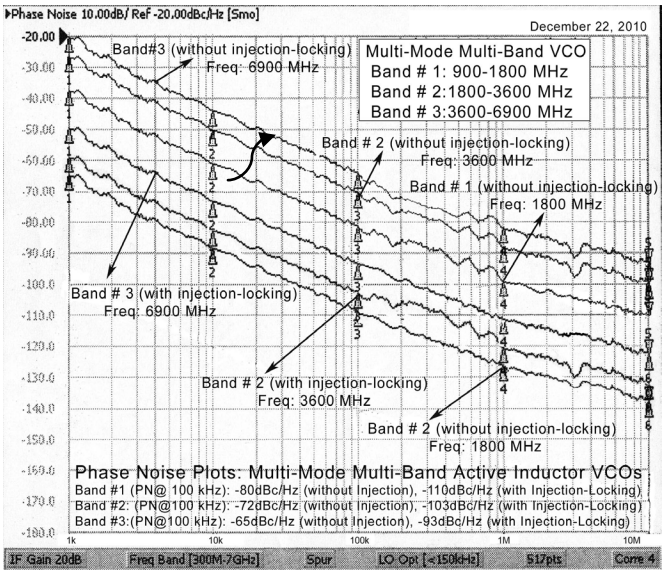


Fig.13 Measured phase noise plots of SiGe HBTs TAIO (Fig. 4)

From (13), the locking range is directly proportional to the maximum amount of the change of the inductance caused by the injected locking signal. For a varactor-tuned injection-locked LC oscillator, the locking range can be described by

$$|\Delta f| \cong f_0 \left[\frac{1}{2}\frac{|\Delta C|}{C}\right]_{\Delta C \ll C} \qquad (14)$$

where ΔC is the maximum amount of the change of the capacitance of the varactor caused by locking signal.

For a given oscillator circuit topology, operating condition, and figure-of-merit, it is important to note that the maximum change of ΔL is higher than ΔC and can be given by

$$\left[\frac{\Delta L}{L}\right]_{\Delta L \ll L} > \left[\frac{|\Delta C|}{C}\right]_{\Delta C \ll C} \qquad (15)$$

From (15), injection-locked inductor-tuned oscillator has a superior ability to adjust and coarse its frequency to follow the locking signal, resulting comparatively higher locking range as compared to varactor-tuned oscillator circuits.

The reason for higher locking range is less impact of injection signal on voltage output nodes of the varactor-tuned oscillator, resulting insignificant change in the capacitance (ΔC) of the varactor devices. On the contrary, inductance of the inductor-tuned oscillator is more sensitive to injected signal, and therefore, significantly varies current of the active device that forms the inductor circuit (Fig. 4). Fig.13 shows the measured phase noise plots with and without injection-locked inductor-tuned oscillators for giving brief insights about improvement in phase noise performance. It can be seen from Fig. 13 that the self-injection locked TAIO offers 15-25 dB better phase noise performance. The measured phase noise plots shown in Fig. 13 for band #1 (0.9-1.8 GHz), band #2 (1.8-3.6 GHz), and band #3 (3.6-6.9 GHz) are typically better than –110 dBc/Hz, -105 dBc/Hz, and -93 dBc/Hz respectively at 100 kHz offset from the respective carrier frequencies. The circuit works on 2V, 18 mA and gives O/P power typically better than -3dBm over respective bands.

The emerging applications of TAI topology continue to be unfolded along with the inception of new design techniques for improving phase noise and dynamic ranges. In spite of the drawbacks such as noise and linearity, active inductors have a significant advantage over passive inductors in terms of die-area, Q-factor, and issues regarding EMI and floor planning.

V. Conclusion

A design strategy centered on improving the locking range and broadband tunability of the realized inductor using SiGe HBTs is discussed, and validated through an example of triple-band injection-locked inductor-tuned oscillator circuit for multi-band multi-standard RF communication systems.

References

[1] S. Re, G. Leuzzi and V. Stornelli, “A new approach to the design of high dynamic range tunable active inductors”, INMMWC, 2008, pp. 25-28.

[2] S. Angkititrakul, H. Hu, and Z. Liang, “Active inductor current balnacing for interleaving multi-phase buck-boost converter”, IEEE APEC 2009, pp. 527-532,

[3] H-Hung Hsieh et.al., “A compact quadrature hybrid MMIC using CMOS active inductors”,IEEE Trans. MTT,Vol.55,pp.1098-1104 June 2007.

[4] R. Mukohopadhyay, et.al, “Investigation of Inductors for Digital Si-CMOS Techs”, pp. 3750-3753, ISCAS 2006.

[5] F. Yuan,“ CMOS active inductor and transformers, principle, implementation and application”, 2008 Springer Science.

[6] A. Tang et al., “ A new constant-Q active inductor with applications in low-noise oscillators”, AICS Proc., Vol. 58, No. 1 pp. 77-80, Jan 2009.

[7] S. Jang, C. Lin, C. Liu, and M. Junag, “ An active-inductor injection locked divider with varaible divisio ratio”, IEEE MW Letters, Vol. 19, No.1 pp. 39-41, Jan 2009.

[8] Y. Zheng and C. E. Saavendra, “Frequency response comparison of two common active circuits”, Progress in EM research letters, Vol. 13, pp. 113-119, 2010.

[9] Y. Zhou and F. Yuan, “A comparative study of lock range of injection-locked active-inductor oscillators”, IEEE MWSCAS 2010, pp. 973-976.

[10] L. Lu, H. Hsieh, Yu Liao, “A wide tuning-range CMOS VCO with a differential tunable active inductor”, IEEE Trans. MTT, vol. 54, pp.3462-3458, Sept 2006.

[11] Z. Safarian and H. Hashemi, “Wideband Multi-Mode CMOS VCO Design Using Coupled Mode Inductors”, IEEE Trans. CS-I, pp. 1830-1843, Aug 2009.

[12] Xiao H. Design of radio-frequency filters and oscillators in deep submicron CMOS technology, Dissertation Portland State Univ. 2008.

A 4.596 GHz Frequency Synthesis Based on a Solid-Mounted Resonator.

Rodolphe Boudot
Vincent Giordano
FEMTO-ST Institute
Time-Frequency Dpt
Besançon, France
Email: rodolphe.boudot@femto-st.fr

Ming Dong Li
Nathalie Rolland
Paul-Alain Rolland
IEMN, CNRS and Universite de Lille 1
Lille, France

Pierre Vincent
CEA-LETI
Grenoble, France.

***Abstract*— We report a 4.596 GHz frequency synthesis based on a 2.1 GHz Solid Mounted Resonator (SMR) voltage controlled oscillator (VCO). The SMR source exhibits a power consumption of about 18 mW, a chip size of 2mm^2 and a phase noise of − 89 dBc/Hz at 2 kHz offset frequency. A synthesizer is constructed from this source to generate a low noise 4.596 GHz signal. The synthesizer is preliminary used as a local oscillator in a compact Cs atomic clock based on Coherent Population Trapping (CPT).**

I. Introduction

Chip Scale Atomic Clocks (CSAC) based on CPT are passive atomic frequency standards and need a high-performance local oscillator (LO) combining an extremely reduced volume, power consumption, frequency-temperature dependence as well a low phase noise not to degrade the clock short term frequency stability [1].

In general, LOs proposed for CSACs are frequency-multiplied low power consumption temperature-controlled (TCXO) or microprocessor-controlled (MCXO) crystal oscillators. Nevertheless, it can be useful to avoid this frequency multiplication process degrading the source phase noise and power consumption. In that sense, Brannon et al. proposed the development for CSAC applications of microwave oscillators based on quarter-wavelength ceramic-filled coaxial resonators [2]. In another way, a great challenge is to realize fully-packaged MEMS oscillators based on bulk acoustic wave resonators (BAW) due to their small size and high Q-factor [3]. Among them, Film Bulk Acoustic Resonators (FBAR), High-overtone Bulk Acoustic Resonators (HBAR) and Solid Mounted Resonators (SMR) are potential candidates. Romisch and Lutwak developed a low power 4.596 GHz Colpitts oscillator for Cs CSAC applications [4]. Yu et al. demonstrated a HBAR-based 3.6 GHz Pierce oscillator with a phase noise of -67 dBc/Hz at 300 Hz fourier frequency, a power consumption of 3.2 mW and a free-running frequency stability of 1.5×10^{-9} at 1 second [5]. A great issue in such oscillators is their high thermal frequency sensitivity. Several solutions were proposed to reduce the frequency-temperature dependence of FBAR [6], [7] and HBAR [8].

This paper presents an original SMR oscillator-based 4.596 GHz frequency synthesis. Such sources have been proposed recently [9]. Section II describes the 2.1 GHz SMR VCO. Section III presents the frequency synthesis to generate the 4.596 GHz signal. Section IV shows the successful detection of CPT resonances using the 4.596 GHz synthesized signal.

II. The SMR-based voltage controlled oscillator

Fig 1 shows the schematic structure of the Solid Mounted Resonator (SMR).

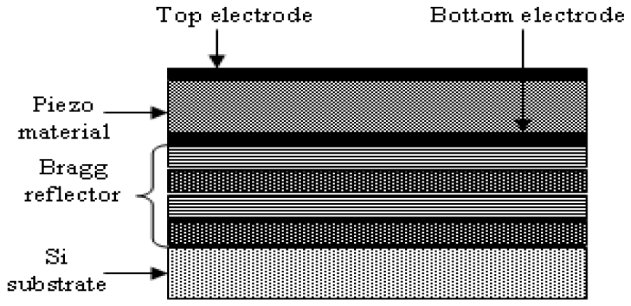


Fig. 1. *Cross-section of the Solid Mounted Resonator.*

It consists of a metal-piezo (Mo-AlN)-metal layered sandwich. The resonator is constructed over a Bragg reflector acoustic mirror structure using alternating layers of low and high acoustic velocity. The Bragg reflector is made of Tungsten (W) and silicon oxide (SiO_2). The so-called "frame" technique [10] is used to confine the acoustic energy in the useful thickness mode.

A quality factor value of 1176 at the series resonance frequency f_s of 2.083GHz is obtained. The Q-factor of the parallel resonance frequency f_p is 502 at 2.125GHz.

The VCO is based on a common-base configuration architecture [11]. The active circuit is a SiGe:C heterojunction bipolar transistor (HBT) using BiCMOS7RF 0.25μm technology processes from ST Microelectronics.

The fine frequency tuning of the oscillator is achieved by an adjustable voltage control supply added to the bias voltage. The coarse frequency tuning is achieved by a DDS circuit. The core oscillator uses a hybrid method by wire-bonding

978-1-61284-110-6/11/$26.00 ©2011 IEEE

the BiCMOS chip (1mm x 1 mm) and the Solid Mounted Resonator (0.35mm x 0.3mm).
The oscillation frequency variation as a function of the voltage tuning is 260 kHz/V. The measured output power variation is 0.5dBm/V.
Fig. 2 reports the variation of the VCO output frequency and power versus the resonator temperature.

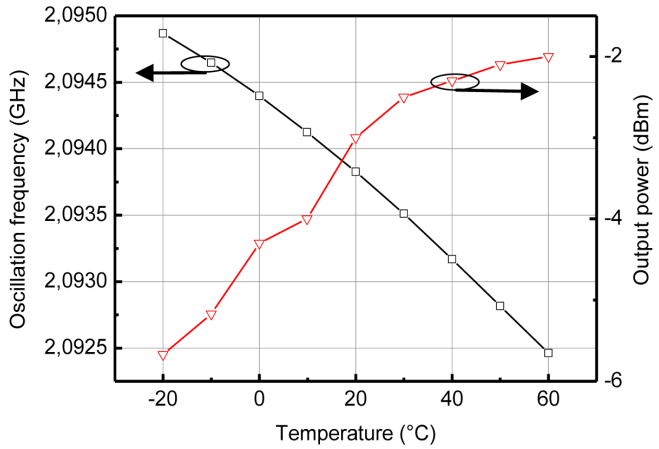


Fig. 2. *Oscillation frequency and output power of the SMR VCO as a function of the temperature.*

The frequency-temperature dependence is found to be -14 ppm/°C. The variation of the VCO output power versus the temperature is 3.6 dBm in the -20-60°C temperature range.

Fig. 3 shows the phase noise performances of the SMR VCO.

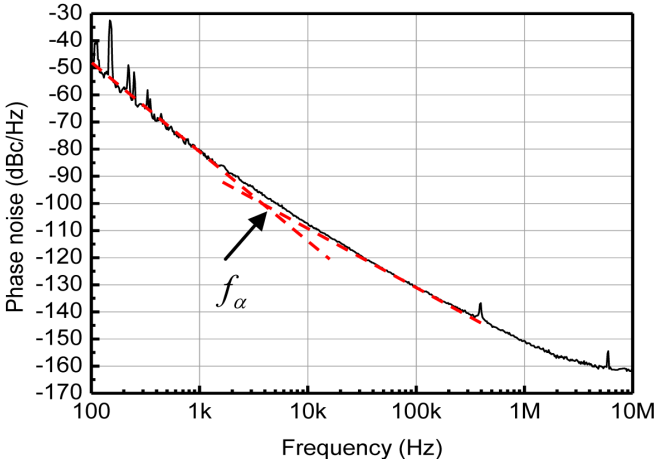


Fig. 3. *Phase noise of the SMR VCO.*

The VCO phase noise is -160 dBc/Hz for $f > 3.5$MHz and is -82 dBc/Hz at 1 kHz offset frequency. A very slight degradation of the VCO phase noise, lower than 1.5 dB, is obtained at f=2 kHz when the voltage control is increased while no significant variation is observed at f= 100 kHz.

III. THE 4.596 GHZ FREQUENCY SYNTHESIS

The description of the 4.596 GHz frequency synthesis based on the 2.1 GHz SMR VCO is described in [11]. The generation of the 4.596 GHz signal is realized by first frequency-doubling the VCO frequency at 4.2 GHz. Simultaneously, a signal of 396 MHz frequency is generated by mixing (mixer M1) the 384 MHz and 12 MHz signals from two Direct Digital Synthesis (DDS) DDS2 and DDS1 respectively. Eventually, the 4.596 GHz signal is obtained by mixing the 4.2 GHz and the 396 MHz signals using the mixer M2.
A synchronization clock signal (77MHz) for DDS and a FPGA is delivered by tracking and frequency dividing the SMR oscillator signal. A Frequency Shift Keying (FSK) modulation at a bit rate of 1kHz and frequency deviation of 3kHz is generated. This modulation is used to lock the final VCO output frequency on the atomic transition frequency. A useful 9.3 MHz RF signal is extracted from the synthesis to characterize easily the atomic clock frequency stability when it is locked. The output power is measured to be about 0 dBm at 4.596 GHz.
Fig. 4 reports phase noise performances of the free-running 2.1 GHz SMR VCO, of the DDS2 and the 4.596 GHz frequency synthesis output. The frequency synthesis phase

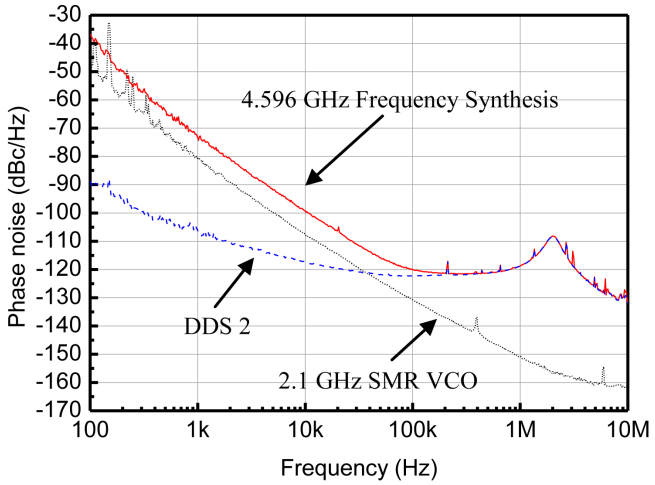


Fig. 4. *Phase noise performances of the 2.1 GHz SMR VCO, of the DDS2 and of the 4.596 GHz output signal of the frequency synthesizer.*

noise is clearly limited by the DDS2 phase noise at a level of -120 dBc/Hz and -132dBc/Hz at 1MHz and 10 MHz offset frequencies respectively for $f > 150$kHz. The phase noise of the 4.596 GHz signal is -76 dBc/Hz and -100 dBc/Hz at f = 1kHz and 10 kHz respectively.

IV. APPLICATIONS TO A MINIATURE CPT ATOMIC CLOCK

The SMR-VCO-based frequency synthesis is used as a local oscillator in a compact CPT atomic clock described in [11]. The CPT resonance was successfully detected. Fig. 5 shows the Allan deviation of the free-running oscillator and locked oscillator.

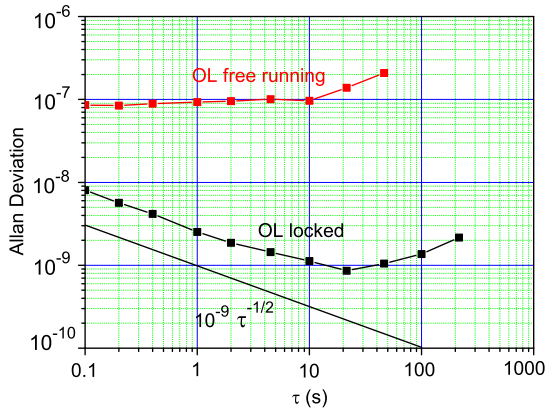


Fig. 5. *Allan deviation of the SMR VCO in free-running or locked configurations.*

In the free-running regime, it is measured to be 9.5×10^{-8} at 1 s integration time and increases after 10 s averaging time. In the locked regime, the BAW oscillator signal fractional frequency instability is measured to be $2.5 \times 10^{-9}\ \tau^{-1/2}$for integration times up to 20 s. This represents an improvement by a factor 38 and 83 compared to the unlocked regime at 1 s and 10 s respectively. In the present case, the clock short term frequency stability limitation is mainly attributed to huge temperature variations in the laboratory room affecting the SMR VCO free-running frequency through its frequency-temperature dependence as well as a clock frequency-locked loop (FLL) bandwidth too small. Unfortunately, this result could not be optimized because of time constraints.

V. CONCLUSIONS

A 2.1 GHz SMR voltage-controlled oscillator with a phase noise of $-$ 89dBc/Hz at f=2 kHz and $-$131dBc/Hz at f=100kHz was presented. The frequency-temperature dependence is $-$14ppm/°C over a range of $-$20°C to 60°C while its power consumption is 18 mW. An original analog-digital 4.596 GHz frequency synthesizer based on the designed SMR VCO was developed. The SMR VCO-based frequency synthesizer was preliminary successfully locked to the CPT resonance of Cs atoms in a mm-scale vapor cell.

ACKNOWLEDGMENT

This work was supported by ANR-MHAST. The authors would like to thank B. Verbeke and C. Begue for their help in electronics.

REFERENCES

[1] C. Audoin, V. Candelier and N. Dimarcq, *IEEE Trans. Instr. Meas.* **40**, (2), 121 (1991).
[2] A. Brannon, J. Breitbarth and Z. Popovic, *Proc. IEEE MTT Symp.* Long Beach, CA, 1535–1538 (2005).
[3] K. M. Lakin, *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.* **52**, (5), 707 (2005).
[4] S. Romisch and R. Lutwak, *Proc. IEEE Int. Freq. Contr. Symp.*, Miami, Fl, pp. 448-451 (2006).
[5] H. Yu, C-Y. Lee, W. Pang, H. Zhang, A. Brannon, J. Kitching and E. S. Kim, *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.* **56**, (2), 400 (2009).
[6] W. Pang, Y. Hongyu, Z. Hao, K. Eun Sok, *IEEE MTT Symp. International Microwave Symposium Digest.* Long Beach, CA, 1279 (2005).
[7] H. Yu, W. Pang, H. Zhang and E. S. Kim, *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.* **54**, (10), 2102 (2007).
[8] T. Baron, D. Gachon, G. Martin, S. Alzuaga, D. Hermelin, J. P. Romand and S. Ballandras, *IEEE International Frequency Control Symposium*, Newport Beach, CA, 652 (2010).
[9] S. Dossou, N. Abele, E. Cesar, P. Ancey, J.F. Carpentier, P. Vincent and J-M. Fournier, *IEEE International Symp. Circ. and Syst.*, Seattle, Wa, 1456 (2008).
[10] R. Thalhammer and R. Aigner, *IEEE MTT Symp. International Microwave Symposium Digest*, Long Beach, CA, 4pp (2005).
[11] R. Boudot, M. D. Li, V. Girodano, N. Rolland, P. A. Rolland and P. Vincent, *Rev. Sci. Instr.* **82**, 034706 (2011).

Adaptive Mode-Coupled Harmonically Tuned Ultra Low Phase Noise VCSO Circuits

Ajay K. Poddar[1,2], *Senior Member IEEE*
[1]Synergy Microwave Corp., NJ 07504, USA
[2]Technical University Munich, Germany

Ulrich L. Rohde[1,2], *Fellow IEEE*
[1]Univ. of Cottbus, BTU Cottbus 03046, Germany
[2]Synergy Microwave Corp., NJ 07504, USA

Abstract— **Adaptive mode-coupled harmonically tuned voltage-controlled SAW oscillators (VCSOs) are versatile performers in phase-locked reference sources, can be supplied in standard and custom frequencies through 5 GHz at a fraction of the size and power consumption of multiplied references. The typical measured phase noise @ 10 kHz offset from the 5 GHz carrier is -140 dBc/ Hz with 500 kHz tuning and RMS jitter less than 10 femto seconds. To author's knowledge, the reported 5 GHz VCSO offers about 10 dB superior phase noise performances as compared to commercially available SAW and crystal multiplied sources for a given size, cost, and power-consumptions.**

I. INTRODUCTION

Reference signal sources set the spectral standards for a wide range of phase-locked-loop (PLL) applications, from clock translators in instrumentation to radar systems. Traditional references sources, such as crystal, ceramic, or surface-acoustic-wave SAW) oscillators, are based on high-quality-factor (high-Q) resonators [1]-[48]. They are generally limited to about 3 GHz and must be multiplied to provide higher reference frequencies. And with multiplication comes degradation in phase-noise, power consumptions, and harmonic performance. In this paper, degenerated adaptive mode-coupling and regenerative noise filtering is being incorporated to achieve excellent spectral performance at fundamental frequencies 3 GHz and above.

In contrast to other high-frequency reference-source solutions, such as crystal-oscillator-based references which employ oscillators, gain blocks, and filtering networks that add to high development costs and reliability concerns, the reported adaptive mode-coupled harmonically tuned VCSO circuits provide simpler, more reliable solutions that save space and power consumption. The novel approach includes a methodology to enhance the mode-injection locking range, and to reduce or eliminate the amount of filtering needed to suppress sub-harmonics and higher-order harmonic products. The approach also reduces the susceptibility to microphonics while retaining low phase noise and a moderate tuning range to compensate for short-term and long-term effects of thermal drift and aging [3]-[11]. The reported reference sources are ideal for use as clock translators for high-speed analog-to-digital converters (ADCs) and direct-digital-synthesizer (DDS) clocks, offer superior figure-of-merit (FOM) in terms of performance, price, and product delivery to satisfy both the technical and current business needs of designers and buyers. So, whether oscillators are used at their specific frequency or tuned to higher frequencies using multipliers or PLLs, the need for accurate, affordable, low-phase noise and stable performance never changes.

II. VCSO TOPOLOGIES

VCSOs (Voltage Controlled SAW Oscillators) are widely used in wireless applications, since the technology features very low phase noise at fixed frequencies through about 1 GHz. However, VCSO circuit operating above 1 GHz and above is generally realized using multiplier approach which is power hungry, large size, and offers degraded phase noise performances. Historically synthesizer developers have relied on crystal based multiplied reference oscillator featuring excellent phase noise performance at 3 GHz and above frequencies. Although, frequency multiplier crystal oscillator can offer low phase noise performance at 3 GHz and above frequencies but at the cost of large size, high noise floor and high power-consumptions. The most dramatic change will occur when optimum loop bandwidth extends to a few megahertz, where reported adaptive mode-coupled harmonically tuned VCSOs noise performance supersede the performance of the conventional SAW and multiplied crystal oscillator at 3 GHz and above.

Furthermore, these characteristics, accompanied with reduced size, power consumption, cost, drift, and jitter inherent to the VCSO topology shown in Fig.1, is likely to secure their domination in the foreseeable future at higher frequency synthesizers and reference frequency standards for the applications in current and 4th generation communication systems.

Fig. 1 shows the typical block diagram of the novel VCSO using adaptive mode-coupled harmonically tuned mechanism with more than 500 kHz tuning for compensating the thermal drift without using heating device, and fabricated in tiny SMD packages (0.5x0.5 inches).

The typical measured phase noise @ 10 kHz offset from the 5 GHz carrier is -140 dBc/ Hz, with RMS jitter less than 10 femto seconds (Fig.2). The reported 5 GHz VCSO offers about 5-10 dB superior phase noise performances as compared to commercially available SAW and crystal multiplied sources for a given size, cost, and power-consumptions.

III. Multiplied Frquency Crystal Oscillator Circuit

As briefly discussed above, designers have often relied on multiplying a low-noise crystal oscillator to achieve higher frequency 1 GHz and higher output signals. The harmonic frequency crystal oscillator approach can provide excellent phase noise, but tend to consume a great deal of power, generate high levels of noise, and are large in size. Commercially available crystal oscillators at 100 MHz and below are available with phase noise of -160 to -176 dBc/Hz offset 20 to 100 kHz from a 100-MHz output.

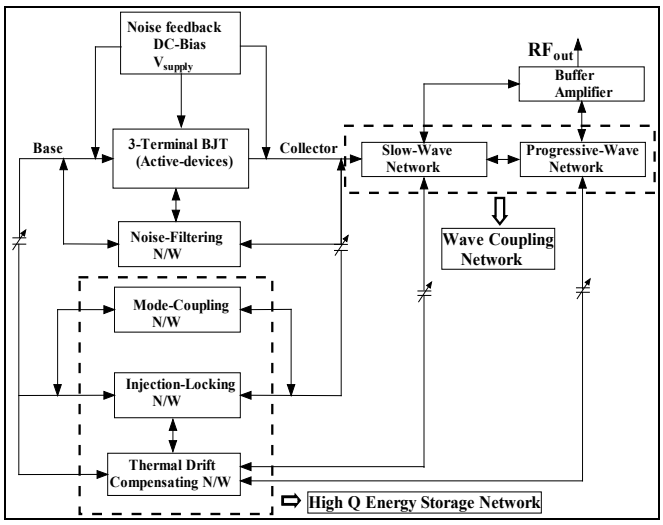


Fig.1. This block diagram shows the functional parts of an adaptive mode-coupled harmonically tuned VCSO for low cost high frequency sources.

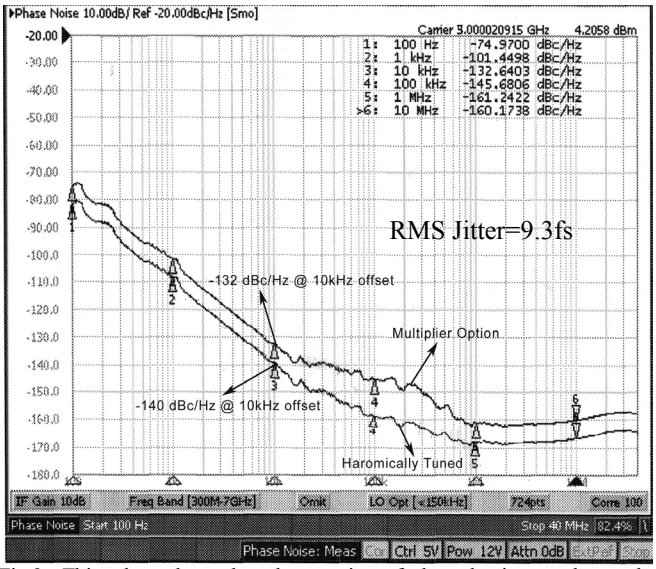


Fig.2. This plots show the phase noise of the adaptive mode-coupled harmonically tuned VCSO was measured at a carrier frequency of 5 GHz and with oscillator bias supply of 30 mA and +12 VDC.

When multiplied 10 times to obtain 1 GHz frequency source, the phase noise of crystal oscillators translates to -140 to -156 dBc/Hz offset 20 to 100 kHz from 1 GHz but can degrade the noise floor by as much of 20 dB.

The availability of low-cost, low-noise voltage controlled saw oscillator (VCSO) clears the way for small, low-noise synthesizers. Part of using these sources involves modeling the surface acoustic wave resonators under large signal drive conditions for the better insights about noise dynamics at close-in phase noise, then developing manufacturable methods for producing high-purity and temperature stable oscillators in chip form.

Fig. 3 compares fundamental 1 GHz VCSO phase noise at 1 GHz to that of a commercially available multiplied crystal oscillator circuit. It can be seen from Fig.3 that fundamental 1 GHz VCSO supersedes the phase noise performance as much as 20 dB at 100 kHz offset from the carrier as compared to commercially available multiplied 1GHz VCXO. The primary drawback of VCSOs is their inability to be tuned by means of tuning varactor diode, therefore, unable to compensate the frequency drifts due to ageing and changes in operating temperature for reference sources applications. Accordingly, there is a need for frequency agile SAW resonators for VCSO applications.

IV. Tunable SAW Resonator For VCSO Circuit

SAW (surface acoustic wave) devices are generally used in frequency ranges above that of the BAW (bulk acoustic wave) crystal such as the AT cut. The SAW involves only the surface of the substrate used. The frequency is set by the distance between the electrode "fingers" (Fig.4). Quartz, LiTaO3, and LiNbO3 are common substrates for such devices. Initial accuracy is dependent on the accuracy of the printing of the "fingers". Initial accuracy can be in the ±50 to 200 ppm range. The approximate temperature constant of quartz is 0.03 ppm/^{0}C² and -20ppm/^{0}C² for LiTaO3.

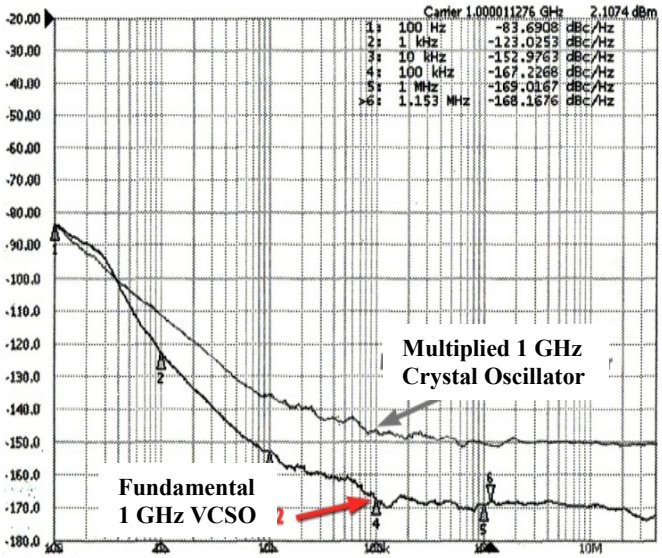


Fig.3. This phase noise plot compares the performance of a 1-GHz VCSO with 1-GHz achieved by multiplying a crystal oscillator.

Upper frequencies of 3.5 GHz are currently obtainable. The lower end is limited primarily by size, since the AT becomes more practical. Currently, surface acoustic wave (SAW) resonator based oscillators are used for high frequency signal sources, which have cost, availability, and frequency stability issues.

The shift to multi-band 3G phones has increased the demand of Surface Acoustic Wave (SAW) filters per handset. This implies a substantial increase in the filtering functions, thereby increasing the filter count in handsets. Combinations of multiband, multimode models are expected to use increasing number of SAWs. As more phones operate on multiple bands, manufacturers are expected to require more SAW filters, thereby increasing the overall SAW filter market size. The research service notes the following product segments: SAW filters, SAW resonators and SAW oscillators. SAW oscillators are widely used in wireless applications, since the technology features very low phase noise at fixed frequencies through about 3 GHz. SAW resonator oscillators are also known for their low microphonic noise (tolerance to vibration), high quality factor (Q), and low jitter.

Generally, it is highly desirable for reference frequency sources to exhibit ultra low phase noise characteristics. This is usually accomplished by using high Q quartz crystals and SAW resonator devices. Unfortunately, SAW resonator based oscillators have several disadvantages, including a limited operating temperature range and limited tuning range (which limits the amount of correction that can be made to compensate for the tolerances of other components in the oscillator circuitry). Coupling a varactor diode or other capacitive element to a high Q-factor frequency controlling element, such as SAW device, does not result in a tunable, low phase noise VCSO circuits. To overcome this limitation, VCSO have been implemented that couple together two or more SAW devices in order to achieve a wider tuning range and low phase noise.

The simplified approach is to reduce the quality factor of the SAW by introducing series or parallel loss resistors. However, the use of multiple SAW elements increases the complexity and the cost of the VCSO circuits [48]. Fig. 5 shows the arrangement of tunable SAW resonator network based on the design tradeoffs between the phase noise and the tuning characteristics of the VCSOs.

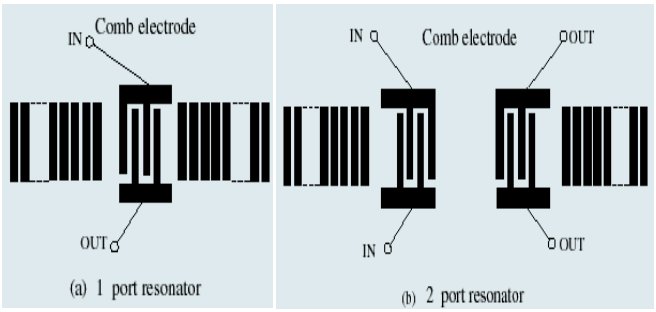


Fig.4 This shows the typical representation of surface acoustic wave resonator: (a) 1-port SAW Resonator, and (b) 2-port SAW resonator

Although, the frequency of VCSO can be changed over a fairly wide range by means of varying the capacitances (C_1, C_2, C_3, C_4, and C_5 as shown in Fig. 5), but the parameters of the oscillator circuit (Fig.6) are changing drastically.

This creates serious problems in terms of phase noise and tuning linearity. For low phase noise, it is necessary to keep the reactance slope as steep as possible by maximizing the phase derivative and group delay [$\partial\varphi/\partial\omega$] of the coupled SAW resonator network. For tuning linearity, it is necessary to keep the reactance slope as straight as possible over the desire operating frequency range. If the reactance curve is too shallow, VCSO (Fig. 6) circuit becomes susceptible to noise and stability perturbation. Furthermore, the straighter the reactance curve is, the greater is the tuning linearity in the desired frequency range.

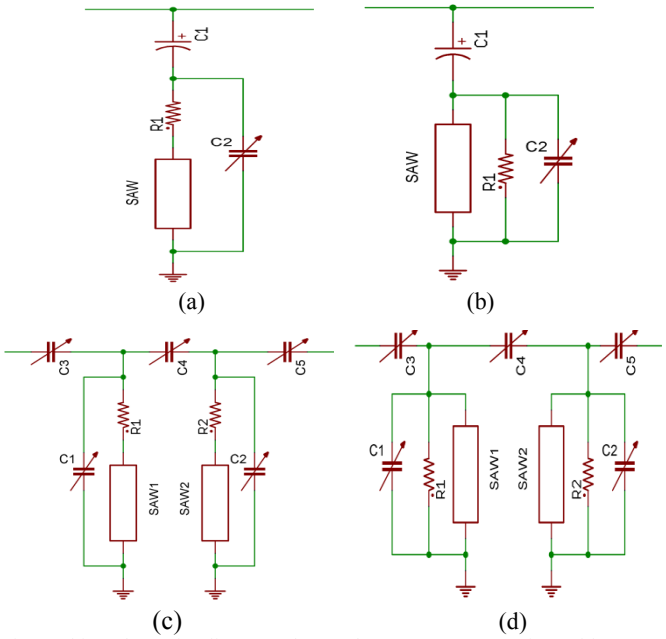


Fig.5 This schematic diagram shows the arrangements of tunable SAW resonator networks: (a) series resistance in conjunction with SAW resonator, (b) parallel series resistance in conjunction with SAW resonator, (c) series resistance in conjunction with capacitively coupled SAW resonator, and (d) parallel resistance in conjunction with capacitively coupled SAW resonator.

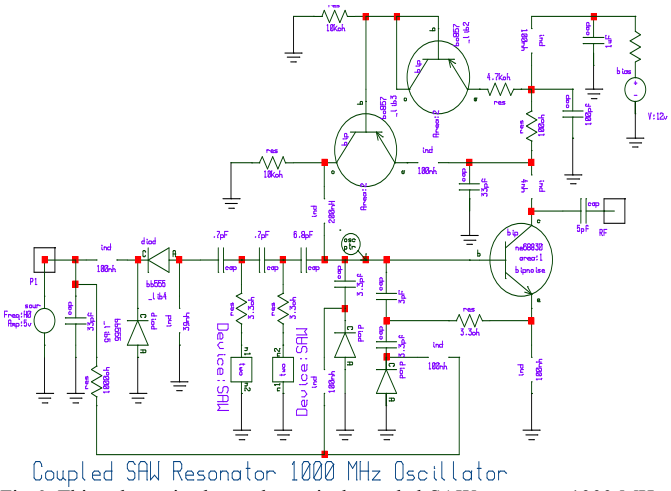


Fig.6. This schematic shows the typical coupled SAW resonator 1000 MHz oscillator

To overcome these problems, a novel technique is to incorporate additional inductor (L) parallel to the SAW resonator that resonates with the C_0 (the internal and package capacitance of SAW resonator) as shown in Fig. 7. Additionally, the value of inductor L2 and C (Fig.7) is adjusted in order to shift the resonance to that of intrinsic SAW resonator (without insertion of L and C) frequency (i.e. 1GHz). This enables the reactance curve linear and steep over the operating frequency range from 999.3-999.8 MHz (Fig.8).

An objective of this paper is to improve the tuning characteristics (on the order of +/-400 PPM) of the SAW resonator without compromising phase noise performance by cascading the mode-coupled harmonically tuned network to compensate the frequency drift caused due to change in operating temperature (-40 to +85 degree Celsius) (Fig.9).

As shown in Fig. 9, the cascaded mode-coupled harmonically tuned VCSO circuit minimizes the phase noise with more than 2 MHz tuning, resulting cost-effective solution. The typical measured phase noise is –152dBc/Hz at 10 kHz offset from the carrier with jitter less than 6fs, this is the reasonably low phase noise solution for this class of inexpensive VCSO solutions (Fig. 10).

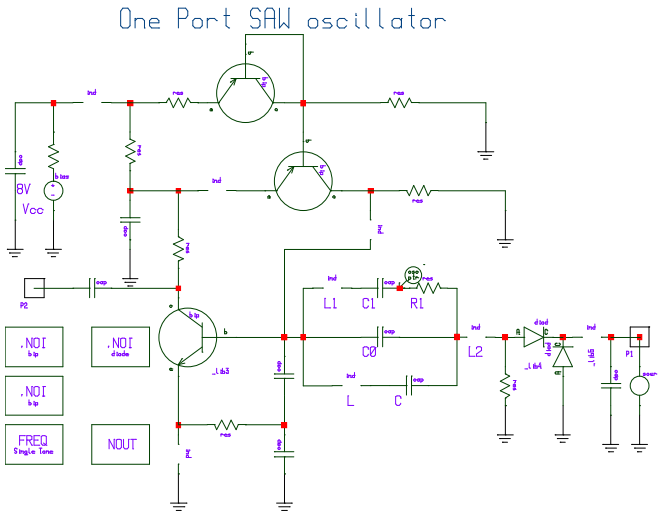


Fig.7. This schematic shows the typical tunable SAW resonator oscillator

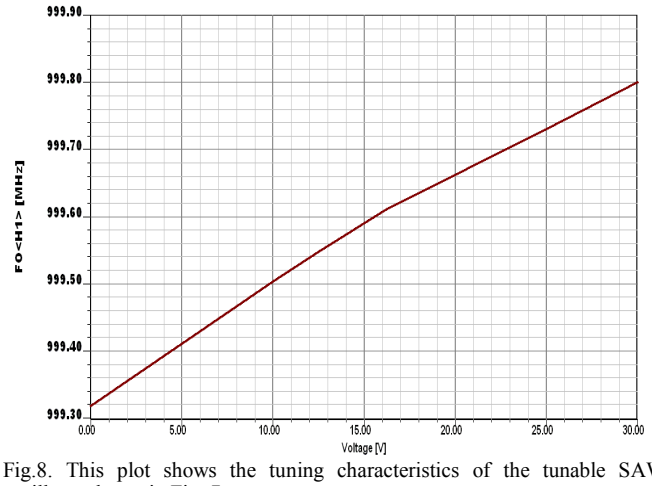


Fig.8. This plot shows the tuning characteristics of the tunable SAW oscillator shown in Fig. 7.

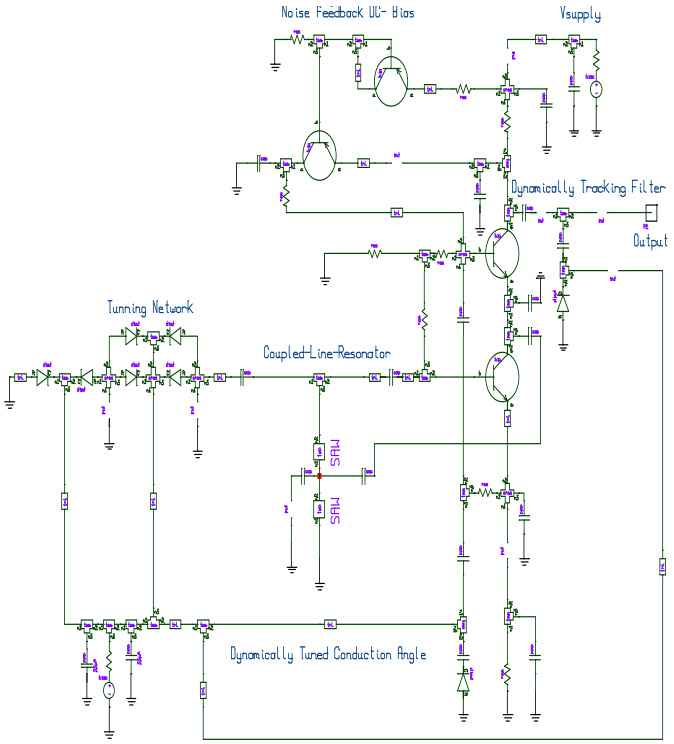


Fig.9. This schematic shows the cascaded mode-coupled harmonically tuned tunable SAW resonator oscillator

V. Example And Validation

The experimental results demonstrate the new design approach for ultra low phase noise VCSO for reference frequency standard applications.

Figure 11 shows the schematic of mode-coupled tunable 1 GHz SAW oscillator. The CAD simulated phase noise confirms the improvement in performance at 1 kHz and higher offset frequency from the carrier (Fig. 12).

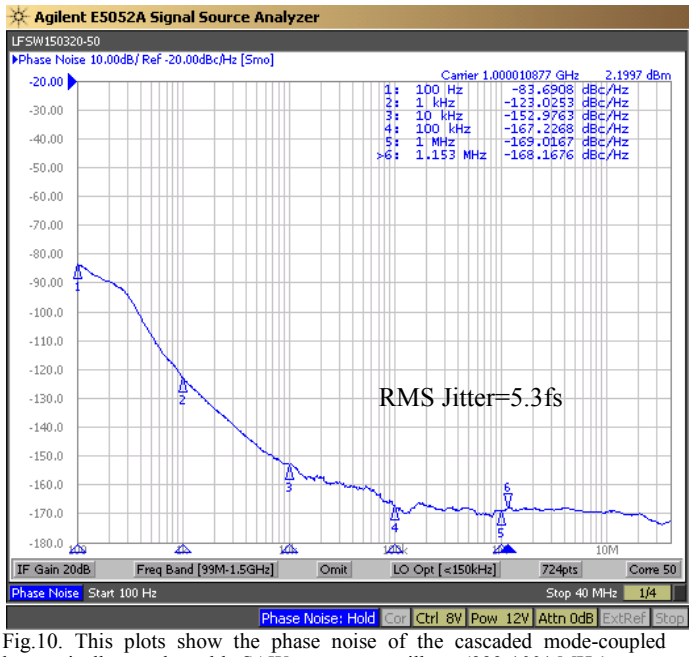


Fig.10. This plots show the phase noise of the cascaded mode-coupled harmonically tuned tunable SAW resonator oscillator (999-1001 MHz)

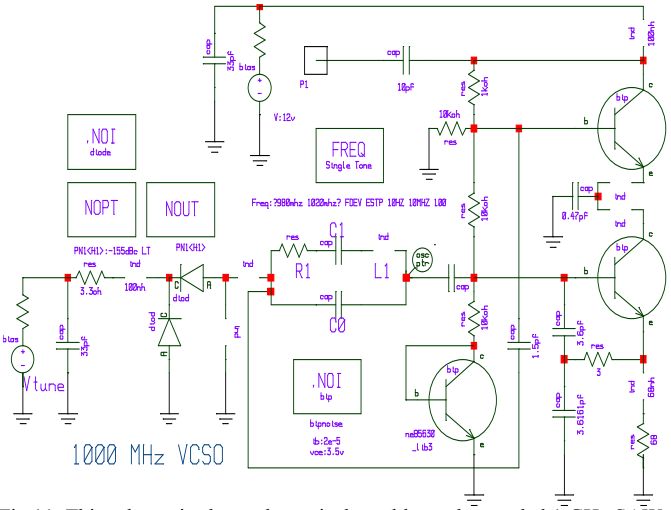


Fig.11. This schematic shows the typical tunable mode-coupled 1 GHz SAW resonator oscillator

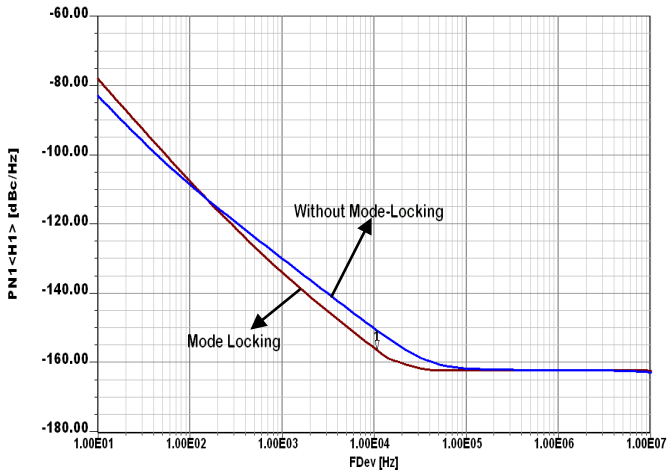


Fig.12. This plots show the phase noise of the mode-coupled tunable 1 GHz SAW resonator oscillator

Figure 13 shows the schematic of hybrid mode tunable 1 GHz SAW oscillator. For comparative analysis and brief insights about the various topologies, Figure 14 shows the measured phase noise plots of tunable SAW oscillators.

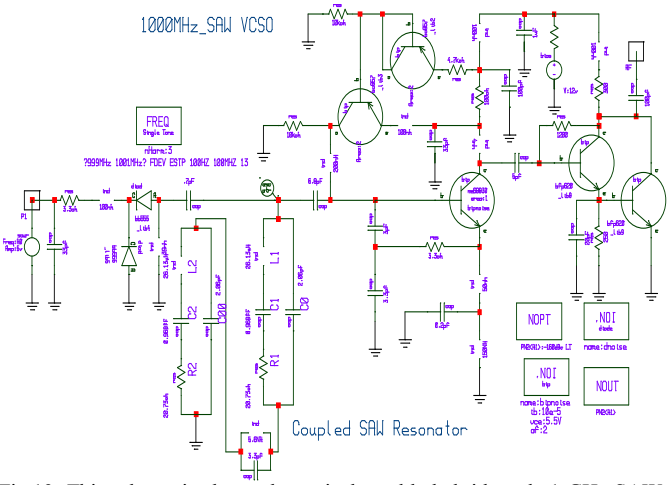
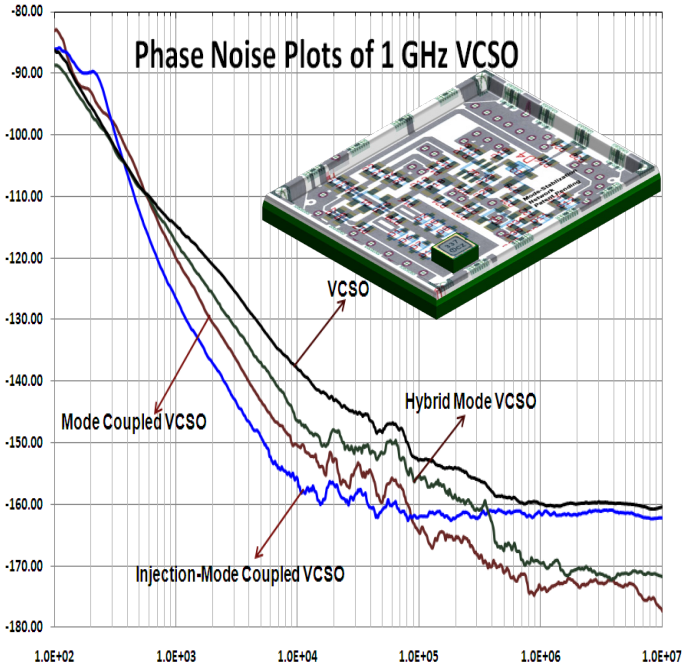


Fig.13. This schematic shows the typical tunable hybrid mode 1 GHz SAW resonator oscillator



It can be seen from Fig.14, injection-mode coupled VCSO offers promissing phase noise with 1 MHz tunung to compensate the ageing and frequency drift. The drawback of this approach is vibration sensitive VCSO circuit, care must be taken for the application where g-sensitive criteria is important yard stick. By proper packaging and securing the SAW resonator in the cavity using double layer PCB for compensating the temperatrre charateristics, including g-insensitive solutions.

VI. Conclusion

The new ultra low phase noise VCSOs offer significant promise in terms of performance, price, and product delivery to satisfy both the technical and business needs for reference frequency sources for synthesizer applications.

References

[1] C.S. Lam, A Review of the Recent Development of MEMS and Crystal Oscillators and Their Impacts on the Frequency Control Products Industry, 2008 IEEE International Ultrasonics Symposium Proceedings, pp. 694-704

[2] P. S. (Paul) Khanna, "Microwave Oscillators: The State of The Technology," *Microwave Journal*, April 2006, pp. 22-42.

[3] U. L. Rohde and A. K. Poddar, "User-Definable Thermal Drift Voltage Controlled Oscillator," United States Patent No.7, 265,642 B2, September 4, 2007.

[4] U. L. Rohde and A. K. Poddar, "Low Thermal Drift Tunable Frequency Voltage Controlled Oscillatory," United States Patent No.7, 262,670 B2, August 28, 2007.

[5] U. L. Rohde and A. K. Poddar, "Tunable Oscillator," US Patent No.7, 292,113, Nov. 6, 2007.

[6] U. L. Rohde and A. K. Poddar, "Tunable Frequency, Low Phase Noise and Low Thermal Drift Oscillator," United States Patent No.*7196591*, March 2007.

[7] U. L. Rohde and A. K. Poddar, "Multi-Octave Band Tunable Coupled-Resonator Oscillator," United States Patent No. 292,113, Nov. 6, 2007.

[8] U. L. Rohde and A. K. Poddar, "Low Noise, Hybrid Tuned Wideband Voltage Controlled Oscillator," United States Patent No. 7,365,612 B2, April 29, 2008.

[9] U. L. Rohde and A. K. Poddar, "Tunable Frequency, Low Phase Noise and Low Thermal Drift Oscillator," United States Patent NO. 7, 545, 229, June 09, 2009.

[10] U. L. Rohde and A. K. Poddar, "User-Definable, Low Cost, Low Phase Hit and Spectral Pure Tunable Oscillator," United States Patent No. 7,586,381 on September 08, 2009.

[11] U. L. Rohde and A. K. Poddar, "User-Definable, Low Cost, Low noise, and phase hit insensitive multi-octave-band tunable oscillator, Phase Hit and Spectral Pure Tunable Oscillator," United States Patent No. 7,605,670, October 20, 2009.

[12] U.L. Rohde, A.K. Poddar, and G. Boeck, *The Design of Modern Microwave Oscillators for Wireless Applications: Theory and Optimization*, Wiley, New York, 2005.

[13] A. K. Poddar, S. K. Koul, and Bharathi Bhat, "Millimeter Wave Evanescent Mode Gunn Diode Oscillator in Suspended Stripline Configuration," Infrared and Millimeter Waves, 22nd International Conference, July 1997, pp. 265-266.

[14] A. K. Poddar, J. K. Bansal, and K. N. Pandey, "Millimeter Wave Evanescent Mode Power Combiner Gunn Oscillator in Suspended Stripline Configuration," Mathematical Methods in Electromagnetic Theory, IEEE/MMET, International Conference, June 1998, pp. 384-386.

[15] U. L. Rohde and A. K. Poddar, "Novel Multi-Coupled Line Resonators Replace Traditional Ceramic Resonators in Oscillators/VCOs," IEEE, International Frequency Control Symposium, IFCS, Florida, June 5-7, 2006.

[16] A. K. Poddar and U. L. Rohde, "A Novel Voltage Controlled Crystal Oscillator (VCXO) Circuits," 9th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services, TELSIKS 2009, Serbia, October 7-9, 2009.

[17] U. L. Rohde and A. K. Poddar, "Electromagnetic Interference and Start-up Dynamics in High Frequency Crystal Oscillator Circuits," 2010 IEEE Sarnoff Symposium, Princeton, NJ, April 12-14, 2010.

[18] U. L. Rohde and A. K. Poddar, "Impact of Radiated EMI in High Frequency Crystal Oscillator," IEEE International Microwave Symposium 2010, Anaheim, CA, May 23-28, 2010.

[19] U. L. Rohde and A. K. Poddar, "Wideband voltage controlled oscillators employing evanescent mode coupled resonators," United States Patent *No*. 71803812, February 2007.

[20] U. L. Rohde, A. K. Poddar, and R. Rebel, "Integrated Low Noise Microwave Wideband Push- Push VCO," United States Patent No. 7,088189, August 2006.

[21] J. Everard and K. Theodoropoulos, "Ultra-Low Phase Noise Ceramic Based Dielectric Resonator Oscillators," *Proceedings of the IEEE International Freq. Control Symposium,* June 4-7, 2006, pp. 869-874.

[22] V. Walkar and I. C. Hunter, "Design of triple mode TE_{01} resonator transmission filters,"*IEEE MWC Let.*, Vol. 12, June 2002, pp. 215-217.

[23] S. Kim, W. Wu, J. Lin, A. Verma, S. Jang, F. Ren, S. Pearton, and J. Gillespie, "A High-Efficiency GaN/AlGaN HEMT Oscillator Operating at L-Band," *Proceedings of the 2006 Asia Pacific Microwave Conference*, Dec. 12-15, 2006, Yokohama, Japan.

[24] S. Romisch and R. Lutwak, "Low-Power, 4.6-GHz, Stable Oscillator for CSAC," *Proceedings of the IEEE International Frequency Control Symposium*, June 4-7, 2006, pp. 448-451.

[25] S. Hamano, K. Kawakami, and T. Takagi, "A Low Phase Noise 19 GHz-band VCO using Two Different Frequency Resonators," *2003 IEEE MTT-S Symposium Digest*, pp. 2189-2192.

[26] Grebennikov, *RF and Microwave Transistor Oscillator Design*, Wiley Ltd., London, England, 2007.

[27] J. K. A. Everard and C. Broomfield, "Reduced Transposed Flicker Noise in Microwave Oscillators Using GaAs Based Feedforward Amplifiers," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control,* Vol. 54, No. 6, June 2007, pp.1108-1117.

[28] P. A. Dallas and J. K. A. Everard, "Characterization of flicker noise in GaAs MESFETs for oscillator applications" *IEEE Tran. on MTT,* Vol. 48, No. 2, Feb. 2000, pp. 245-257.

[29] J. Browne, " Frequency Synthesizers Tune Communication Systems, " Microwaves & RF, March 2006.

[30] E. Rubiola and V. Giordano, " On the 1/*f* Frequency Noise in Ultra-Stable Quartz Oscillators", 2006 IEEE FCS.

[31] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum, " *IEEE*, *Proc.* pp. 329-332, 1966.

[32] S. Galliou, F. Sthal, and M. Mourey, " Enhanced Phase Noise Model For Quartz Crystal Oscillators", 2002 IEEE FCS.

[34] M. M. Driscoll, "Oscillator AM-to-FM Noise Conversion Due to the Dynamic Frequency- Drive Sensitivity of the Crystal Resonator", 2008 IEEE Frequency Control Symposium FCS, pp. 672-676.

[35] S. Galliou, F. Sthal, N. Gufflet, and M. Mourey," Predicting Phase Noise in Crystal Oscillators", IEEE FCS 2002.

[36] M. M. Driscoll, "Low Noise, VHF Crystal-Controlled Oscillator Utilizing Dual, SC-Cut Resonators", IEEE Tran. on UFFC, Vol. UFFC-33, No. 6, pp. 698-704, Nov. 1986.

[37] W-Thai Hsu," Resonator Miniaturization for Oscillators", 2008 IEEE Int'l Frequency Control Symposium, pp. 392-395.

[38] G. Ernst et. al., "Reducing Phase Noise Degradation Due to Mechanical Vibration on High Performance Quartz Resonator Osc. for Gateway Applications", 2008 IEEE FCS.

[39] G. Gonzalez, *Foundations of Oscillator Circuit*, Ch-IV, 2007 Artech House, Inc., London.

[40] Y. S. Shmaliy, "One-Port Noise Model of a Crystal Oscillator", *IEEE Trans.* on *U, F* and *FCS*, Vol. 51, No. 1, Jan. 2004.

[41 D. V. Bogomolov, R. Boroditsky, " Low Phase Noise UHF TCXO", Proc. Of the 2005 *IEEE IFCS*, pp. 509-511, 2005.

[42] X. Huang, W. W, Feng Tan, and W. Fu, " High-Frequency Overtone TCXO Based on Mixing of Dual Crystal Oscillators", *IEEE Trans.* on *U, F,* and *FC,* Vol. 54, No. 6, pp. 1103-1107, June 2007.

[43] K. Hennessy, "Quartz Crystal Basics: From Raw Materials to Oscillators", High Frequency Electronics, pp. 54-58, Dec. 2007.

[44] F. Sthal, S. Galliou, N. Gufflet, and M. Mourey, " Predicting Phase Noise in Crystal Oscillators", *IEEE Trans.* on *U, F,* and *FC,* Vol. 52, No. 1, pp. 27-30, June 2005.

[45] A. Grebennikov, RF and Microwave Transistor Oscillator Design, John Wiley & Sons Ltd., 2007, UK.

[46] Randall W. Rhea, *Oscillator Design And Computer Simulation,* 2nd ed., SciTech Publishing, Raleigh, NV, 2006.

[47] U. L. Rohde and A. K. Poddar, "Emerging Technology and Market Trends for Frequency Controlled Circuits and Timing Devices", Microwave Journal, pp. 20-44. May 2010.

[48] A. R. Northam, "Low Phase Noise, Wide Tuning Range Oscillator Utilizing a one port SAW Resonator and method of operation", US Patent No. 6, 239, 664 B1, May 2001..

ULISS: A Mobile Cryogenic Ultra-Stable Oscillator

S. Grop, B. Dubois, P-Y. Bourgeois, Y. Kersalé, E. Rubiola, G. Haye and V. Giordano.
FEMTO-ST Institute, Time & Frequency Dpt.
UMR 6174 CNRS - Université de Franche-Comté - ENSMM
26 Chemin de l'Épitaphe, 25000 Besançon, France
e-mail: giordano@femto-st.fr
web site: uliss-st.com

***Abstract*—We recently demonstrated a Cryogenic Sapphire Oscillator (CSO) presenting a short term frequency stability better than 3×10^{-15} for 1 s $\leq \tau \leq$ 1,000 s and achieving 4.5×10^{-15} for one day integration [1], [2]. This CSO was designed and built in the frame of a research contract funded by the European Space Agency (ESA). It incorporates a pulse-tube cooler instead of a bath cryostat – thus eliminating the need for regular supplies and manual transferring of liquid helium. The advent of reliable and cryocooled CSO opens the possibility to implement such an ultra-stable reference not only in metrological laboratories with liquid helium facilities but also in remote sites like base stations for space navigation, VBLI antenna sites, ...**

In a new project: ULISS (Ultra-Low Instability Signal Source), funded by Regional and European Institutions, we are developping a new cryocooled oscillator specially designed to be transportable. The ULISS oscillator will be compared and tested in various laboratories, observatories and industrial companies throughout Europe.

In this paper we present the status of the ULISS project and the last frequency stability and phase noise characterisations of the current CSO working for more than one year in our lab.

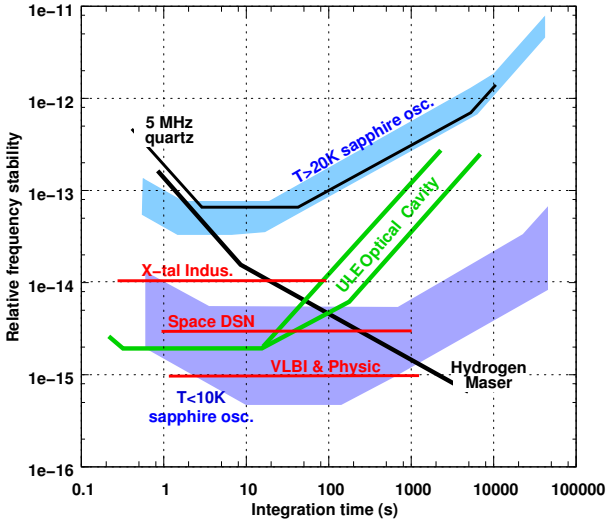


Figure 1. Relative frequency stability of some clocks and oscillators.

I. INTRODUCTION

The ever-increasing need for better time synchronization in number of High-Tech and scientific applications calls for the development of new frequency references of improved stability. The figure 1 shows the relative frequency stability as a function of the integration time τ for various clocks and reference oscillators as well as the needs newly expressed by various users:

The blue area in the figure 1 represents the typical performance of microwave oscillators incorporating a dielectric sapphire resonator cooled in a liquid helium bath. The Cryogenic Sapphire Oscillator (CSO) can exhibit the highest short-term frequency stability ($< 1\times10^{-15}$ for 10 s$< \tau <$ 100 s). CSO is the only available technology able to meet the requirements expressed by various users (red lines in Fig.1). Few liquid helium cryogenic sapphire oscillators have been built in a small number of worldwide research laboratories [3], [4], [5], [6]. Some of these instruments have been used as local oscillator for atomic fountain clocks [7] and for fundamental physical experiments [8]. Nevertheless they are still experimental prototypes that are difficult to implement outside a well-equipped laboratory. Moreover the liquid helium cryostat has to be refilled periodically adding a heavy maintenance constraint. Yet, various applications could benefit of the performances offered by the cryogenic sapphire oscillator:

- Space: The interplanetary navigation is based on microwave signal exchanges between the space vehicle and a ground station equipped with a high stability timing reference. Ranging and Dopplertracking resolutions are limited by the clock instabilities. To be able to met the stability requirements for the future space missions, the European Space Agency (ESA) sponsored the construction of a cryocooled–CSO. This project (ELISA project [9]) coordinated by the Femto-ST Institute was finalised in April 2010. The cryocooled–CSO achieves better than 3×10^{-15} frequency stability over 1 s to 1,000 s while insuring at least one year autonomy.

- Metrology: primary frequency standards. In an atomic clock, a local oscillator (LO) is frequency locked on the resonance frequency of an atomic transition. The best current atomic clocks using cold atoms operate sequentially. Indeed time is needed to cool the atoms before sensing the atomic transition. During this period of time no information is available for the frequency of the LO which is then free-running. Consequently the overall clock frequency stability is degraded by the short-term instability of the LO. Better than 1×10^{-14} over 1–100 s is required.

- Metrology: frequency stability certification. National Metrological Institutes have in charge to maintain the

standards and to establish certifications for industrial products and laboratory instruments. The frequency stability is one of the major characteristics of the signal source that could impact on the security and performances of number of systems. It is then a metrological quantity that has to be measured with the best resolution available. A standard presenting frequency stability better than 1×10^{-14} over 1–1,000 s will notably increase the current calibration capability making it compatible with all current industrial requirements.

- Very Long Baseline Interferometry: VLBI allows data from widely separated radio telescopes to be combined to get exceptional astronomical measurement resolution providing each radio telescope is equipped with an ultra-stable clock.

- Industry of high performance clocks and oscillators: To estimate, without ambiguity the frequency performances of state-of-the-art oscillators and clocks requires an ultra-stable reference. Some industrial companies making high performance quartz oscillators are interested by a versatile frequency reference of less than 1×10^{-14} frequency instability.

- Fundamental Physics. Special Relativity is one basic pillar of modern physics. Each basic theory has to be tested as its best, because it is so fundamentally important. The best tests of Special Relativity are based on ultra-stable oscillators for which we are locking for a frequency variation as a function of the speed and orientation of the laboratory. The frequency fluctuations of the oscillator limit the measurement resolution.

For all these applications, the autonomy and reliability of the frequency reference are major requirements. In all cases, liquid helium bath has to be rejected.

II. The ULISS project

Our work on the field of cryogenic oscillators was initiated 15 years ago, in the former Laboratoire de Physique et de Métrologie des Oscillateurs (LPMO) which becomes part of the Time and Frequency Department of the Femto-ST Institute at its creation in 2004. The project ELISA funded by ESA represents the outcome of all these efforts: the objective set by ESA was plenty fulfilled and we demonstrated for the first time the possibility to use a cryocooler while maintaining a state-of-the-art frequency stability.

We think that today the cryocooled–CSO technology is mature enough to be proposed as a reliable frequency reference for various applications. The ULISS (Ultra-Low Instability Signal Source) project answers our wish to see the CSO–technology used by other. This project is funded by the Fonds Européen de Développement Régional (FEDER), the Région de Franche-Comté and OSEO. It has been offically launched the 1st of March 2011 and contains two main parts:

- We are building a new CSO specially designed to be transportable by road. The new oscillator will be transferred (by road) successively to several potential users' sites. It will be tested in the real experimental conditions and critical result analysis will be carried out with the user. Hence the potentiality to improve their experiments as well as the reliability and autonomy of the CSO will be clearly demonstrated. These international tests will also give the opportunity to establish precisely the requirements specific to each potential user.

- A business unit: ULISS® has been created and is currently managed by the Université de Franche-Comté with the objective to commercially propose the CSO technology. The first purpose of this entity is the implementation of the first prototype, i.e. ELISA, in the ESA ground station situated in Malargüe (Argentina).

Visit our web site: www.uliss-st.com.

III. CSO performances

The CSO architecture is depicted in the figure 2. The heart of the instrument is a Whispering Gallery Mode Sapphire resonator stabilized at its thermal inversion point, i.e 6.1K. Such a low temperature is achieved in a closed–cycle Pulse–Tube Cryocooler specillay designed to ensure a high thermal stability ($\pm$1mK) and a low level of mechanical vibration: axial displacement less than 2 μm at the PT cycle frequency ($\approx 1Hz$). The CSO is designed to oscillate at 9.99 GHz $\pm$5 MHz. The difference with the 10 GHz round frequency is compensated by a low noise DDS. VHF and RF signals are synthetized from the stabilized 10 GHz [10]. The system is complemented with a Phase–COmparator (PCO) enabling to lock the synthized outputs to an incoming 100 MHz signal.

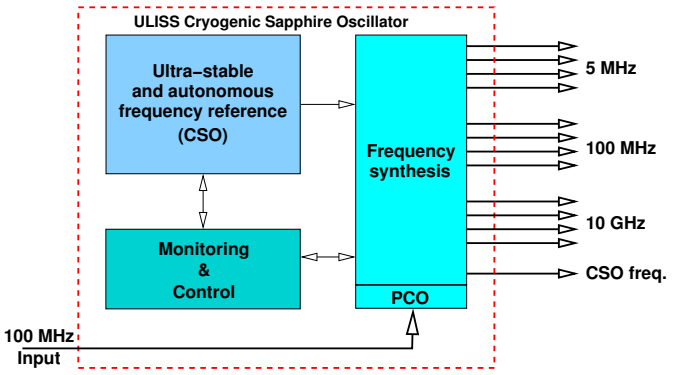


Figure 2. ULISS CSO architecture

The typical frequency stability performance of the CSO is shown in the figure 3. It is compared with the ADEV of other frequency references as active and passive Hydrogen Maser. The CSO bridges the gap between (unstable) high spectral purity oscillators and H maser. Here, we show fully conservative results. In all our measurements we do not remove drift, drift rate etc and we show the total noise of two

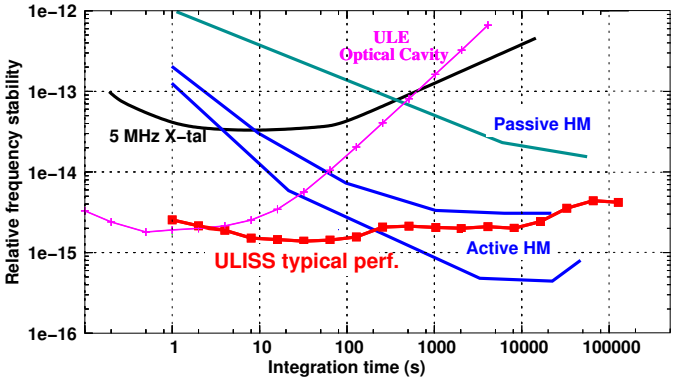


Figure 3. ULISS typical stability compared to some other clocks and oscillators.

oscillators, instead of taking away 3 dB under the assumption of two equal units.

IV. An example of application

As a Metrology laboratory associated with the Laboratoire National d'Essai et de Métrologie (LNE), a first evident application for us is the characterization of ultra-stable oscillators. Short term frequency stability of Quartz X-tal oscillator progresses only very slowly for more than 10 years and seems limited to $0.5 - 1 \times 10^{-13}$. In the X-tal oscillator, the location of the main source of noise is still not known and could be the quartz resonator frequency fluctuations and/or the sustaining amplifier phase noise. Even if some effort have been deployed to improve the resolution of measurement set-ups [11], [12], the limiting cause has still not be revealed. The ambiguity of the results remains due to the fact that the current instruments do not allow for the characteristaion of individual oscillator or resonator.

The CSO overcomes this issue, and offers an unprecedented resolution in the short term frequency stability measurement. Owing to the Uliss high stability, the frequency-stability measurement of most oscillators becomes straightforward as demonstrated by the experiment schematized in figure 4:

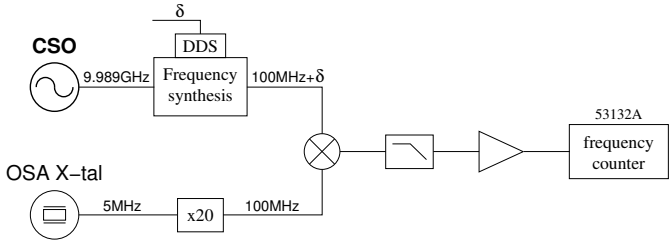


Figure 4. Simple measurement set-up used to evaluate the ADEV of a state-of-the-art X-tal oscillator.

An ultra-stable quartz X-tal oscillator from Oscilloquartz (modele 8607-B) [13] is simply compared to the CSO. The measurement is made at 100 MHz to ensure a sufficient resolution at short integration times. The beatnote frequency is adjusted by acting on the DDS command to optimize the counter operation. The raw result of this comparaison is shown in figure 5. 1s gate time samples have been recorded during 4,800 s, and ADEV computed without any post-treatment.

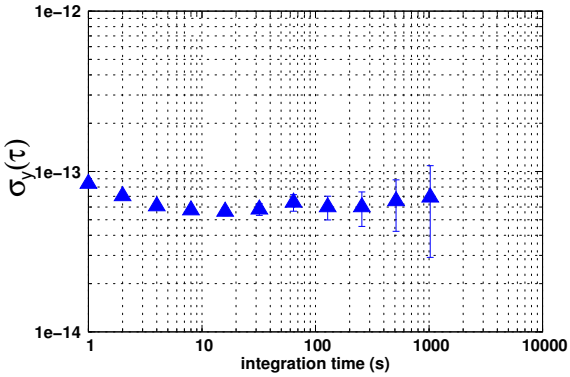


Figure 5. Relative frequency stability of a new Oscilloquartz 5 MHz X-al oscillator as measured with the CSO.

As revealed by the CSO, the short-term frequency stability of this X-tal oscillator constitutes a world record. The measured ADEV is limited by a flicker floor of 5×10^{-14} and still remains better than 8×10^{-14} at $\tau = 1,000$ s.

V. Acknowledgements

The authors want to thank M. Jacques Chauvin from Oscilloquartz to provide ultra-stable X-tal oscillator and to participate to the measurement compaign.

The ULISS team wish to express acknowledgements to the "Fond Européen de développement Régional " (FEDER), "le Conseil Général de Franche–Comté" and OSEO for their support.

References

[1] S. Grop, P.-Y. Bourgeois, N. Bazin, Y. Kersalé, E. Rubiola, C. Langham, M. Oxborrow, D. Clapton, S. Walker, J. D. Vicente, and V. Giordano, "A crycooled 10 GHz oscillator with 10^{-15} frequency stability," *Review of Scientific Instruments*, vol. 81, pp. 025102–1–7, 2010.

[2] S. Grop, P. Y. Bourgeois, R. Boudot, Y. Kersalé, E. Rubiola, and V. Giordano, "10 GHz cryocooled sapphire oscillator with extremely low phase noise," *Electronics Letters*, vol. 46, pp. 420–422, 18th March 2010.

[3] A. Luiten, A. G. Mann, A. Giles, and D. Blair, "Ultra-stable sapphire resonator-oscillator," *IEEE Transactions on Instrumentation and Measurement*, vol. 42, pp. 439–443, Apr. 1993.

[4] G. Marra, D. Henderson, and M. Oxborrow, "Frequency stability and phase noise of a pair of x-band cryogenic sapphire oscillators," *Meas. Sci. Technol.*, vol. 18, pp. 1224–1228, 2007.

[5] K. Watabe, J. Hartnett, C. R. Locke, G. Santarelli, S. Yanagimachi, T. Ikegami, and S. Ohshima, "Progress in the development of cryogenic sapphire resonator oscillator at NMIJ/AIST," in *Proc. 20th European Frequency and Time Forum*, (Braunschweig, Germany), pp. 92–95, March. 27–30 2006.

[6] G. Dick, D. Santiago, and R. Wang, "Temperature compensated sapphire resonator for ultra–stable oscillator capability at temperatures above 77K," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 42, no. 5, pp. 812–819, 1995.

[7] G. Santarelli, P. Laurent, P. Lemonde, A. Clairon, A. G. Mann, S. Chang, A. N. Luiten, and C. Salomon, "Quantum projection noise in an atomic fountain: A high stability cesium frequency standard," *Physical Review Letters*, vol. 82, June 1999.

[8] P. Wolf, S. Bize, A. Clairon, A. N. Luiten, G. Santarelli, and M. E. Tobar, "Test of Lorentz invariance using a microwave resonator," *Physical Review Letters*, vol. 90, pp. 060402–1–4, Feb. 14 2003.

[9] S. Grop, P. Bourgeois, N. Bazin, E. Rubiola, C. Langham, M. Oxborrow, W. Schafer, J. De Vicente, Y. Kersalé, and V. Giordano, "High Stability Cryocooled 10 GHz Oscillator for the European Space Agency," in *Proceedings of the 2010 IEEE International Frequency Control Symposium*, (Newport beach CA, US), 2010.

[10] S. Grop, P.-Y. Bourgeois, E. Rubiola, W. Schäfer, J. De Vicente, Y. Kersalé, and V. Giordano, "Frequency synthesis chain for the ESA deep space network," *Electronics Letters*, vol. 47, pp. 386–388, Mar. 17 2011.

[11] E. Rubiola, J. Groslambert, M. Brunet, and V. Giordano, "Flicker noise measurement of HF quartz resonators," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 47, pp. 361–368, Mar. 2000.

[12] A. Kuna, J. Cermak, L. Sojdr, P. Salzenstein, and F. Lefebvre, "Lowest flicker-frequency floor measured on BVA oscillators," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 57, no. 3, pp. 548–551, 2010.

[13] `http://www.oscilloquartz.com/`.

A Multi-Channel time interval measurement for Time Comparison

SHI Shao-hua [1,2,3], LI Xiao-hui [1,3], ZHANG Hui-jun [1,2,3], ZHAO Zhi-xiong [1,3]

[1] National Time Service Center, Chinese Academy of Science, Lintong, Shaanxi, 710600, China

[2] The Graduate School of the Chinese Academy of Science, Beijing, 100049, China

[3] Key Lab. of Time & Frequency Primary Standards, CAS, Lintong, Shaanxi, 710600, China

***Abstract*—This paper presents improvements on a multi-channel time interval measurement for Time Comparison. The measurement has eight channels and measurements the time interval between nine clocks at the same tine. The measurement design is based on a commercially available GPX ASIC from acam-messelectronic gmbh provides a maximum of 86ps resolution per channel. In this measurement precision the measurement range of system is less than 10us and goes beyond the requirement of the time comparison. We extend the measurement range to endless by the internal start re-trigger and external interrupt. The architecture details of this system as shown in figure 1. We also develop a software runs on pc can control the measurement and display the measurement result in real time. We analyze the resolution variation of the measurements and the present linearity plots finally. The calculate results indicate that the precision per channel just a little decline after the measurement range was extended.**

I. INTRODUCTION

Time comparison is crucial in many areas of research and application such as timekeeping, navigation and positioning, laser distance measurement and so on. Multi-channel, high-resolution and low-power time interval measurement is demanded by time comparison for measuring time intervals between several received time signals to the standard time reference signal.

We developed a multi-channel time interval measurement based on time digital converter for comparing time between the reference clock and other clocks in a certain project. Time digital converter (TDC) is one of the most important tools to measure time intervals precisely [1] between several time signals. We used a commercially available time to digital converter to provide a two-level resolution for expediting progress of the measurement. Measurement system is communicated with personal computer (PC) by serial interface, this configuration provides a simply and efficient way of using the software running in the PC controlled the measurement, displaying the experimental results and storing these data in the hard disk of PC [2]. Such design is easy and efficient and convenient to modify. The experimental results are presented and the performance of this measurement is discussed in the end.

II. DESIGN

In order to measure the intervals between the reference clock and other clocks in the time standard assembly, we had developed an eight channel time interval measurement based on the TDC-GPX time to digital converter ASIC. The measurement is comes in 2U form factor for an optimized proprietary backplane bus. The selection of TDC-GPX for the time comparison of our project was motivated by the existing experience with it.

The principle of high resolution of the TDC-GPX is based on digital tapped delay line within the phase-locked loop. The line is composed of a number of delay cells, i.e. logic gate unit, each having the same propagation delay. It can be picoseconds resolution by modern CMOS technology. The use of the PLL ring oscillator can be regarded as a multiphase clock driving the clock inputs of the edge-triggered D-type flip-flops of the associated register, as shown in figure 1. Then the delay time can be precisely determined by counting the number of logic gate that start signal pass through.

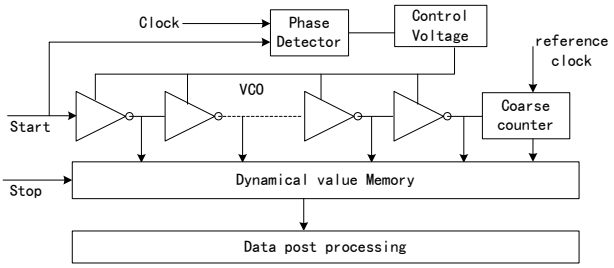


Figure 1. the principle of high resolution of the TDC.

This chip provides 8 channels with a typical resolution of 81ps least significant bit and can be slightly adjusted. The double-pulse resolution is 5.5 ns with 32-fold multi-hit capability, resulting in a peak hit rate of 182 MHz per channel. Over all channels a continuous rate of 40 MHz can be achieved.

It measures the time difference between each of the eight stop inputs and one common start input uses digital delay line techniques, and the total measuring range is 17 bit. In order to keep delays in the internal delay locked loop constant of this chip, it also had a phase locked loop circuit for the core voltage to compensating temperature drifts and manufacturing tolerances. The interval measurement results are stored in internal FIFOs of the chip, which are read out by a 28-bit parallel port clocked with 40 MHz.

Fig 2 depicts the proposed major part of the measurement according to the block diagram. The hardware architecture of the system consists of a power supply circuit, a pretreatment circuit, a time to digital converter circuit, a main-controller circuit, a reference frequency standard, power and resolution adjustment circuit and some other circuits.

Supported by the National Natural Science Foundation of China

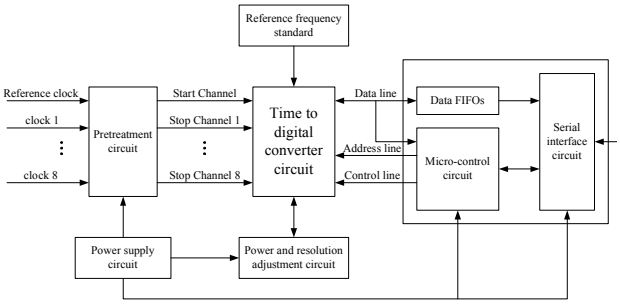


Figure 2. the hardware architecture

The pretreatment circuit has nine input interface and nine output interface. The time signal of clock was converted to TTL signal or LVTTL signal. So that it could be measured by the time to digital converter. The power supply circuit provides 3.3V direct current to pretreatment circuit, power and resolution adjustment circuit, micro-control circuit, serial interface circuit and so on.

The time to digital converter circuit is the core of the hardware of measurement which includes a high-resolution measurement circuit based on TDC-GPX. The principle of high resolution measurement part is based on digital tapped delay line within the phase-locked loop [3]. The TDC circuit consists of a start input channel and eight stop input channel which can accept TTL or LVTTL signals by the SMA connector.

The main-controller circuit consists of data FIFOs, micro-control circuit and serial interface circuit. This circuit connects to the TDC circuit by data bus include address line, data line, control line and some flag bit output lines and so on. The micro-control circuit is based on a 32-bit micro-controller which working frequency is up to 40 MHz [4]. The operation to the TDC circuit is directly accomplished by this micro-controller such as power-up reset, configuration of the registers, initiating measurement of input signal and readout the measurement result. The data FIFOs was a 28 ×100 Bit register. The results of time intervals were pre-stored individually in the register and are read out all together by the personal computer. The interface with the computer becomes very simple using the serial interface circuit, control data and measurement results transmit between the main-control circuit and personal computer by the serial RS232 interface.

The reference standard circuit provides a 40 MHz sine wave frequency signal as reference clock signal of the TDC circuit which is converting four times by the 10MHz high resolution frequency signal. The jitter of the reference was controlled less than 80ps. Because the jitter of the reference clock signal has a direct affection on precise measurement when the internal start retrigger signal of TDC circuit is used.

In principle the high resolution of the TDC circuit is derived from the internal gate propagation times. The logic gate propagation time depends upon voltage, temperature and the manufacturing process, the resolution of the TDC is not known and must be calculated via calibration measurements [5]. In addition, the resolution is not stable with the swing of voltage and temperature. So a power and resolution adjustment circuit is indispensable to adjust quartz-accurately and absolutely temperature stable and regulates the core voltage of the TDC, so that the resolution of TDC is set exactly to the programmed value [6].

Furthermore the implemented software consists of three main components, the device driver, the firmware of micro-controller and an application program runs on PC. The block diagram of software as shown in fig. 3.

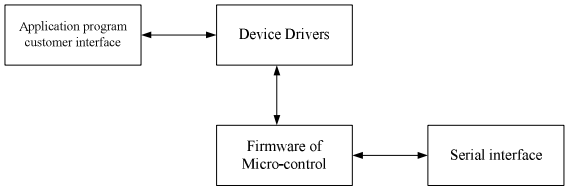


Figure 3. The software architecture block diagram.

The device driver is basic segment of the software which communicates with the measurement through the computer bus to the measurement hardware. When an application software program invokes a routine in the driver, the driver should issues commands to the device. Once the device sends data back to the driver, the driver may invoke routines in the original calling software. Driver provides the interrupt handling required for the serial interface and is hardware-dependent and operating-system-specific.

The firmware is involved with very basic low-level operations without which the measurement would be completely non-functional. It resides in the ROM or RAM of the micro-controller, and directly operates the time to digital circuit and so on. The firmware implements the power-up reset, the registers configuration, enabling the measurement and readout the measurement data from the time to digital circuit.

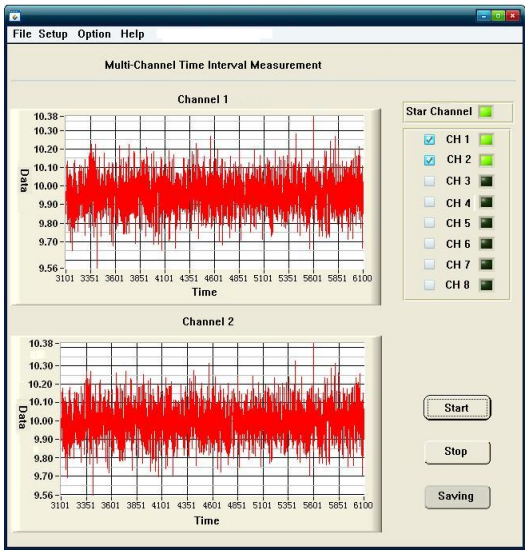


Figure 4. The customer interface of application program

The customer interface of application program as showed in fig. 4. It automatically detects when the measurement device is being prepared, initializes the serial communication and starts the data acquisition. The software displays the result of the measurement which read from the firmware and plots the

graphics of these data, save all the measurement result in the hard disk of the personal computer.

III. EXPERIMENTAL RESULT

Measurements have been carried out with the help of master clock of national service center in order to check if the measurement conforms to the requirements with regard to resolution and measurement range for time comparison.

For the resolution and range measurement, the master clock of NTSC produces 10Mhz sine wave frequency was reduplicate four times by the measurement circuit as the reference clock. A 1pps signal synchronized with the master clock as the start signal. And the 1pps output from master clock is divided into the two same signals and added some time delay. The two divided 1pps are connected to two stop channel respectively. The TDC is configured as the I-mode and the resolution is actually 130ps and the measurement range is from 200 ns to 30s. It accepts only one stop event per channel, therefore the next measurement is made after Master-Reset.

A. Short measure range

The short measure rang result as shown in figure 5. The delay between the start pulse and stop pulse is 1251275835 ns (include the delay of cable and channel) and the measurement time is 1200 second. The mean of result is 1251275836.479 ns; the maximum positive offset from mean is 0.39 ns and the maximum negative offset from mean is 0.42 ns. The standard deviation of the result is 134 ps.

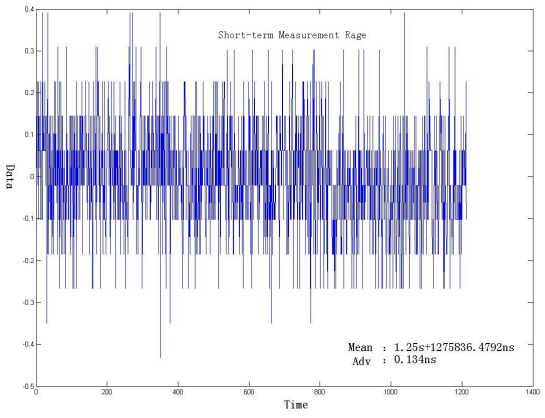


Figure 5. The short-term measure rang result

B. Long measure range

For the long measure rang test, the delay between the start pulse and stop pulse is 29S + 51 ns (also include the delay of cable and channel); the measurement time is 14000 second. All the results are shown in figure 6. The statistical results are showed that all of the readings are between the maximum 29S + 52.8ns and the minimum 29S + 50.1ns. The mean of result is 29S + 51.3237 ns and the standard deviation of the result is 316ps.

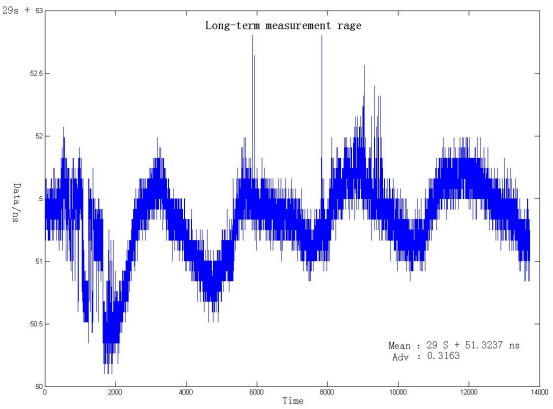


Figure 6. The long-term measure rang result.

IV. CONCLUSION

The measurement has eight input channels plus a common channel and realizes common-start and common-stop operations. It can provide a maximum of 134ps resolution per channel with the measurement range approximately 1.25 second and a maximum of 316ps resolution per channel with the measurement range approximately 30 second. The experimental result show that the resolution per channel was a little decline along with measurement range was extended from 1s to 30s. We will improve the precision of this multi-channel time interval measurement with a large measurement range and extend the measurement rage to endless in the future work.

REFERENCES

[1] A.S. Yousif, J.W. Haslett, A fine resolution TDC Architecture for next generation PET imaging[J]. nuclear science, 2004, 54(5): 1574-1575.

[2] Gao Wu, Gao Deyan, Wei Tingcun, Hu-guo, C.YannHu. A high-resolution multi-channel time-to-digital converter(TDC) for high-energy physics and biomedical imaging applications[C]. IEEE conference on industrial electronics and applications, 2009: 1133-1139.

[3] A. Boujard, D. Bloyet, M. Tripon, A digital TDC with a reduced number of delay line cells[J], Nuclear Instruments and Methods in Physics Research A, 486, pp. 803-812.

[4] Szplet R, Kalisz J, Szymanowski R. Interpolating time counter with 100ps resolution on a single FPGA device [J]. IEEE Transaction on Instrumentation and Measurement, 2000, 49 (4): 879-883.

[5] Andrew N. Novick, Michael A. Lombardi, Victor S. Zhang, A High Performance Multi-Channel Time-Interval Counter With An Integrated GPS Receiver, in Proc. 31st Precise Time and Time Interval Meeting , 1999: 561-567.

[6] Jinyuan Wu, An FPGA wave union TDC for time-of-flight application [C]. Nuclear science symposium conference record, 2009:299-303.

Vibration Sensitivity of Optical Components: A Survey

A. Hati, C. W. Nelson, D. A. Howe, *Senior Member, IEEE*
National Institute of Standards and Technology,
Boulder, CO 80305 USA
e-mail: archita@boulder.nist.gov

***Abstract*—Building optical fiber-based systems presents different challenges than free-space architectures due to the inherent vibration sensitivity of the fiber and the associated components. A survey of the vibration sensitivity of an assortment of commonly used fiber-based optical components is presented to identify problematic parts of a fiber-based design. The measurement of vibration sensitivity is challenging due to the difficulty of separating the sensitivity of the components from the measurement apparatus itself. The noise introduced by the interconnecting fibers bridging between the stationary measurement system and the vibrating device under test can dominate and mask the noise of the device being measured. We propose and demonstrate a novel technique to measure the vibration sensitivity of fiber-based optical components. It uses a common-arm counter-propagating frequency-shifted interferometer that cancels the vibration-induced phase noise of the interconnecting fibers. The proposed technique improves the vibration-induced phase noise floor by more than 30 dB compared to a conventional frequency-shifted Mach-Zehnder interferometer and allows measurement of low vibration sensitive devices.**

***Index Terms*— optical fiber, interferometer, phase noise, vibration sensitivity**

I. Introduction

There is a growing need for low phase modulated (PM) noise and low vibration sensitive oscillators for many applications such as radar, navigation, spectroscopy, and timing. In recent years, ultra-low PM noise microwave signals have been generated by dividing optical signals from stabilized lasers. This division from the optical to microwave domain has resulted in extremely low PM noise of -101 dBrad^2/Hz or lower at 1 Hz offset from a 10 GHz carrier [1]. Opto-electronic oscillators (OEO) have also been used to generate low noise microwave signals by use of a modulated optical carrier [2]. Presently, the lowest PM noise has been achieved only in quiet, low-vibration laboratory environments. The vibration sensitivity of an OEO is on the order of 10^{-8} to 10^{-10} per g (1 g is the acceleration of gravity near the earth's surface, approximately 9.8 m/s^2), and arises mostly from the length fluctuation of the OEO's long optical fiber that acts as a resonator [3-5]. Extensive work involving vibration insensitive Fabry–Pérot cavities have been reported [6-8]. However, there is little information on the vibration sensitivity of fiber-based optical components [9-11]. The focus of this paper is to study the noise performances of an assortment of such optical components under vibration.

A conventional frequency shifted Mach-Zehnder interferometer (MZI) for measuring vibration sensitivity of fiber-based optical devices is shown in Fig. 1. During the course of setting up a measurement of vibration sensitivity, the interconnecting fibers (F_1 and F_2 in Fig. 1) that deliver the optical signal to and from the device under test (DUT) were so sensitive to vibration that the measurement of a given device could not be made with assurance. Since these delivery fibers cross between the stationary reference frame and the moving frame of the vibration actuator, they undergo not only vibration due to the actuator, but also bending and stretching between the two frames. Mechanical distortion of the core and surrounding cladding causes fluctuations in the phase of the optical signal passing though the fiber [3]. Many fiber-based optical components are connected via integrally included optical fiber 'pigtails' (F_A and F_B in Fig. 1). The sensitivity of these pigtail fibers cannot be ignored and must be viewed as an integral part of the DUT. The effect of these pigtail fibers are not to be confused with, and should be separated from, the delivery fibers that connect the DUT on the vibration actuator to the measurement interferometer. Unfortunately, in many cases the delivery fibers that bridge the stationary and vibratory frames undergo larger phase fluctuations than those experienced by the DUT being evaluated. In this letter, we propose and experimentally demonstrate a novel scheme for reducing the vibration effect on the interconnecting delivery fibers while measuring the vibration sensitivity of an assortment of optical fiber-based components. Sections II and III constitute the technical portion of this paper. Section II describes the setup for reducing the vibration effect on the interconnecting fibers while measuring a component under vibration. Section III discusses measurements results of the optical component's vibration sensitivity.

II. Measurement Methods

In order to measure the vibration sensitivity of a component accurately, it is important to know the noise floor of the measurement system. The initial measurement technique employed to measure the PM noise, and hence the sensitivity of optical components to vibration, is shown in Fig. 1. It consists of a frequency-shifted MZI with an erbium-doped

fiber laser at 1550 nm as the optical source. A 40 MHz acousto-optic modulator (AOM) is connected to the reference path of the interferometer, shifting the frequency of the laser by 40 MHz and translating the output of the MZI from baseband to radio frequency for analysis. The DUT, mounted on a vibration platform, is inserted in the non-shifted measurement path of the MZI. The lengths of the two paths are kept almost equal, to minimize differential delay. Finally, the 40 MHz beat at the output of the photo-detector (PD) is amplified and measured on a direct-digital PM noise measurement system. All fibers used in the system are single-mode SMF-28.

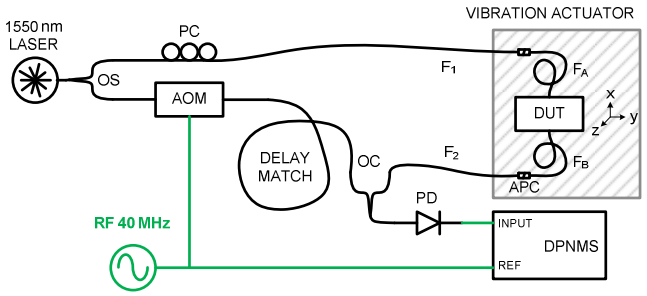


Fig. 1. Block diagram of a conventional frequency-shifted MZI for residual vibration sensitivity measurement of optical devices. PC: Polarization Controller, OS: Optical Splitter, OC: Optical Combiner, AOM: Acousto-Optic Modulators, PD: Photo detector, APC: Angled Physical Contact Connector. DUT: Device Under Test, DPNMS: Digital Phase Noise Measurement System, F_1, F_2: Interconnection Fibers, F_A, F_B: DUT Fiber Pigtails. Green lines represent radio frequency signals and black lines represent optical signals. The z-axis is perpendicular to the page.

Traditionally, vibration sensitivity is given by Γ_y, which is defined as the ratio of fractional frequency fluctuations to acceleration. For devices such as resonators and delay lines, fluctuations in the length (l) or delay (τ) are often normalized as $\delta l/l$ or $\delta\tau/\tau$. The vibration-induced phase fluctuations observed for short fibers are not necessarily proportional to their lengths. Therefore, an alternate vibration-phase sensitivity, Γ_φ, for two-port devices can be defined as [12]

$$\Gamma_\varphi = \sqrt{\frac{S_\varphi(f)}{S_g(f)}} \quad \text{rad/g}, \qquad (1)$$

where, $S_\varphi(f)$ is the double-sideband phase noise in rad^2/Hz and $S_g(f)$ is the root mean square (rms) acceleration power spectral density (PSD). Vibrational effects that are distributed with length are best described using Γ_y. Localized or spot effects due to vibration should not be normalized by total length and can be described with Γ_φ. Conversion to the traditional Γ_y can easily be made with use of the carrier frequency.

To determine the noise floor of the conventional MZI measurement system under vibration, the DUT is removed, and fibers F_1 and F_2 are connected and secured to the vibration platform. For this test, a random white vibration profile of $S_g(f)$ = 1 mg^2/Hz is used for 10 Hz $\le f \le$ 1000 Hz. The vibration phase sensitivity of the conventional MZI is shown by the top green curve in Fig 3. The noise floor due to the bending and stretching (length fluctuations) of interconnecting fibers was excessive and thus prohibited the measurement of low-vibration sensitive devices. A different approach was therefore necessary that compensated for fibers F_1 and F_2 while retaining sensitivity of the DUT connected between them. A novel technique using a common-arm counter-propagating frequency-shifted Mach-Zehnder interferometer (CACP-MZI) is proposed and shown in Figure 2. In this new method the conventional MZI is modified with the addition of four circulators. Two of these circulators, A and D, are mounted on the stationary measurement system, and the remaining two circulators, B and C, are mounted on the vibrating platform. The optical signal is split into two paths by use of a 50/50 coupler. The forward measurement signal path, represented by the blue arrows, propagates through the circulators A, B, the DUT, C and, D as shown in Fig 2. Similarly, the reverse reference path represented by the red arrows is frequency-shifted by the AOM and propagates through circulators D, C, B and A respectively, while bypassing the DUT. The forward and reverse signals combine at the PD, creating a 40 MHz beat signal that is then analyzed. The noise introduced from the flexing interconnecting fibers, F_1 and F_2, is common to both the forward measurement and reverse reference paths and thus cancels in the MZI, producing a lower noise floor. It should be noted that in CACP-MZI, there is a short uncompensated signal path between ports 2 - 3 of circulator B, ports 1 - 2 of circulator C and fiber F_3 that contributes to the noise floor.

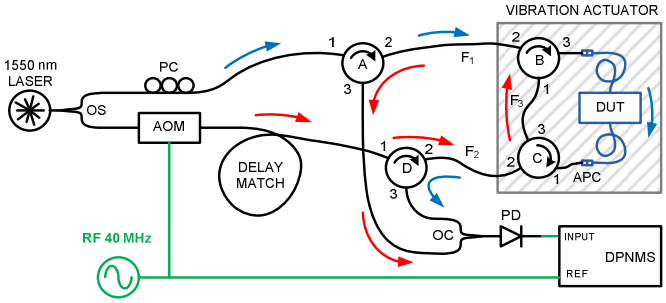


Fig. 2: Common-arm counter-propagating frequency-shifted interferometer for measuring vibration sensitivity of fiber-based optical devices.

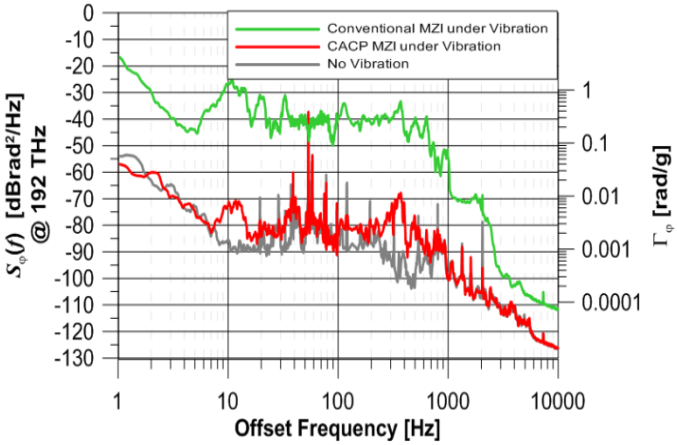


Fig. 3. The vibration sensitivity and residual phase noise floor of the conventional MZI and CACP-MZI under vibration. A random vibration with $S_g(f)$ = 1 mg^2/Hz is used for 10 Hz $\le f \le$ 1000 Hz. The bottom gray curve shows the noise floor measured under no vibration.

The noise floor for CACP-MZI system is measured under the same vibration condition as that of conventional MZI by replacing the DUT with 10 cm of SMF-28 fiber between ports 3 and 1 of circulators B and C. The noise floor for CACP-MZI (shown in red in Fig. 3) is limited primarily by the noise floor under no vibration. An improvement of more than 30 dB over conventional MZI noise floor is observed.

III. Test Results

The optical components considered for vibration sensitivity testing are listed in Table I and shown in Fig. 4. The fiber pigtails associated with these components are all 9/125/250 µm SMF-28 fiber. The buffer type and length for each DUT measured are indicated in Table I. For these DUTs, the associated buffer consisted of one of three types, namely, 900 µm tight buffer, 900 µm loose tube, and 3 mm jacketed buffer shown in 5a.

TABLE 1 - Devices considered for vibration test

Device Under Test	Buffer Type	Pigtail Length (m)
Power Splitter	900 um Loose Tube	1
Circulator	900 um Loose Tube	1
Etalon Filter (Finesse ~ 100, BW = 0.9 nm)	900 um Tight Buffer	1
Erbium-doped Fiber Amplifier (Gain = 23 dB, P1dB = 17 dBm)	900 um Tight Buffer	2
80 MHz AOM	3 mm Jacketed	2
Fiber potted in RTV Silicone	250 um Bare Fiber	0.1

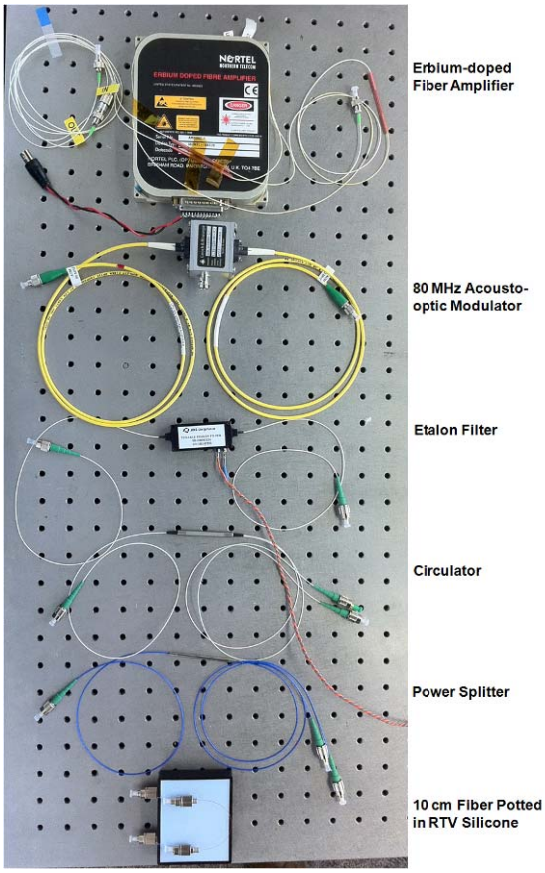


Figure 4. Picture of fiber-based optical components considered for the vibration tests at 1550 nm.

In order to understand the sensitivity contribution of the pigtail fibers to the DUT, we first measure Γ_φ of 1 m long optical patch cords of each buffer type. The patch cord is coiled in a 10 cm diameter and is taped on the shaker. For this test, $S_g(f)$ = 100 µg^2/Hz from 10 Hz $\leq f \leq$ 1000 Hz is used. For each fiber patch cord, Γ_φ varied depending on the amount of stress induced by the tape securing the fiber coil. Under different mounting conditions, about one to two orders of magnitude variation in Γ_φ is observed for a given fiber. The min-max plots showing this variation of Γ_φ for these three kinds of patch cords are shown in Fig. 4. The result indicates that the loose tube fibers, when constrained with mounting tape, can be less sensitive to vibration than the tight buffered fibers.

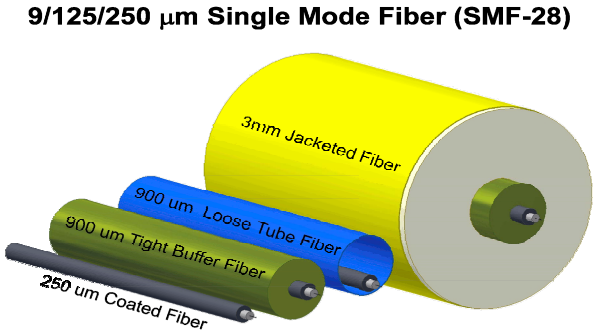


Figure 5a. Depiction of 9/125/250 µm SMF-28 fiber with different buffer types.

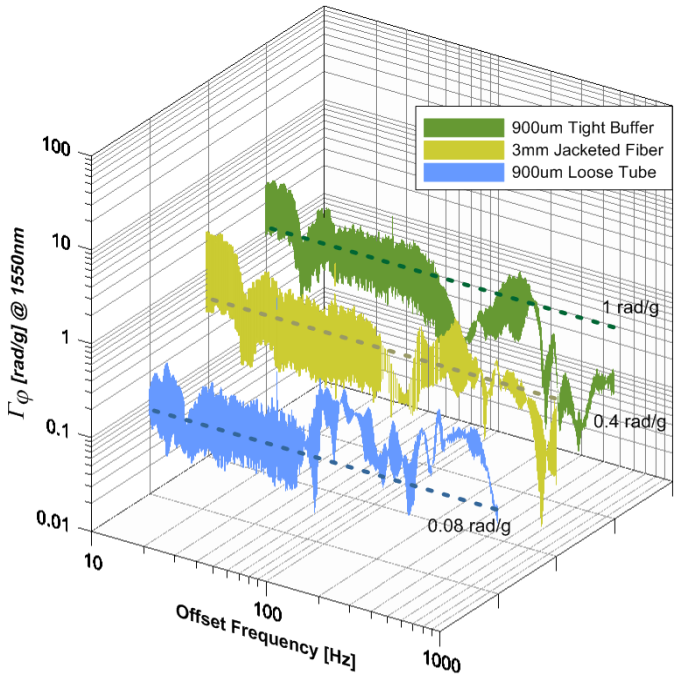


Fig. 5b. Min-Max plot showing the variation of the vibration sensitivity of 1 m long optical patch cords. The resonances above 100 Hz are contributions from the custom built vibration actuator and are not features of the DUT.

Next, the vibration sensitivity of fiber-based optical components are measured under the same vibration conditions, and results are shown in Fig. 6. The vibration is applied only along the z-axis, as shown in Fig. 1. Comparing Fig. 5b and Fig. 6 indicates that the vibration sensitivity of the power splitter, the Etalon filter and the AOM arises mostly

from the fiber pigtails, whereas for the circulator and the EDFA, their high sensitivity comes from the device itself, and not from the fiber pigtails. We also tested a 10 cm bare fiber potted in room temperature vulcanized (RTV) silicone. The sensitivity to vibration is found to be nearly two orders of magnitude worse than a non-potted bare fiber of the same length.

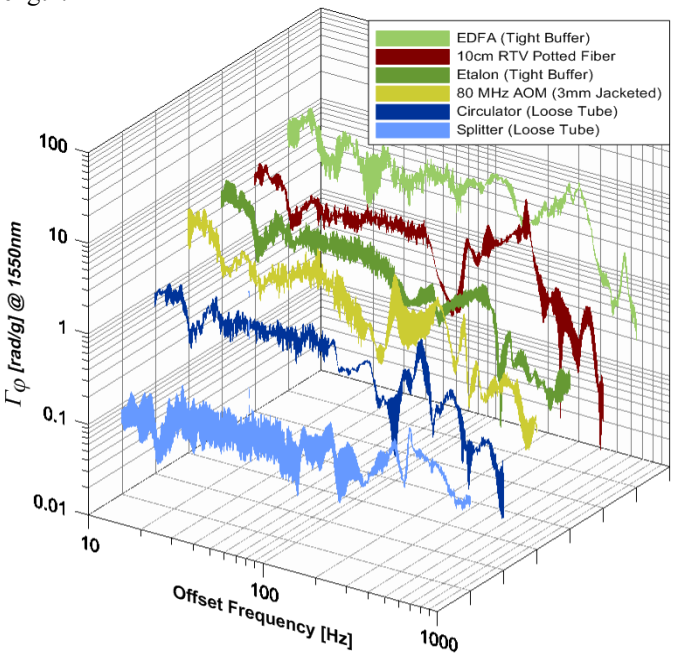


Fig. 5. Min-Max plot of the vibration sensitivity of the various optical components listed in Table I.

IV. Conclusions

We propose and demonstrate a technique to measure the vibration sensitivity of optical fiber-based components. It uses a common-arm counter-propagating MZI, that reduces the vibration-induced PM noise of the interconnecting fibers used in the measurement system. To our knowledge, this is the first implementation of this novel topology. This technique improves the vibration-induced noise floor typically by 30 to 40 dB compared to that of a conventional MZI approach. The proposed system has a sensitivity of nearly 1 mrad/g at 192 THz; therefore, it allows the measurement of low-vibration sensitive optical components. Vibration sensitivity of many optical components is dominated by the sensitivity of the associated fibers pigtails.

Constraining the optical fibers to prevent movement is often necessary for repeatability and consistency in fiber-based systems. Stress induced in the fiber due to this constrainment can dramatically increase vibration sensitivity. Different types of buffers used to protect the fiber from other environmental effects can either transmit or isolate the mounting stress to the underlying bare fiber. Loose tube buffering has shown less sensitivity to mounting stress than other buffer types.

Careful design of fiber-spool geometry has been used to minimize vibration sensitivity for long lengths of fiber [11]. The sensitivity of the spool has been reduced to levels approaching, or possibly below, the sensitivity due to the interconnecting fibers. If a vibration-insensitive spool of 1 km length were constructed, the connecting 3 mm jacketed fibers (Γ_φ = 0.4 rad/g) would limit the overall fractional vibration sensitivity of the spool to $\Gamma_y = 7\times10^{-11}$/g.

Methods similar to CACP-MZI can also be implemented either with polarization or frequency division multiplexing to combine the reference and measurement path in a common fiber. These common-arm MZI methods may find additional applications when it is critical to separate environmental noise of the delivery fibers from a remote fiber-based sensor.

In the future, a similar technique will be used to reduce the vibration sensitivity of coaxial cables in the microwave frequency range [12].

Acknowledgement

The authors gratefully thank Corey Barnes, Jason DeSalvo, Scott Diddams, Jeramy Hughes, Danielle Lirette, Franklyn Quinlan, Giorgio Santarelli, Bill Swann and Jennifer Taylor for useful discussions, assistance with equipments and preparation of this manuscript.

References

[1] T. Fortier, M. Kirchner, F. Quinlan, J. Taylor, J.C. Bergquist, Y. Jiang, A. Ludlow, C.W. Oates, T. Rosenband, and S.A. Diddams, "Photonic generation of ultrastable microwave signals," arXiv:1101.3616v2.

[2] X.S. Yao and L. Maleki, "Optoelectronic microwave oscillator," J. Opt. Soc. Amer. B, Vol. 13, no. 8, pp. 1725-1735, 1996.

[3] N. Ashby, D.A. Howe, J. Taylor, A. Hati, C.W. Nelson, " Optical fiber vibration and acceleration model," Proc. 2007 IEEE Intl. Freq. Cont. Symp., pp 547-541, 2007.

[4] J. Taylor, C.W. Nelson, A. Hati, N. Ashby, and D. A. Howe, "Vibration-induced PM noise measurements of a rigid optical fiber spool ," Proc. 2008 IEEE Intl. Freq. Cont. Symp., pp. 807-810, 2008.

[5] A. Hati, C.W. Nelson, J. Taylor, N. Ashby and D.A. Howe, "Cancellation of vibration-induced phase noise in optical fibers," IEEE Photonics Tech. Lett., Vol. 20, pp. 1842-1844, 2008.

[6] B. C. Young, F. C. Cruz, W. M. Itano, and J. C. Bergquist, "Visible lasers with subhertz linewidths," Phys. Rev. Lett. vol. 82, pp. 3799–3802, 1999.

[7] J. Millo, S. Dawkins, R. Chicireanu, D.V. Magalhaes, C. Mandache, D. Holleville, M. Lours, S. Bize, P. Lemonde, G. Santarelli, "Ultra-stable optical cavities: Design and experiments at LNE-SYRTE," Proc. 2008 IEEE Intl. Freq. Cont. Symp., pp. 110-114, 2008.

[8] D. Leibrandt, M. Thorpe, M. Notcutt, R.E. Drullinger, T. Rosenband, and J.C.Bergquist, "Spherical Reference Cavities for Frequency Stabilization of Lasers in Non-Laboratory Environments," Opt. Exp., vol. 19, pp. 3471-3482, 2011.

[9] F. Kefelian, H. Jiang, P. Lemonde, G. Santarelli, "Ultra Low Frequency Noise Laser Stabilized On Optical Fiber Spool", Proc. 2009 IEEE Intl. Freq. Cont. Symp, pp. 815-817, 2009.

[10] E. Baumann, F. R. Giorgetta, J. W. Nicholson, W. C. Swann, I. Coddington, and N. R. Newbury, "High-performance, vibration-immune, fiber-laser frequency comb," Optics Letters, vol. 34, Issue 5, pp. 638-640, 2009.

[11] T. Li, A. Haboucha, H. Jiang, J. L. Dournaux, D. Koné, C. Nelson, A. Hati, A. Clairon, P. Lemonde and G. Santarelli, "Low Vibration Sensitivity Fiber Spools for Laser Stabilization," presented at joint IFCS/EFTF meeting 2011.

[12] A.Hati, C.W. Nelson and D. A. Howe, "Vibration-induced PM and AM noise in microwave components," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 56, no. 3, pp. 2050–2059, 2009.

The Use of Multi-variance for Likelihood Weighted Drift Estimation in a Disciplined Oscillator Algorithm

Gregory Weaver, Mihran Miranian, Jeffrey Garstecki
Space Department
Johns Hopkins University Applied Physics Laboratory
Laurel, MD, USA
gregory.weaver@jhuapl.edu

***Abstract*—Our recent research at JHU/APL has focused on advancing algorithms and software control developments for a next generation, disciplined USO based clock system. This research shows promise that incorporating an in-situ control system for removing USO deterministic frequency error based on multi-variance methods and a Bayesian maximum likelihood estimation (MLE) process could mitigate the burden of post processing spacecraft clock data.**

We found the effects caused by sudden frequency changes and perturbations could be diminished with a dynamic, multi-variance algorithm used to augment the drift estimation process. We used the complementary aspects of the Allan and Hadamard variance to form an MLE function. We then used the output of the MLE function to drive the weighted decision process of a simulated frequency correction control algorithm.

Our paper will discuss the mathematical approach used to formulate the multi-variance MLE and discuss our findings on its implementation in the disciplining algorithm developed from the research. Using laboratory collected USO data, we will show simulated USO timekeeping performance to a maximum time interval error of < 1 μs over 70 days of operation and discuss the impact of frequency anomalies and short term rate changes to the behavior of the multi-variance approach.

I. Disciplined System for Improved USO Timekeeping

The Johns Hopkins University Applied Physics Laboratory (JHU/APL) has been providing ultra-stable oscillators (USOs) for decades for use in mission critical space applications where best performing frequency stability of less than 3E-13 in the 1 to 1000 s time interval is required for deep space science data return. The New Horizons mission and the Gravity Recovery and Interior Laboratory (GRAIL) are two most recent examples. The New Horizons mission to Pluto and the Kuiper Belt will use the frequency stability of the on-board USOs for a radio science experiment (REX) to determine any residual atmosphere that may exist from Pluto's most recent perihelion. The GRAIL mission will use the frequency stability of the USO in a Dual One Way Ranging (DOWR) instrument deployed in each of two tandem flying spacecraft to map the gravitational field of the Moon.

As described in these two examples, USOs are generally used as very-low phase noise frequency sources in science instruments. However, years of frequency data collected by ground mission operations has consistently shown that the frequency drift associated to quartz resonator aging over long periods of space operation can become good enough for consideration as timekeeping references. For example, frequency data collected from USO A on the New Horizons spacecraft determines a current frequency drift of -3.4E-12/day for a total accumulated frequency change of -19E-9 after 1800 days since launch, shown in Fig. 1 [1]. Also noticeable in Fig. 1 is a slow monotonic improvement in frequency drift with few frequency perturbations or anomalies, the most observable at day 400 from ionizing radiation associated with the spacecraft's encounter with Jupiter's magnetopause [2]. Such regular occurrence in USO performance for long-term frequency stability has motivated JHU/APL to support research into the development of a disciplined USO timekeeping system (TKS).

A disciplined USO TKS uses a remotely located reference of higher accuracy, in our case the master clock of ground

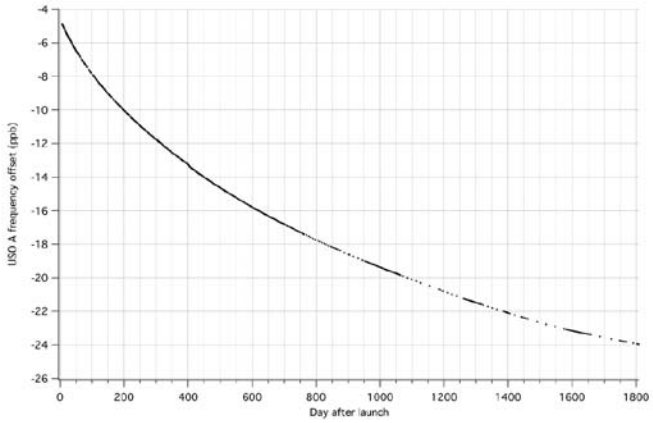


Figure 1. New Horizons USO A frequency since launch, Jan. 19 2006

control operations, to periodically correct the frequency of the USO output and maintain synchronization of the spacecraft to mission time. A disciplined USO TKS must perform two important functions. First, it must be able to transfer and then hold synchronization to mission time within some maximum time interval error (MTIE). Second, it must correct the USO frequency by removing its error during the time interval when access to the ground control master clock is not available. In USO's, this error can be partitioned between changing environmental conditions and dynamic drift (aging). These error contributions must be uniquely characterized to the USO frequency source in use. The information of the characterized USO is then used by the clock model of the disciplining control system to reduce the frequency error effects on the accumulated time.

Fig. 2 describes a notional system block diagram of our disciplined USO TKS. Our recent research shows, through simulation on archived USO test data, a possible USO timekeeping performance of less than 1 μs over 70 days of operation. This simulated performance improvement was accomplished through the development of three key technologies, indicated within the red dashed area of Fig. 2:

Steering algorithm for frequency drift correction using multi-variance based maximum likelihood estimator (MLE) function,

Clock model interface to control the operation of a Diophantine frequency synthesizer (DFS) [3],

Virtual state machine (VSM) autonomy controller to supervise the update implementation to the clock model DFS interface and alarm possible incidences of frequency perturbations or anomalies, as determined by the MLE algorithm [4].

This paper will discuss the theory, development and implementation of the MLE function from our research. We will describe our use of a three-state clock model derived in [5] and forward how its separation of deterministic frequency coefficients from Gaussian noise diffusion coefficients motivated our use of the dynamic Allan and Hadamard variances to detect USO frequency anomalies and perturbations.

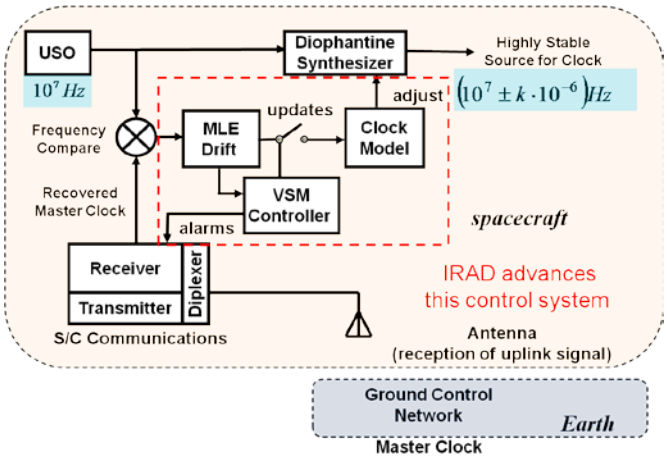


Figure 2. Block diagram of JHU/APL disciplined USO system

The information from these two related statistics was used to form a measure of likelihood that the frequency data collected over the most recent USO characterization interval was representative of previous clock model innovations. Dependent on this measure of statistical likelihood the influence of the data was selectively weighted for updating the clock model. It was found that the higher order power law noise of spectral density $S_y(f)$ proportional to f^{α}, where α is less than -2 are indicative of USO frequency data that is highly influenced by short-term frequency instability and should not be used with high significance in the clock model innovation. It was also found that a dynamic test comparing the statistics of the most recent characterization interval to a cumulative moving average could be used to check for the incidence of increased white FM, $\alpha = 0$, caused by outlier phase pops and frequency jumps.

II. The Clock Model and Its Relationship with the Allan and Related Variances

We use the title of [5] for this section heading as it clearly presents the motivation of our work in the construction of the MLE function to be described later in the paper. In [5], the authors derive a description of a three-state clock model, $d\mathbf{X}(t)$, from a generalized set of stochastic ODEs that represent the standard rate algebra of a clock with frequency offset, drift and a linear change of drift (μ_1, μ_2, μ_3 respectively). Gaussian noise diffusion processes are assigned to each equation with diffusion coefficients for white FM, random walk FM and random run FM (σ_1, σ_2, σ_3 respectively). The authors continue by showing that the solution of this set of strictly linear equations can be found in a closed form with a useful separation of the deterministic and stochastic parts:

$$\mathbf{X}(t) = \mathbf{\Phi}_{t,0}\mathbf{X}(t_0) + \mathbf{B}_t\mathbf{M} + \mathbf{G}_t, \tag{1}$$

where the matrix $\mathbf{\Phi}_{t,s}$ drives the transition $\mathbf{\Phi}_{t\to s,0}$, $\mathbf{M}$ is a vector of μ_1, μ_2, μ_3 and $\mathbf{G}_t$ is a vector that innovates only the stochastic part of $\mathbf{X}(t)$. Moreover, the authors develop an approach, offered by the separation of the deterministic and stochastic parts, to produce the clock model mean, $E[\mathbf{X}(t)]$, expressed in terms of a distribution of the vector $\mathbf{M}$ and clock model covariance, $E[\mathbf{G}_t\,\mathbf{G}_t^{\mathrm{T}}]$, expressed in terms of the distribution of the diffusion coefficients (σ_1, σ_2, σ_3). The mean, shown in (2) and the covariance, shown in (3) of the clock model are reproduced from [5], as follows:

$$E[\mathbf{X}(t)] = \begin{bmatrix} c_1 + (c_2+\mu_1)\,t + (c_3+\mu_2)\,\dfrac{t^2}{2} + \mu_3\dfrac{t^3}{6} \\ c_2 + (c_3+\mu_2)\,t + \mu_3\dfrac{t^2}{2} \\ c_3 + \mu_3 t \end{bmatrix} \tag{2}$$

$$E[\mathbf{G}_t\mathbf{G}_t^T] = \begin{bmatrix} \sigma_1^2 t + \sigma_2^2\frac{t^3}{3} + \sigma_3^2\frac{t^5}{20} & \sigma_2^2\frac{t^2}{2} + \sigma_3^2\frac{t^4}{8} & \sigma_3^2\frac{t^3}{6} \\ \sigma_2^2\frac{t^2}{2} + \sigma_3^2\frac{t^4}{8} & \sigma_2^2 t + \sigma_3^2\frac{t^3}{3} & \sigma_3^2\frac{t^2}{2} \\ \sigma_3^2\frac{t^3}{6} & \sigma_3^2\frac{t^2}{2} & \sigma_3^2 t \end{bmatrix}. \quad (3)$$

In subsequent sections of [5], the authors use the definitions of the Allan, $\sigma^2_y(\tau)$, and Hadamard, $H\sigma^2_y(\tau)$, variances expressed as expectations of the mean and covariance of $\mathbf{X}(t)$ to find the relationship of the clock model deterministic and stochastic coefficients to these two variances. The Allan, shown in (4) and the Hadamard, shown in (5) are reproduced from [5], as follows:

$$\sigma_y^2(t_k,\tau) = \left(\frac{\sigma_1^2}{\tau} + \frac{\sigma_2^2\tau}{3} + \frac{\sigma_3^2\tau^3}{20}\right) + \sigma_3^2\left(\frac{\tau^3}{3} + \frac{\tau^2 t_k}{2}\right) + \frac{\tau^2}{2}\left[c_3 + \mu_3(\tau + t_k)\right]^2 \quad (4)$$

$$H\sigma_y^2(\tau) = \frac{\sigma_1^2}{\tau} + \frac{\sigma_2^2\tau}{6} + \frac{11}{120}\sigma_3^2\tau^3 + \frac{\mu_3^2\tau^4}{6} \quad (5)$$

With respect to our work, two highly interesting observations can be made from (4) and (5). First, the deterministic linear change of drift coefficient, μ_3, is separable from the noise related parts of the statistics and contains contrasting information in the form of differing power law orders, τ in the case of the Allan deviation and τ^2 in the case of the Hadamard deviation. μ_3 is an important parameter for our interest in disciplining a USO since its dynamic behavior is the principle mechanism for the required periodic correction of our clock model process. Referring to Fig. 1, we expect that the μ_3 parameter should be small in respect to an update interval on the order of 12 to 24 hours. Therefore, we use this insight to form a piecewise comparative ratio, with the deviations segmented by octaves,

$$\mathrm{A} = \sum_{i=0}^{m}\frac{\sigma_y(2^{i+1}\tau,\mathrm{t_k})}{\sigma_y(2^i\tau,\mathrm{t_k})} \quad \mathrm{H} = \sum_{i=0}^{m}\frac{\mathrm{H}\sigma_y(2^{i+1}\tau,\mathrm{t_k})}{\mathrm{H}\sigma_y(2^i\tau,\mathrm{t_k})} \text{ to form } \frac{\mathrm{H}}{\mathrm{A}}, \quad (6)$$

where m is an integer limit to the number of $2^i\tau$, octave segments, that can be made in a t_k interval of data, such that $2^m\tau$ is accepted within the $\chi^2_{n,p}$ equivalent degrees of freedom, n, for a chosen confidence interval, p. We can then use this composite statistic to check dynamic intervals, t_k, for clock data where the relation $\frac{\mathrm{H}}{\mathrm{A}}$ exceeds a value, hypothetically indicating the incidence of a pronounced level of μ_3.

The second observation made from (4) and (5) is that as the deterministic coefficients of linear frequency drift approach zero, the power law relationships of $\sigma_y(t_k,\tau) \rightarrow H\sigma_y(\tau)$. In this manner, we are not so concerned about determining the level of the diffusion noise coefficients, σ_1, σ_2, σ_3, but that the relation $\frac{\mathrm{H}}{\mathrm{A}}$ approaches one. We interpret this result in two scenarios.

For a quartz-based USO, it is typically expected that c3 (or μ2) should always exist, but gradually diminish in a monotonic sense, as shown in Fig. 1, with some small value (slow function) of μ_3. So, in the very long term, the Allan deviation should approach the Hadamard and each of these identify the power law noises to a similar extent, although the Allan variance is twice as strong at identifying random walk FM, σ_2, than the Hadamard. For this scenario, the relation $\frac{\mathrm{H}}{\mathrm{A}}$ would start at a value less than one and evolve toward one. We consider this the nominal case.

The second scenario would be a slope reversal in the apparent drift of the USO. This sometimes is called "hockey-stick" aging where the drift starts in one sense (positive or negative) over an initial period in the order of a few weeks and then turns to the opposite sense over the longer term operation, usually in a yearly time frame. The cause for "hockey-stick" can be for various reasons, but it always means that two (or more) processes are competing in the composition of the USOs drift characteristic. These competing mechanisms may not be both intrinsic to the quartz resonator. For example slow temperature stabilization could overwhelm the apparent aging in the short term and then, once the system becomes quiescent, the resonator aging becomes dominant. It has been asserted that accumulated dose effects of ionizing radiation can be observed in space-based USOs as a competing process [6].

For our disciplining system, determining a possible "hockey-stick" drift evolution is highly desired as misidentifying this scenario as the nominal case would degrade the accuracy and the extent of autonomy of our USO timekeeping system. Our described work, then, concentrates on examining the dynamic behavior of the relation $\frac{\mathrm{H}}{\mathrm{A}}$ over data collection intervals, t_k to t_{k+1}. We venture from the basis of dynamic variance provided from [5] to formulate a maximum likelihood function (MLE) that we use to weight our periodic estimation of the deterministic coefficients, μ_1, μ_2, μ_3, and differentiate drift reversals and other short term frequency anomalies from the expected nominal case demonstrated in Fig.1.

III. A Maximum Likelihood Estimation Function (MLE) for Weighted Drift Estimation

Significant and elegant work has been accomplished, as in [7, 8 and 9] toward the use of multiple variances (multi-variance), Bayesian inference and the moment condition to separate, identify and determine the diffusion coefficients from a mixture of power law noises and deterministic drift in the spectral density of frequency sources such as USOs. We

adopted the notions of this work, in a less sophisticated approach, to construct a likelihood function using the relation $\frac{H}{A}$, along with our a priori knowledge about the evolution of $\frac{H}{A}$ from the two USO drift scenarios expressed in the previous section.

We consider our approach multi-variance in that $H\sigma^2_y(\tau)$ is shown in (5) to estimate the power law diffusion coefficients at different sensitivities than the $\sigma^2_y(\tau)$ in (4). In particular, the $H\sigma^2_y(\tau)$ assesses σ_1 for white FM the same as $\sigma^2_y(\tau)$; the $H\sigma^2_y(\tau)$ is half as sensitive for σ_2 random walk FM and about 1/3 as sensitive to σ_3 random run FM. In a strict multi-variance approach these definite relationships could be used to minimize uncertainty in power law identification and improve the accuracy of diffusion coefficient estimates. We did not perform this strict technique, rather simply opting for the unique identification of the $H\sigma^2_y(\tau)$ to the change of linear drift, μ_3, as a power of τ^4 and the knowledge that the classic, non-dynamic $\sigma^2_y(\tau)$ is not capable of resolving power law noises greater the random walk FM. As will be explained later in our implementation sections, we did use the aspect that the $H\sigma^2_y(\tau)$ and the $\sigma^2_y(\tau)$ react equally to white FM as a trap for phase pops and frequency jumps.

We consider our approach Bayesian, in that we estimate the extent of an unknown quantity, the change of linear drift μ_3, from measurements identifying μ_3 and the power law noises as found in the relation $\frac{H}{A}$. Further from the work of [5], we can express likelihood on the relation $\frac{H}{A}$ by setting μ_3 and certain power law noises to fixed values with an expectation of how μ_3 should evolve given our two a priori drift scenarios, the nominal case and the "hockey-stick" case. We did not perform the mathematics to derive the probability densities $p(\mu_3 \mid \frac{H}{A})$, $p(\frac{H}{A} \mid \mu_3)$ and the a prior probability density $\pi(\mu_3)$. However, we believe certain unique noise properties of the USO can nonetheless enable us to express the information provided by the relation $\frac{H}{A}$ into a Bayesian theory formulation.

Specifically, high quality USOs such as in New Horizons show a sharp contrast in the transition from flicker FM (F FM) with $\ln\tau^0$ dependency to linear drift with τ dependency in the Allan deviation. As the drift improves the extent of the flicker FM floor appears to increase with small evidence of random walk FM with $\tau^{1/2}$ dependency. This characteristic in USO noise offers the ability of the Allan deviation as segmented in (6) to highly distinguish USO drift from noise. This USO noise characteristic offers the additional advantage that we can place a high pass cut-off, through our selection of the base collection interval τ_0, to be aligned within the latter part of the flicker FM floor, nominally just high enough in frequency that the relation $\frac{H}{A}$ starts before the expected USO drift and nominally only considers noise with power law dependencies of $\tau^{\geq 0}$. We therefore will treat white FM as an anomaly in the expected result.

Finally we consider the $H\sigma^2_y(\tau)$ a sufficient variance as defined by the moment condition because its insensitivity to linear drift, $\mu 2$, provides for convergence for all power law noises up to random run FM. Also, since our timekeeping system is expected to act on daily or shorter updating periods, this places a low frequency limit that should cut-off the effects of higher order noises greater the random run FM.

We set four expected values, a minimum and a maximum for both the A and H defined in (6), to develop our Bayesian approach into a function that can operate as a maximum likelihood estimator (MLE) for the discrimination of the linear change of drift, μ_3. We construct the MLE with a bipolar selection effect for shifting the nominal case to the negative side of the function and the "hockey-stick" or high likelihood of μ_3 toward the positive. In this manner, we consider the positive indication in the hypothesis that μ_3 exists as anomalous to the expected nominal case and any use of the clock data collected in that update period should be deemphasized in the revision of the deterministic coefficients of the clock model. The four expected values are considered test limits, given as follows;

$$\mathsf{A}; \mathrm{E}_{\max}\left[\sum_{i=0}^{m}\frac{\sigma_y(2^{i+1}\tau, \mathrm{t}_k)}{\sigma_y(2^i\tau, \mathrm{t}_k)}\right] \Rightarrow 2m, \text{ drift} >> \text{noise} \quad (7)$$

$$\mathsf{B}; \mathrm{E}_{\min}\left[\sum_{i=0}^{m}\frac{\sigma_y(2^{i+1}\tau, \mathrm{t}_k)}{\sigma_y(2^i\tau, \mathrm{t}_k)}\right] \leq \sqrt{2}m, \text{ drift} < \text{noise} \quad (8)$$

$$\mathsf{C}; \mathrm{E}_{\max}\left[\sum_{i=0}^{m}\frac{\mathrm{H}\sigma_y(2^{i+1}\tau, \mathrm{t}_k)}{\mathrm{H}\sigma_y(2^i\tau, \mathrm{t}_k)}\right] \Rightarrow 4m, \mu_3 \text{ dominant} \quad (9)$$

$$\mathsf{D}; \mathrm{E}_{\min}\left[\sum_{i=0}^{m}\frac{\mathrm{H}\sigma_y(2^{i+1}\tau, \mathrm{t}_k)}{\mathrm{H}\sigma_y(2^i\tau, \mathrm{t}_k)}\right] \leq \sqrt{2}m, \sigma_2 \text{ or less} \quad (10)$$

Test limits A and B, (7) and (8), are the maximum and minimum expected values for the Allan deviation function A in (6). In test limit A, the expected value represents the result of the summed, octave-sectioned power law slopes where drift dominates over any other power law noises. Test limit B is the expected value of A in (6) where drift has diminished beyond $2^m\tau$ and only a random walk noise may

be evident. Again *m* is the maximum number of segments that can be statistically formed with a given confidence level in the update period, t_k. Test limits C and D, (9) and (10), are the maximum and minimum expected values for the Hadamard deviation function H in (6). Test limit C is the expected value when the change of linear drift, μ_3, dominates over any other power law slopes. Test limit D is similar to test limit B, but since H in (6) is insensitive to drift, D is also the likely expected result.

We use the test result limits set by A, B and C, D to construct a four zoned hypothesis based on the a prior knowledge about the two drift scenarios for USOs. Fig. 3 depicts this conception of our four hypothesis zones and the dynamic evolution of the USO drift scenarios. In Fig. 3, the four hypothesis zones are laid out in a quasi-coordinate manner with the dynamic evolution of the USO drift, as measured by the Allan deviation, shown in the vertical axis and unexpected changes in power law noise or the strong incidence of change of linear drift, μ_3 in the horizontal axis. The visualization of the dynamics of the USO drift and noise process is a vector moving from the nominal expected zone A, D (aging drift and low mostly- stationary noise processes) toward one of the other hypothesis zones. From our nominal case a priori scenario, such as shown in Fig. 1, we expect this vector to travel mostly vertical toward B.

The discrimination of the MLE function is then constructed to determine short term shifts of the visualized USO drift and noise process vector of Fig. 3 toward the horizontal axis. Such shifts are considered short-term anomalies and the significance of the clock data in the most recent t_k interval should be reduced in revising the clock model.

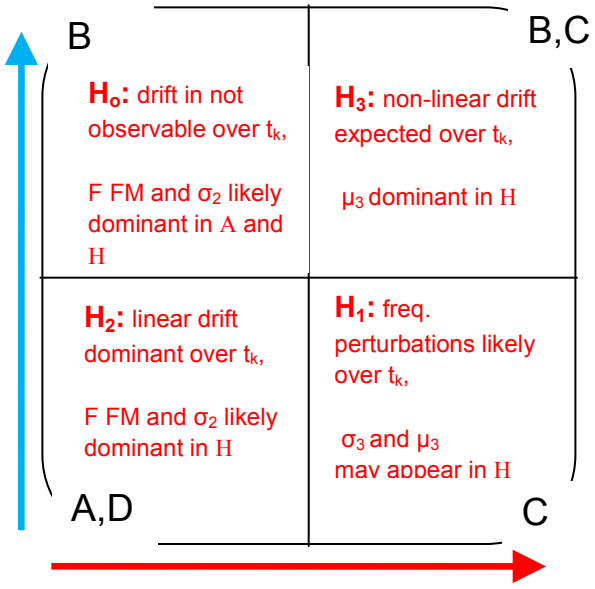


Figure 3. Four zones of MLE hypothesis based on USO drift evolution scenarios. Vertical blue arrow indicates nominal case evolution. Horizontal red arrow indicates "hockey-stick" case.

The test limits A, B and C, D, (7-10), and the four hypothesis zones H_0, H_1, H_2 and H_3 are assembled in the typical method of expressing an MLE function as the natural log of the test measurement ratios, in our case $\ln(\frac{H}{A})$ to provide a monotonic expression with the following formulations;

$$H_0 : \ln D - \ln B = 0, \quad \text{F FM or } \sigma_2 \text{ , drift not observable} \tag{11}$$

$$H_1 : \ln C - \ln A = \ln 2, \quad \sigma_3 \text{ or } \mu_3 \text{ , freq. perturbations likely} \tag{12}$$

$$H_2 : \ln D - \ln A = \ln \frac{1}{\sqrt{2}}, \text{ F FM or } \sigma_2 \text{ , linear drift dominates} \tag{13}$$

$$H_3 : \ln C - \ln B = \ln 2, \quad \mu_3 \text{ dominant , non - linear drift expected} \tag{14}$$

Fig. 4 graphically shows the MLE function $\ln(\frac{H}{A})$ as the relation $\frac{H}{A}$ changes from μ_2 drift dominated Allan deviation with low noise, as found by A of (6), moves to higher order μ_3 dominated linear change of drift as found by H in (6). Along the bipolar, monotonically increasing MLE function, the hypothesis zones are placed by the expected test limits (7-10) and their associated hypothesis definitions (11-14). Along the vertical secondary axis of Fig. 4 is placed the value of parameter W, which represents the assigned weighting to be applied to the update of the clock model coefficients, μ_1, μ_2, μ_3 using the function;

$$\mu_x(t_{k+1}) = \frac{W\mu_x(t_k) + \mu_x(t_{k+1})}{W+1} \tag{15}$$

The value of W in (15) gives the highest significance to the most recent clock data, t_{k+1}, when the MLE is less than $\ln\frac{1}{\sqrt{2}}$ or identified as likely in H_2, our nominal USO case.

The weighting function proceeds to de-emphasize the significance of the clock data in t_{k+1}as the MLE function moves positive. We assign a mid-grade weighting to W as the function transitions through zero, since the MLE function can have an ambiguity in the resolution of our two a priori scenarios in this region. In the nominal case, the drift slowly diminishes over a long time and it is expected that the MLE approach zero; as stated earlier, the power law noise of the Allan deviation should approach the Hadamard. However, in the "hockey-stick" scenario or with the introduction of a new competing drift process with a reversal in sign, the MLE function will also temporarily approach zero, but likely in the short term. We therefore diminish the possible outcome of the ambiguity by introducing a precautionary mid-grade significance. Above an MLE function value of zero, we give the data for t_{k+1} the least significance. Clearly with this approach, the number of weighting levels for W and the delimiter values associated to the MLE could be tailored.

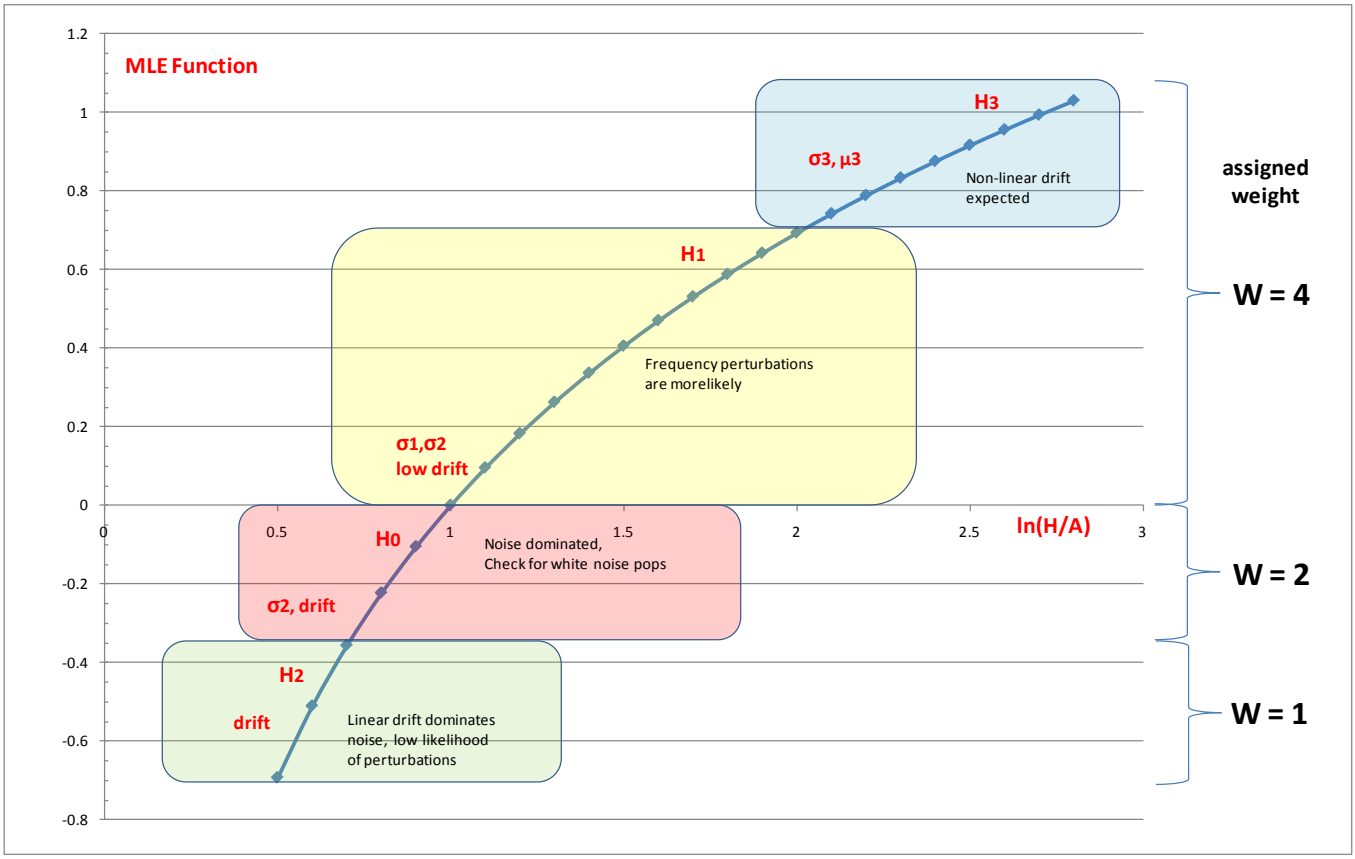


Figure 4. Graph of MLE function with hypothesis zones and weighting values for W on the secondary vertical axis

However, we chose to operate in this simple manner to examine how the process might work on real USO clock data, as will be discussed in the remaining sections.

But before we finish our paper with operational results, there is a special case we must consider in the MLE function associated with the incidence of white FM from phase pops or frequency jumps. We discussed earlier in this section that the base collection interval τ_0 was chosen to provide a high pass cut-off to filter the USO clock data so that the MLE function mostly operates on power law noises of $\tau^{\geq 0}$. Nonetheless in all real clock data, including the research USO data to be presented, there are instances of discontinuous phase and frequency that will issue large white FM noise into the collected data for that interval. Our MLE function is vulnerable to this condition since the Allan and Hadamard deviation are equally sensitive to white FM, σ1 in (4) and (5). Phase pops or frequency jumps that impart a white FM dominance into the data could wrongly place the MLE function into H_0 or H_2 and result in assigning the data more significance than desired.

We design a trap for this condition by using a dynamic analysis test, which compares the function A in (6) for the most recent collected data interval, t_{k+1} against a cumulative moving average of this function from previous intervals. A similar technique is used for H in (6). The expressions for this process are shown as follows;

$$\text{With } A(\tau, t_{k+1}, m) = \sum_{i=0}^{m} \frac{\sigma_y(2^{i+1}\tau, t_{k+1})}{\sigma_y(2^i\tau, t_{k+1})} \text{ and}$$

$$H(\tau, t_{k+1}, m) = \sum_{i=0}^{m} \frac{H\sigma_y(2^{i+1}\tau, t_{k+1})}{H\sigma_y(2^i\tau, t_{k+1})}$$

$$E\left[A(\tau, t_{k+1}, m) - \sum_{1}^{k} \frac{A(\tau, t_k, m)}{k}\right] = \pm\left|2-\sqrt{2}\right| \quad (16)$$

$$E\left[H(\tau, t_{k+1}, m) - \sum_{1}^{k} \frac{H(\tau, t_k, m)}{k}\right] = \pm\left|\sqrt{2}-1\right| \quad (17)$$

In (16) and (17) expected limits are set for the dynamic analysis tests. The limits are based empirically on our nominal dynamic drift scenario; the power law change from t_{k+1} to t_k is low compared to our update period. In (16), we expect the summed power law will change no more than one octave slope segment from drift to random walk FM. Likewise, in (17) the change is no more than one octave segment from random walk to flicker FM. In application to our research data, we only used the negative side of the test limits, such that when a result of (16) and (17) coincidentally fall outside the negative limits, we trap t_{k+1} as a white FM event and set W equal to 4. The A and H results for t_{k+1} are also not brought into the cumulative moving average of these functions.

IV. USO Frequency Data used in Research and Dynamic Analysis Results

We applied our theory of operation of the MLE function into an algorithm that would process periodically collected clock data. We set the processing period with a dynamic window variable called t_k, where t_k is set by an integer multiplication of the base τ_0 of the data measurement. For our research into the algorithm's effectiveness, we operated on 72 days of archived USO data. The USO data are phase measurements of a USO compared to a hydrogen maser with the USO maintained at 30 C in a space simulation environment with vacuum levels below 50 µtorr. The USO was powered on several days prior to the start of the data collection with vacuum pump down also occurring in this period. The phase collection measurement interval, τ_0, was 600 s.

Fig. 5 shows the relative frequency of the USO, in red, over the 72 day collection period with a fitted diffusion equation, $a+b(\tau_0 + c)^{½}$ shown in green. Fig. 6 shows the residual of the USO data with the fitted diffusion equation subtracted to more closely examine the dynamic behavior of the USO. The inspection of Figs. 5 and 6 reveals several aspects of the data that make it useful for our purpose of assessing the MLE based algorithm's performance in determining the μ_1, μ_2, μ_3 updates to our dynamic clock model for the purpose of disciplining the USO.

In Figs. 5 and 6, the USO data appears to contain the "hockey-stick" scenario of competing drift mechanisms. Additionally, the USO data contains various phase pops and short-term frequency perturbations that are traceable to anomalies in the vacuum and temperature control systems of the space simulation chamber, as sometimes occurs in long-term operation. The USO data also contains gaps in the data caused by discontinuity in the service provided by the maser signal. These gaps were subsequently filled into the data by least squares quadratic fits. The combined occurrence of all

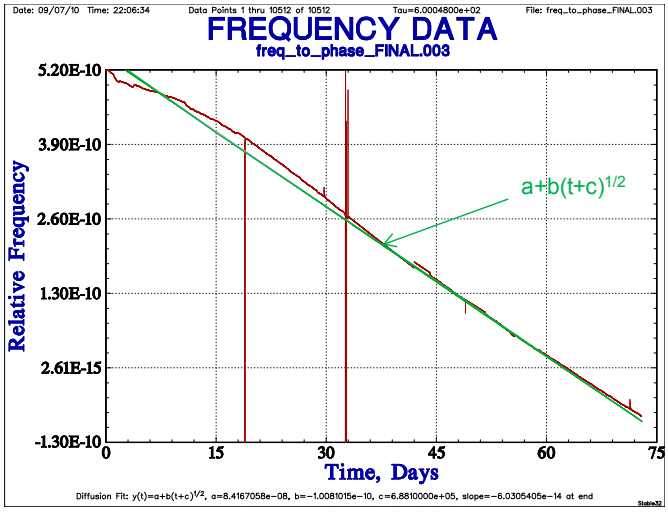


Figure 5. USO frequency data with diffusion equation

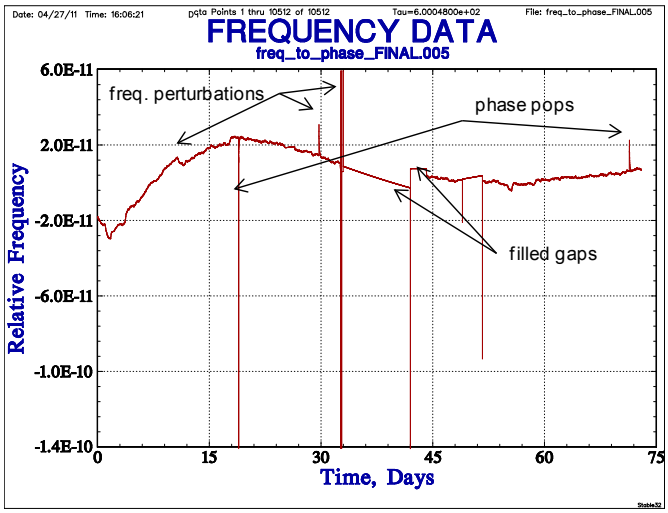


Figure 6. Residual frequency difference of USO data from diffusion equation

these data artifacts provides a challenging model for our disciplined USO algorithm research in contrast to the nominal scenario data such as shown in Fig. 1 from USO A of the New Horizons spacecraft.

We estimate the dynamic Allan and Hadamard variances for the USO data using one day windows, or t_k equal to $144(\tau_0)$. Fig. 7 shows the dynamic Allan deviation in a three dimensional presentation with τ and t_k (slices) shown on the floor projection and the frequency stability axis normal to this time plane. As discussed, we estimate the variance in τ by increasing octave sections from τ_0 to $2^5\tau_0$. Fig. 7 also shows a two dimensional inset with just the stability and t_k (slices) axis. The dynamic Allan deviation of the USO data shows a mostly constant drift with the stability proportional to τ. We observe the aspect of the "hockey-stick" process in the first fifteen t_k slices, noting the power law relationship dipping to lower order noise processes. We also see the effects of certain frequency anomalies coordinate with the indications shown in Fig. 6. These appear to have a mostly whitening effect as expected. The filled data gaps have an artificial appearance as the numerical noise from the calculated fits is not significant.

Fig. 8 shows similar results for the Hadamard deviation. In these presentations, we observe a mostly random walk FM. We also notice the effects of certain frequency anomalies similar to the Allan deviation of Fig. 7, although certain anomalies, such as t_k slices 29, 56 and 72 show indication of higher order noises expected with evidence of the linear change of drift, µ3. Figs. 9 and 10 show the density of the two point power law slopes for the Allan and Hadamard deviations, for each octave sections from τ_0 to $2^5\tau_0$. With careful inspection, the effect of the "hockey-stick" process in the first fifteen tk slices is apparent along with evidence of the frequency anomalies. These analysis results provide confidence that our MLE function should produce the effective discrimination desired.

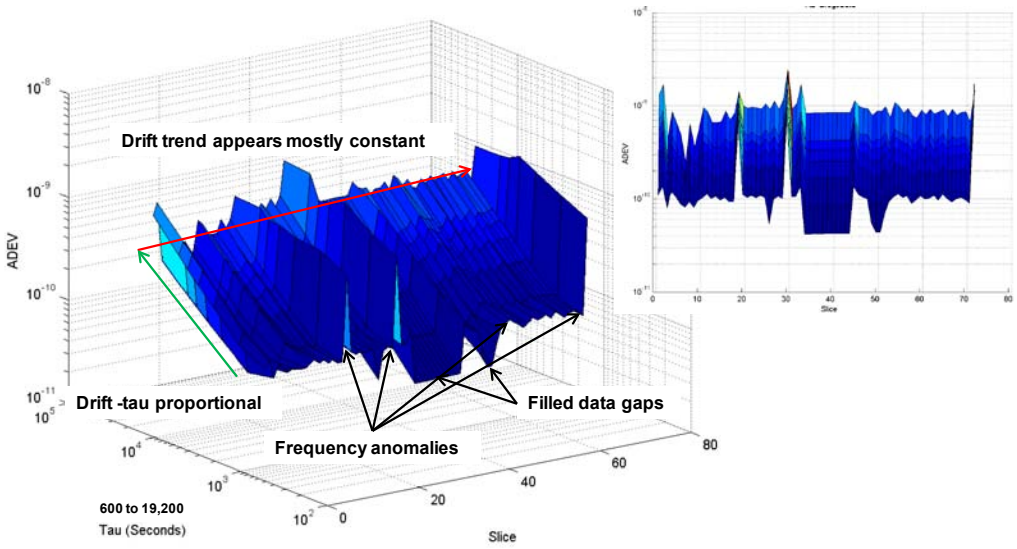


Figure 7. Dynamic Allan deviation: one day windows, $t_k = 144(\tau_0)$

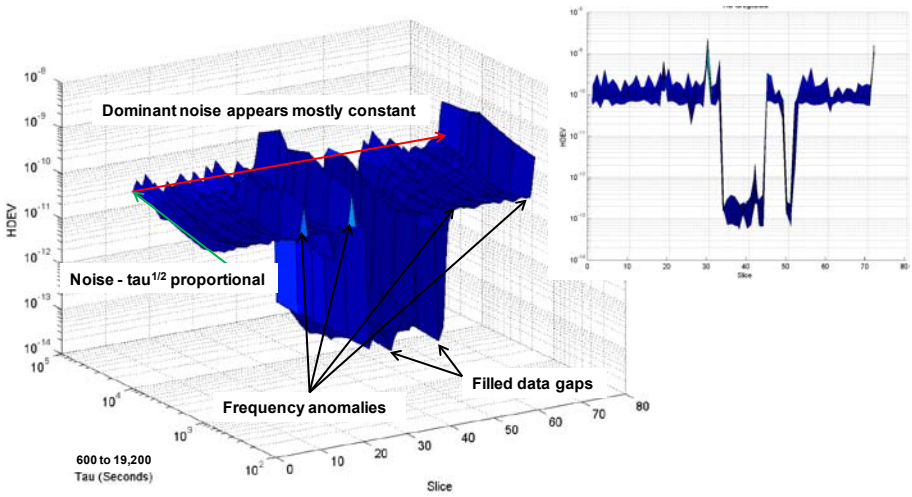


Figure 8. Dynamic Hadamard deviation: one day windows, $t_k = 144(\tau_0)$

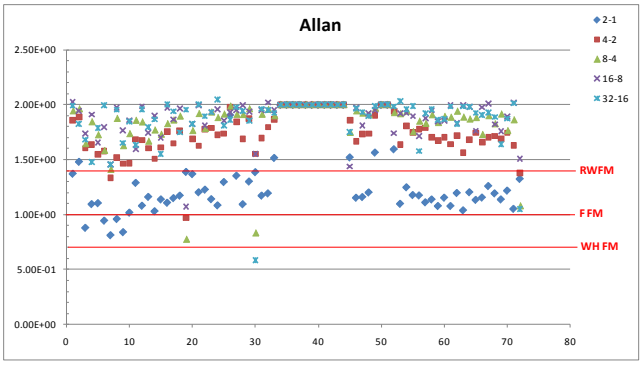


Figure 9. Density of octave power slopes for Allan deviation

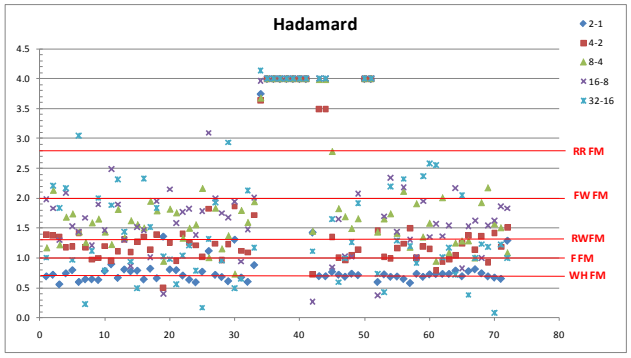


Figure 10. Density of octave power slopes for Hadamard deviation

We conclude by showing the performance of our MLE based algorithm operating on the USO data. Figs. 11 and 12 are similar in presentation as Figs. 5 and 6. However, in the case of Figs. 11 and 12, the green line represents the propagated clock model based on the MLE algorithm's calculation of the frequency offset, drift and a linear change of drift coefficients (μ_1, μ_2, μ_3 respectively). Fig. 12 shows the residual apparent corrected frequency with the propagated simulated frequency subtracted from the USO data. Finally, Fig. 13 shows the estimated timekeeping by converting the residual data from Fig. 12 to phase. The simulation indicates an effective timekeeping of the disciplined USO of about 1μs, including the effects of short-term frequency perturbations, phase pops and data gaps representative of the absence of a master reference clock. With the inclusion of a periodic phase adjustment, which would be available in our implemented system, the possible performance could maintain an MTIE of less than 1 μs over this same 72 day operation period.

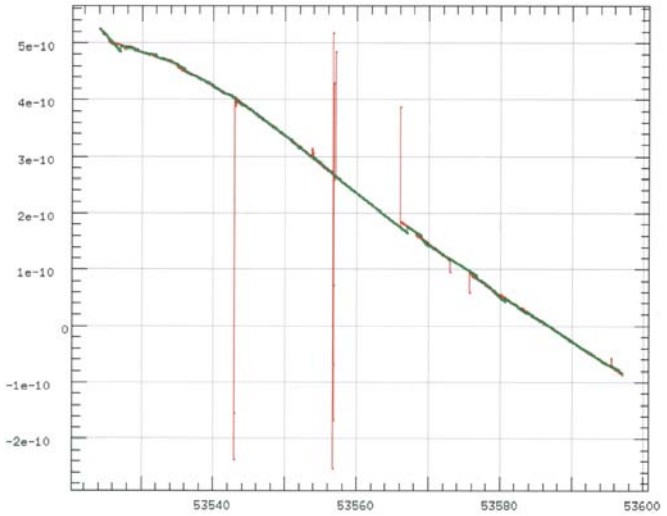


Figure 11. USO frequency data with propagated clock model in green

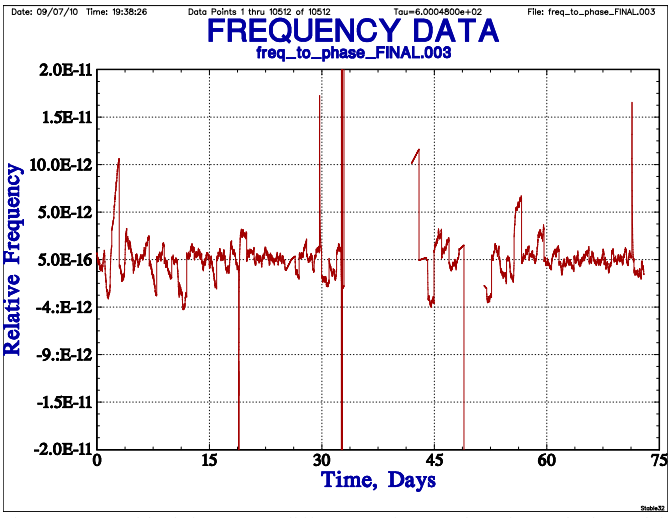


Figure 12. Residual frequency difference of USO data from clock model propagation

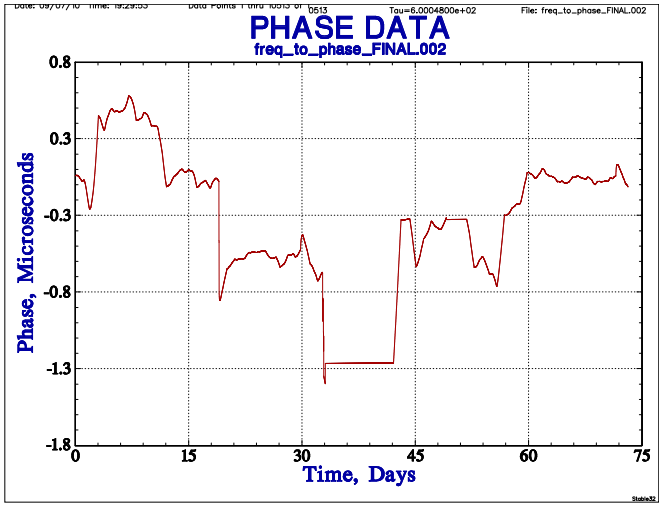


Figure 13. Simulated timekeeping performance of MLE algorithm

REFERENCES

[1] J. R. Jensen, "New Horizons USO A and USO B frequency drift histories from launch," JHU/APL internal memo., unpublished, February 2011.

[2] J. R. Jensen, and G. L. Weaver, "The In-flight frequency behavior of two ultra-stable oscillators on-board the new horizons spacecraft," 39th Annual Precise Time and Time Interval (PTTI) Meeting, , pp. 79–93, November 2007.

[3] P. Sotiriadis, "Diophantine frequency synthesis," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 53, No. 11, pp. 1988–1998November 2006.

[4] G. Cancro, W. Innanen, R. Turner, C. Monaco, and M. Trela, "Uploadable executable specification concept for spacecraft autonomy systems," IEEE Aerospace Conf. Proc., paper #1201, March 2007.

[5] C. Zucca, and P. Tavella, "The clock model and its relationship with the Allan and related variances", IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 52, no. 2, pp. 289–296, February 2005.

[6] M. Bloch, O. Mancini, and T. McClelland, "Effects of radiation on performance of space-borne quartz crystal oscillators," Proc. European Frequency and Time Forum & IEEE International Frequency Control Symposium - 2009 Joint Conference, pp. 171–174, April 2009.

[7] F. Vernotte, E. Lantz, J. Groslambert, and J. J. Gagnepain, "Oscillator noise analysis: multivariance measurement," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 42, no. 2, pp. 342–350, April 1993.

[8] F. Vernotte, and G. Zalamansky, "A Bayesian method for oscillator stability analysis", IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 46, no. 4, pp. 508–513, April 2002.

[9] F. Vernotte, "Application of the moment condition to noise simulation and to stability analysis," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 49, no. 2, pp. 289–296, February 2005.

Reliability of Next Generation High Performance pMEMS™ Resonator Oscillators

Harmeet Bhugra, Seungbae Lee, Wanling Pan, Ye Wang, Dino Lei
Integrated Device Technology Inc. (IDTI)
San Jose, CA 95138
Email: {Harmeet.Bhugra, Seungbae.Lee, Wanling.Pan, Ye.Wang, Dino.Lei}@idt.com

***Abstract*—Long term reliability of piezoelectric micro-electro mechanical systems (pMEMS™) resonators and oscillators has been characterized for high performance frequency reference applications. In particular, high-*Q* piezoelectric bulk acoustic wave (BAW) resonators exhibit aging (frequency drift) within measurement system accuracy (±2ppm) under two long-term measurement conditions: over 9 months at 25°C, and over 1400 hours at 125°C.**

I. INTRODUCTION

Quartz crystal resonator oscillators have been used for frequency reference applications due to their excellent performance, such as high *Q*s and stable frequencies for the past few decades. Over the past few years, however, many efforts have been underway to replace quartz with other technologies with [1][2][3] or without MEMS resonators [4]. These disruptive technologies enable reductions of package sizes and cost, and in some cases enable significant improvement in performance and reliability. Among commercialized oscillator products based on MEMS resonators, electrostatically transduced vibrating resonator oscillators have been the first to be introduced into the timing markets. Compared to quartz counterparts, however, their performance needs to be improved especially for high-end frequency reference applications.

As a member of MEMS resonator family, pMEMS™ resonator has emerged as one of the strong candidates for replacing quartz technology for high-end timing and frequency reference products [1][5][6]. These pMEMS™ resonator oscillators have demonstrated short term frequency stability and phase noise on par with quartz crystal oscillators. However, for successful adoption in timing and frequency reference markets these pMEMS™ resonators and oscillators need to pass stringent qualification tests such as the ones specified by JEDEC standards [7] among others [8][9]. Robustness of these low phase noise oscillators has been demonstrated by subjecting these plastic packaged devices to a large range of reliability tests [7]. Focus for this paper is on aging, i.e., frequency drift over time.

Oscillator frequency drift over time is mostly a function of the resonator, i.e., quartz resonators or in this case pMEMS™ resonators. This paper presents the aging (frequency drift) test results of the pMEMS™ resonators and oscillators at both normal (25°C) and accelerated temperature conditions (125°C).

Figure 1. Teardown of a pMEMS™ oscillator sample in a plastic package

II. PMEMS™ RESONATOR

As the name implies, the pMEMS™ resonator is composed of a piezoelectric material with single crystal silicon, taking advantage of both worlds, i.e., combining advantages of piezoelectric quartz resonators with advantages of MEMS resonators. Specifically, pMEMS™ utilizes a piezoelectric material for its high coupling between electrical and mechanical interaction, and a single crystal silicon for the high quality factor as the main vibration body that dominates the frequency selective performance [1][6]. Figure 2 presents the schematic of a pMEMS™ resonator with a piezoelectric layer and a single crystal silicon layer. By electrically exciting bottom and top electrode layers using ac-electric signals, the piezoelectric material vibrates mechanically at the designed frequency transferring the acoustic wave energy to single crystal silicon layer. Making use of its excellent frequency selectivity, the single crystal silicon layer in turn bounces most of the acoustic energy back to the piezoelectric layer where

978-1-61284-110-6/11/$26.00 ©2011 IEEE

the acoustic energy changes into electric signals sensed at the top electrode. Due to the nature of silicon as the main element of the vibrating structure, this pMEMS™ resonator provides sufficient quality factors (Q's > 10,000)[1]. With the proper selection of the piezoelectric material as the transformer of the two energy domains, it eliminates the need for external DC biases required in capacitively transduced MEMS resonators [10][11], and enhances high energy transmission ratio enough to implement practical high frequency timing and frequency reference products.

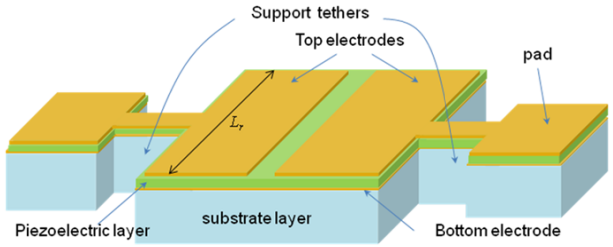


Figure 2. Perspective schematic view of a pMEMS™ resonator

For the aging reliability tests in this paper, very high frequency (>100 MHz) pMEMS™ resonators were selected. All the selected pMEMS™ resonators vibrate extensionally in lateral direction at the following frequency,

$$f_0 = \frac{1}{2 \cdot L_r}\sqrt{\frac{E_{eff}}{\rho_{eff}}} \qquad (1)$$

where L_r is the lateral length, E_{eff} the effective Young's modulus, and ρ_{eff} the effective density of the composite pMEMS™ resonator. Once the materials of pMEMS™ are set, the only parameter that determines the frequency is the lateral dimension of pMEMS™ resonator, meaning that the resonator frequency is easily designed by a single device layout and implemented with etchings of the composite material layers.

III. Setup for Reliability Tests

In order to accommodate multiple pMEMS™ resonators in a single aging setup, multiplexers, 3499A and 34980A from Agilent, are employed to switch both RF signals and low frequency temperature sensors, respectively. Each multiplexer for RF signals requires five quad dual switch modules (N2268A from Agilent), four of which were used for first stage of switching and the last one for multiplexing the first stage four switch modules. Each switch module has 50Ω characteristic impedance with SMA connections. The multiplexer for accurate temperature measurements required one low frequency switch module (34921A from Agilent), and it not only switches the RTD (Resistance Temperature Detector) sensors but also converts the RTD resistance reading to temperature data. Therefore, each set of multiplexers can take up to 16 units in parallel, and the implemented aging setup consisted of two sets of multiplexers and network analyzers. The overall measurement system shown in Figure 3 is controlled by an automatic program through GPIB interface to take measurements and extract data, such as resonance frequency, quality factor, insertion loss, etc., from each pMEMS™ resonator. Due to the high sensitivity of resonance frequency on temperature it is very important for the pMEMS™ resonators to be accurately measured with respect to surrounding temperature. Therefore, in this aging test setup, temperature data are measured from platinum RTD sensor readings at the time of frequency measurements and then later used to extract aging data.

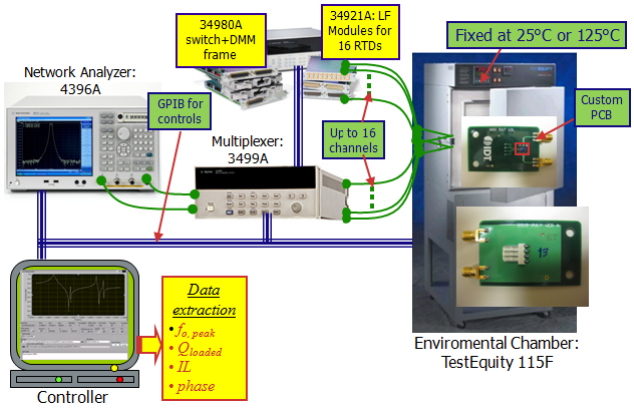
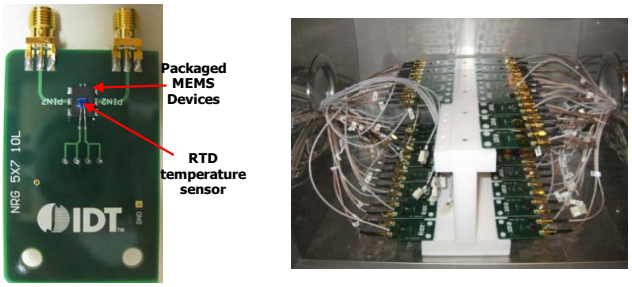


Figure 3. Schematic of Aging Test Setup

Figure 4 (a) shows a customized PCB board where a packaged pMEMS™ resonator / oscillator and a platinum RTD are soldered. Note that an RTD platinum sensor is epoxy-glued on top of the plastic package to measure the temperature at the time of a single frequency sweep as accurately as possible. Total of 32 boards were mounted to a Teflon rack and put into a temperature chamber shown in Figure 4 (b).



(a) Test board (b) Test boards on a rack

Figure 4. PCB test boards mounted on a rack inside temperature chamber

All the materials used in this setup can withstand high temperature of 125°C to cover acceleration tests. The reliability test systems have been set up at IDT's long term MEMS test lab shown in Figure 5.

Figure 5. MEMS lab for the long term aging tests of pMEMS™ resonators and oscillators at Integrated Device Technology, Inc.

IV. Results and Discussions

A. Aging of pMEMS™ resonators at 25°C

In the aging setup, ten pMEMS™ resonators were measured in the temperature chamber set to 25°C for approximately nine months, while their frequency data was recorded automatically with open-loop configurations. To extract the aging performance of the pMEMS™ resonators, the frequency fluctuations with respect to temperature variations within the temperature chamber were corrected using linear least square regression.

Figure 6 below presents the plot for extracted aging data for ten pMEMS™ resonators vibrating at high frequency. All of the ten pMEMS™ resonators regardless of the device locations have frequency variations less than ±2ppm for nine months.

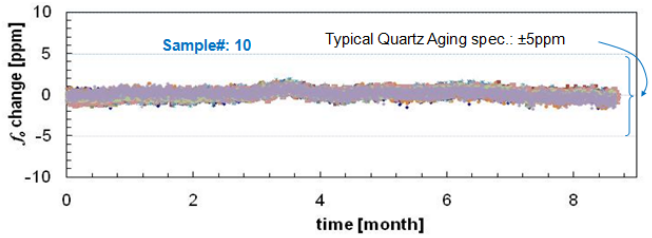


Figure 6. Aging measurements of 10 pMEMS™ resonators / oscillators over nine months at 25°C (Measurement accuracy ±2ppm).

Of the ±2ppm frequency changes, ±1ppm is attributed to the accuracy of the network analyzer. We observed the frequency changes with discontinuous quantum leaps, which correspond to the frequency resolution of the network analyzer at a given sweep span with a fixed number of data points. It is worth noting that the compensated frequencies are somewhat correlated with temperature fluctuations, suggesting that the rest of ±1ppm of the total ±2ppm frequency variation likely stems from two factors as follows.

1. Temperature gradient between the pMEMS™ resonator and an RTD sensor location: RTD temperature sensor measures the temperature on top of the plastic package not at the actual pMEMS™ resonator inside of the package.
2. Inaccurate timing of the temperature sensing: Temperature reading is based on the average by the integration of digital multi-meter 34980A, which means that the reading does not represent the temperature exactly at the time of resonance frequency detection. To make it worse, as the temperature easily changes ±0.02°C in a short period of time, there is an uncertainty in the measured temperature.

B. Acceleration tests for pMEMS™ resonators at 125°C

In order to accelerate the aging tests, ten packaged pMEMS™ resonators were placed for more than 1400 hours inside the temperature chamber set to 125°C which is the temperature required for accelerated burn-in tests in the JEDEC standards [7]. Figure 7 presents the extracted aging data showing less than ±2ppm frequency variation at 125°C. This suggests that the pMEMS™ resonators are very robust and can withstand high temperatures over long periods of time that exceeds any commercial application requirements.

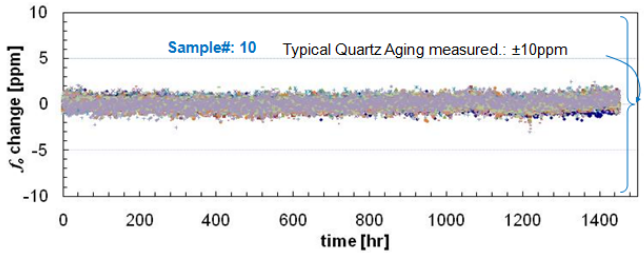


Figure 7. Accelerated aging measurements of 10 pMEMS™ resonators / oscillators over 1400 Hours at 125°C (Measurement accuracy ±2ppm).

For comparison, a typical AT-cut quartz crystal resonator was measured along with pMEMS™ resonators at 125°C. Since it was not easy to find quartz crystal resonators that were commercially available and rated for 125°C operation, a quartz resonator specified at 85°C was selected. The quartz measurements exhibit frequency drifts of ±10ppm with a negative slope over 1400 hours compared to ± 2ppm measured for pMEMS™ devices. These ±10ppm quartz oscillator measurements are comparable to burn-in test data published by multiple commercial crystal oscillator vendors.

V. Conclusions

Robustness of low phase noise pMEMS™ resonators and oscillators has been demonstrated by subjecting these plastic packaged devices to long term aging (frequency drift) tests. pMEMS™ devices exhibit aging (frequency drift) comparable to quartz at 25°C and significantly better at 125°C (accelerated tests). Long term reliability tests presented in this paper suggest that the pMEMS™ resonator oscillators are as robust as quartz resonator oscillators in the long-term at nominal temperatures (25°C) and much better over a wide temperature range.

References

[1] S. Humad, R. Abdolvand, G. Ho, G. Piazza, and F. Ayazi, "High frequency micromechanical piezo-on-silicon block resonators," IEEE Int. Electron Devices Meeting, Dec. 2003, pp.39.3.1-39.3.4.

[2] W.-T. Hsu and A. R. Brown, "Frequency trimming for MEMS resonator oscillators," Proceedings of 2007 IEEE Frequency Control Symposium, Switzerland, May 29-June 1, 2007, pp.1088-1091.

[3] B. H. Stark, E. Radza, P. Gupta, J. Sharma, G. Peramaih, S. Deng, S. Suen, R. Sheridan, A. Patridge, and M. Lutz, "An ultra-thin packaged MEMS oscillator," Technical Digest of the IEEE Solid-State Sensor Actuator Workshop, Hilton Head, SC, June 1-5, 2008, pp.6-9.

[4] M. S. McCorquodale, *et al*, "A silicon die as a frequency source," Proceedings of 2010 IEEE Frequency Control Symposium, Newport Beach, California, June 2-4, 2010, pp.103-108.

[5] M. Pardo, L. Sorenson, W. Pan, and F. Ayazi, "Phase noise shaping via forced nonlinearity in piezoelectrically actuated silicon micromechanical oscillators," 24th IEEE International Conference on Micro Electro Mechanical Systems, Cancun, Mexico, Jan. 23-27, 2011, pp.780-783.

[6] A. Jaakkola, P. Rosenberg, A. Nurmela, T. Pensala, T. Riekkinen, J. Dekker, T. Mattila, and A. Alastalo, "Piezotransduced single-crystal silicon BAW resonators," IEEE Ultrasonics Symposium, New York, NY, U.S.A., Oct. 28-31, 2007, pp.1653-1656.

[7] JEDEC SOLID STATE TECHNOLOGY ASSOCIATION, JESD47H (Revision of JESD47G.01, April 2010), "Stress-Test-Driven Qualification of Integrated Circuits", Feb 2011.

[8] R. N. Candler, M. A. Hopcroft, B. Kim, W. Park, R. Melamud, M. Agarwal, G. Yama, A. Patridge, M. Lutz, and T. Kenny, "Long-term and accelerated life testing of a novel single-wafer vacuum encapsulation for mems resonators," *J. of Microelectromechanical Systems*, pp. 1446–1456, 2006.

[9] V. Kaajakari, J. Liihamaki, A. Oja, H. Seppa, S. Pietikainen, V. Kokkala, and H. Kuisma, "Stability of wafer level vacuum encapsulated single-crystal silicon resonators," *The 13th International Conference on Solid-State Sensors, Actuators and Microsystems* (Transduers'05), Seoul, Korea, June 5-9, 2005, pp.916-919.

[10] V. Kaajakari, T. Mattila, A. Oja, J. Kiihamaki, H. Kattelus, M. Koskenvuori, P. Rantakari, I. Tittonen, and H. Seppa, "Square-extensional mode single-crystal silicon micromechanical RF-resonator," *The 12th International Conference on Solid-State Sensors, Actuators and Microsystems* (Transduers'03), Boston,MA, USA., June 8-12, 2003, pp.951-954.

[11] F. Bannon, J. Clark, and C. T.-C. Nguyen, "High-frequency micromechanical filters," *IEEE, J. Solid-State Circuits*, Vol. 35, no. 4, pp.512-526, April 2000.

Slow Wave Resonator Based Tunable Oscillators

Ajay K. Poddar[1,2], *Senior Member IEEE*
[1]Synergy Microwave Corp., NJ 07504, USA
[2]Technical University Munich, Germany

Ulrich L. Rohde[1,2], *Fellow IEEE*
[1]Univ. of Cottbus, BTU Cottbus 03046, Germany
[2]Synergy Microwave Corp., NJ 07504, USA

Abstract—**Demands for high-speed multimedia data communications increasing telescopically, pushing wireless transmission data rate with 2Gbps or higher. In these high-speed data communication systems, the bit error rate (BER) characteristic is dependent on a phase noise performance of the VCO. To meet the above challenges, it is necessary for these systems to have high performance VCO with low phase noise, low cost, wide tuning range, and high DC-to-RF conversion efficiency. The tiny VCO developed using SWR (Slow Wave Resonator) and mode-coupling techniques support high DC-to-RF conversion efficiency and low phase noise in compact size and also easily amenable for integration in the IC (integrated circuit) form. The typical measured phase noise @ 10 kHz offset from the 12 GHz carrier is -110 dBc/ Hz with 5V, 30 mA DC bias using slow wave dynamics in multi-mode regimes.**

I. INTRODUCTION

Tunable oscillators are instrumental in the operation of many systems, from commercial communications to military radars. And, while many characteristics define the performance of a tunable RF/microwave oscillator, one of the more difficult parameters to optimize is phase noise, which is critical to the performance of so many systems. Communication systems rely on low-phase-noise signal sources such as tunable oscillators or phase-lock-loop (PLL) synthesizers for reliable voice communications and to ensure transmitted data integrity. Constraints on cost, size, power-efficiency, performances are particularly demanding. As data requirements increase beyond 2 Gb/s, the phase noise of one possible tunable source solution, the voltage-controlled oscillator (VCO), becomes critical for achieving acceptable bit-error-rate (BER) performance [1]-[22].

VCOs are versatile, tunable sources, their resonators can be formed in a variety of ways, including by printing coupled resonators. VCOs reported in this paper using printed coupled slow wave resonator (SWR) measure just 0.3 x 0.3 in. but can match the phase-noise performance of more expensive oscillators with high-quality-factor (high-Q) resonators. The SWR network approach also lends itself to semiconductor fabrication processes, to produce high-performance VCOs in integrated-circuit (IC) form. The tiny SWR VCOs reported in this paper are ideal sources for miniaturized PLL frequency synthesizers and provide outstanding bandwidth versus phase noise performances stability by locking in frequency to a reference source as part of a PLL frequency synthesizer. In addition to this, these novel VCO solutions are inherently immune to shock and vibration due to printed coupled resonators. It is to note that SWR design parameters (coupling coefficient, multi-mode sensitivity, and Q factor) can be obtained using full wave EM simulation. It has been found that degenerate modes play important role in improving the Q-factor, therefore, rigorous mathematical treatment is necessary for the minimization of the spurious and jitters in pass-band. The theoretical treatment and experimental validation of a new class of slow wave resonator suggests the possibility of cost-effective alternative of Dielectric and YIG resonator based signal sources [1]-[9] .

II. THEORY

As the frequency band for the wireless communication shifts higher, generation of the power efficient ultra low noise wideband and thermal stable compact signal sources with low cost become more and more challenging due to the frequency limitations of the active devices. The technique of harmonic generation has been proven to be a cost-effective solution to extend tuning range of the oscillator circuits [2]-[5]. These techniques are based on the nonlinearity of the transconductance of the active device under large signal drive conditions but it introduces distortions and has poor phase noise performances [6]-[7].

YIG oscillators are well known for broadband configurable signal source but at the cost of size, power and integrability in IC form [8]-[11]. In view of the limitations of known YIG resonators for integrability and power-effective operation, the present work describes a novel and innovative topology as slow wave resonator (SWR) VCO, which supports wideband tunability in compact size and easily amenable for integration in the integrated chip form. SWR VCO is planar structure and broadband in nature, and is well suited for MMIC fabrication with affordable price. With the potential to enable wide operational bandwidths, eliminate discrete resonators (YIG sphere), and produce high Q planar resonators for ultra low noise VCOs in the context of a planar fabrication process compatible with existing IC and MMIC processes, SWR VCO is widely believed to be promising technology to meet the present and future requirements [22].

To solve the frequency-agility issue, radio manufactures that serve the point-to-point and point-to-multi-point markets would prefer to use SWR VCO that could emulate

978-1-61284-110-6/11/$26.00 ©2011 IEEE

configurability, wideband tunability and low phase noise, yet eliminate the crude YIG based signal source; and deliver higher transmission speeds, cost and power-effective solutions for present and later generation of the communication systems [2].

A. Multi-Coupled Slow-Wave Resonator

A printed coupled resonator network realized by slow wave dynamics is attractive due to compact size and provides a wide spurious-free band. In addition to this, the physical layout of slow wave resonators enables the implementation of optimum EM couplings [10]. As a result, resonators are not only compact size due to slow wave effect but also exhibit high frequency selectivity, wider stop-band resulting from dispersion phenomena. In order to facilitate a broad yet precise description of the VCO topology and layout of the SWR structure is selected in such a way that it minimizes the effect caused by temperature and mechanical stresses, and supports uniform negative resistance over wide tuning range.

In slow-wave propagation, as its name implies, the electromagnetic wave is transmitted in the guided-wave media with a slower phase velocity, namely, shorter guided wavelength, at a specified operating frequency [8]. This is achieved by modifying electric and magnetic energy storage in the guided-wave media. Slow-wave resonators are attractive due to several reasons: the slow-wave effect makes them very compact and can support evanescent mode coupling for obtaining Q-multiplier effect at resonant condition, thus low phase noise signal source solutions. Additionally, the geometrical configuration of SWR resonators makes possible the implementation of compact layout and amenable for RFIC/MMIC realizations [9]-[10].

A slow wave structure is one of the concepts that can be used for reducing the size of the circuits, especially the length of the transmission lines. For the electrical circuits, the periodic shunt loading of the transmission lines can behave as a slow wave structure. This periodic shunt loading reduces the phase velocity, thereby increasing the effective electric length of the line. In conventional transmission lines, the phase velocity V_p is controlled only by the dielectric material and can be expressed as [14]

$$V_P = f \times \lambda = \frac{c_0}{\sqrt{\mu_r \varepsilon_{reff}}} \tag{1}$$

where c_0 is the velocity of light, μ_r is the effective relative permeability, and ε_{reff} is the effective relative permittivity.

From (1), by increasing the effective relative permittivity of dielectric material at a given operating frequency decelerate the propagation electromagnetic (EM) waves in a guided medium. The Slow Wave Effect can be realized by adding periodical shields that decelerate the propagation electromagnetic (EM) waves in a guided medium.

From (1), an equivalent inductance L_{eq}, equivalent capacitance C_{eq}, and propagation constant β can be described as

$$\beta = \omega\sqrt{L_{eq}C_{eq}} \Rightarrow V_P = \frac{1}{\sqrt{L_{eq}C_{eq}}} \tag{2}$$

From(2), both distributed inductance (L) and capacitance (C) along the transmission line in the guided medium can be increased for the realization of slow-wave effects owing to increase of the propagation constant β, consequently, the loss of the resonator can be reduced [8].

Fig. 1 shows the typical simplified topology of a single-coupled slow-wave resonator. This consists of a slow-wave resonator loaded at its near and far ends with series capacitors C_S. The resonator itself is made up of a transmission line of characteristic impedance Z_0 and electrical length 'θ', periodically loaded by shunt capacitors C_p. The electrical length θ of the unloaded transmission line is defined at frequency *f* as [10]

$$\theta = \beta d = \frac{2\pi f\sqrt{\varepsilon_{reff}}}{C} d \Rightarrow \quad \theta\big|_{f=f_0} = \theta_0 \tag{3}$$

where ε_{reff} is the effective relative permittivity , d is the physical length of the unloaded line, and C is the free space light velocity. At the center frequency f_0, the electrical length is θ_0. By using *ABCD* (cascade) matrices, the equivalent characteristic impedance Z_{eq} of a coupled slow-wave resonator can be easily extracted.

As shown in Fig. 1, the real estate of SWR layout is quite large, therefore not suitable for the application where size is the constraint. Fig. 2 shows the typical layout of the printed slow wave resonator, capacitively coupled between two arms of the resonator leads to the slow-wave effect. For unloaded transmission line there were no dispersion in the phase velocity, however, for the periodically loaded line the phase velocity is frequency dependant.

As illustrated in Fig. 3, the transmission line is periodically loaded with identical open stub elements for the realization of high Q-factor SWR structure. The input impedance Z_{in2} for a lossless line can be given by [9]

$$Z_{in2} = Z_0 \frac{Z_L + jZ_0\tan(\beta l)}{Z_0 + jZ_L\tan(\beta l)}, \quad where \quad Z_0 = \sqrt{\frac{L}{C}} \tag{4}$$

$$Z_{in2}\big|_{\tan(\beta l)\to 0} = \frac{Z_0 Z_L}{Z_0} \Rightarrow Z_{in2} \propto Z_L \tag{5}$$

where L and C are the inductance and capacitance per unit length of the line, Z_0 and β are the characteristic impedance and phase constant of the open stub, respectively.

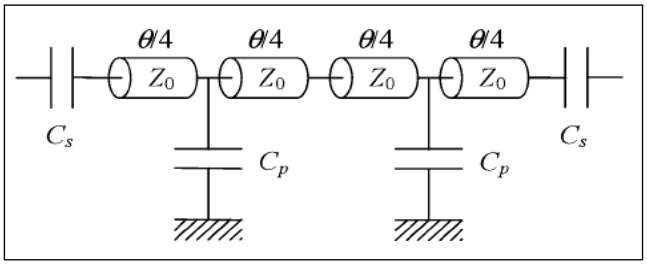


Fig. 1. A typical topology of a single-coupled slow-wave resonator loaded at the near and far ends with series capacitors C_s [10]

From (5), for infinitesimal value of $\tan(\beta l)$, the input impedance Z_{in2} is proportional to Z_L, therefore, $Z_{in2} \rightarrow \infty$ or 0 for corresponding $Z_L = \infty$ or 0. Under these cases, the slow-wave periodic structure loaded by Z_{in2} in Fig. 3(b) provides passband ($Z_{in2} \rightarrow \infty$) and stopband ($Z_{in2} \rightarrow 0$) characteristics.

As illustrated in Fig.3, the layout size is large and not attractive for tiny VCO applications. Fig. 4 shows the modified version of SWR, where transmission line is loaded by a square ring resonator with a line-to-ring coupling structure.

As depicted in Fig. 4(b), Z_{in3} is the input impedance looking into the transmission line l_b toward the ring resonator with the line-to-ring coupling.

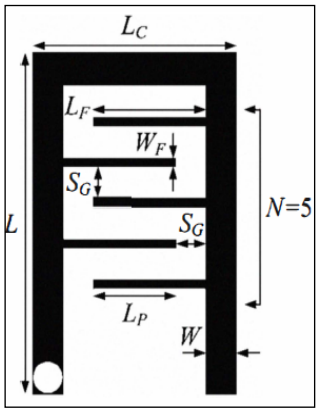


Fig.2 A typical slow wave printed coupled resonator (no of section, N=5) [8]

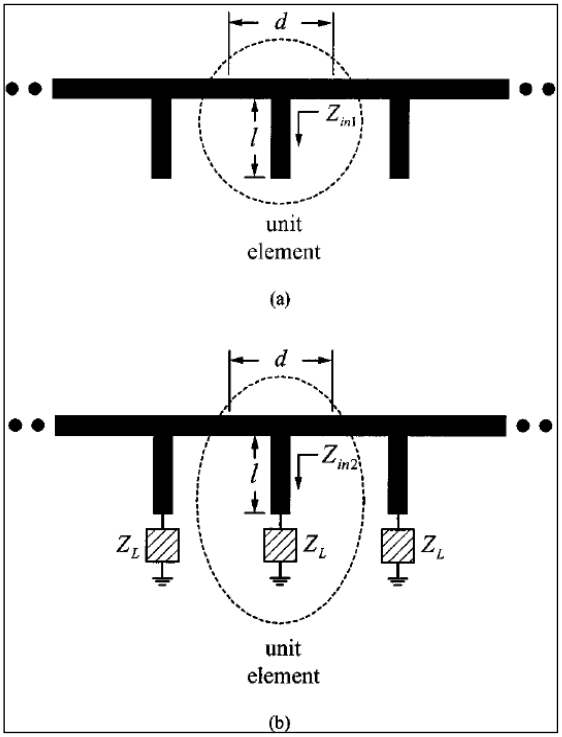


Fig. 3. Slow-wave periodic structure. (a) Conventional type. (b) With loading Z_L at open end [9].

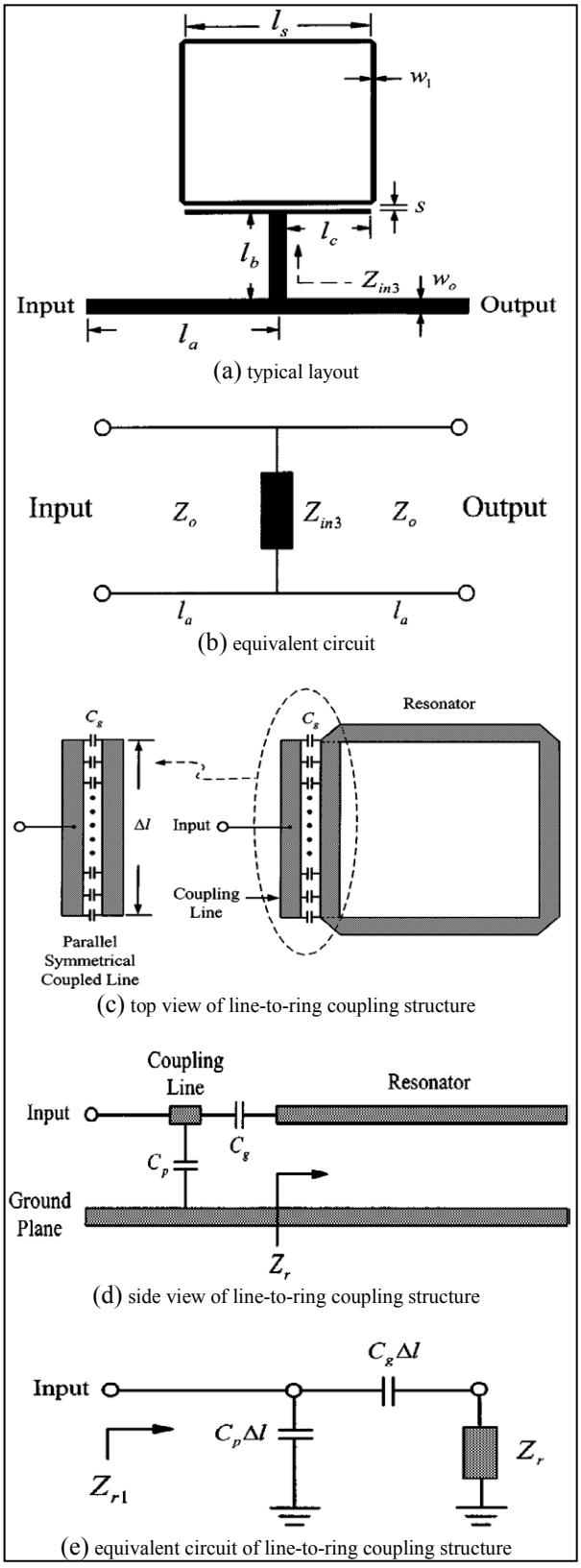


Fig. 4 Slow-wave ring resonator with one coupling gap (a) typical layout, (b) equivalent circuit, (c) top view of line-to-ring coupling structure, (d) side view of line-to-ring coupling structure, and (e) equivalent circuit of line-to-ring coupling structure [9].

The input impedance Z_{r1} looks into the line-to-ring coupling structure toward the ring resonator. The input impedance Z_{in3} is [9]

$$Z_{in3} = Z_0 \frac{Z_{r1} + jZ_0 \tan(\beta l_b)}{Z_0 + jZ_{r1} \tan(\beta l_b)} \quad (5)$$

Where

$$Z_{r1} = \frac{(Z_r + Z_g)Z_p}{(Z_r + Z_g + Z_p)}, \quad Z_g = \frac{1}{j\omega C_g \Delta l}, \quad Z_p = \frac{1}{j\omega C_p \Delta l}, \quad \omega = 2\pi f, \quad (6)$$

The parallel f_p and series f_s resonances of the slow wave ring resonator as shown in Fig.4 can be obtained by setting $|Y_{in3}| = 0$ and $|Z_{in3}| = 0$.

Figures 5(a), (b), and (c) show the typical block diagram and layout of the VCO using mode-coupled SWR, which validate a novel SWR (Slow Wave Resonator) approach, using a SiGe hetro-junction-bipolar-transistor (HBT) active device fabricated on low loss RF dielectric substrate material with a dielectric constant of 3.38 and thickness of 22 mils printed structure.

The SWR structure is modeled in 3-D EM (Electromagnetic) CAD simulator and incorporated into optimized nonlinear oscillator circuit to enable configurable and low phase noise operation over the band. This enables SWR structure to set up optimum standing waves (within the resonator) and the noise impedance transfer function over the tuning range by controlling m_{opt} (by optimizing injection locking) and β_{opt} (by optimizing mode tuning).

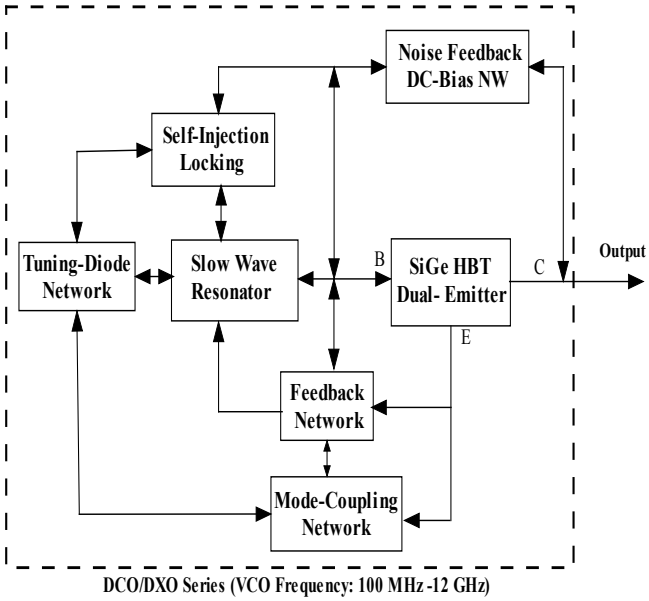
The nonlinear circuit contains the oscillator's active device, with S-parameters. This partitioning of the oscillator into its modeled component parts works quite well, and the combination of the S-parameters and the nonlinear circuit model agrees closely with the measured data in the circuits we already built. The S-parameters used for these transistors are large-signal S-parameters, which improve the optimization cycles using commercial CAD tools such as ADS 2010 (Agilent) and Ansys (Nexxim) to the limits allowed by physics.

B. High Q-Factor Slow-Wave Resonator VCO

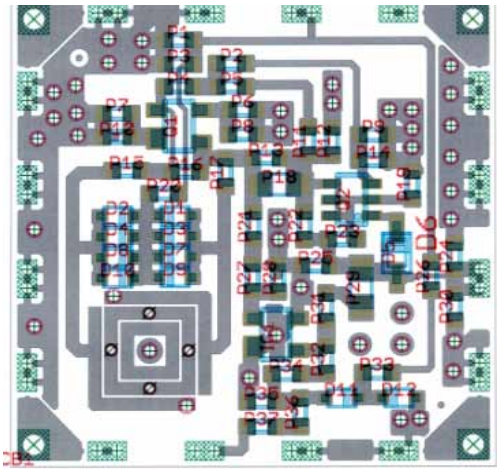
The *Q* (quality) factor of the Printed transmission line resonator can be enhanced by introducing slow wave coupling mechanism (electric/magnetic/hybrid).

Figure (6) illustrates the layout of the typical electric, magnetic, hybrid-coupling planar resonator networks, and oscillator circuits for comparative analysis. As described in Figure (6), the coupling dynamics can be characterized by proximity effect through the fringing fields, which exponentially decays outside the region; electric and magnetic field intensity tends to concentrate near the side having maximum field distribution. The coupling factor β (β_e: electric, β_m: magnetic and β_h: hybrid) can be described as

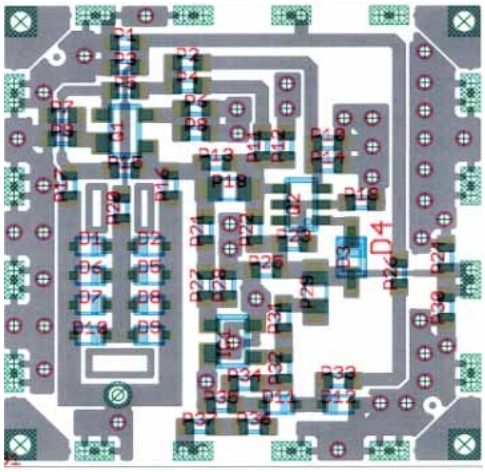
$$\beta_e \cong \frac{coupled-electrical \quad energy}{stored-energy \ of \ uncoupled-resonator} \cong \frac{f_{me}^2 - f_{ee}^2}{f_{me}^2 + f_{ee}^2} \cong \frac{C_{me}}{C} \quad (7)$$



(a) block diagram of SWR VCO



(b) layout mode-coupled SWR VCO



(c) progressive wave coupled SWR VCO

Fig.5. (a) A typical Block diagram of SWR VCO, (b) layout mode-coupled SWR VCO, and (d) layout of progressive wave coupled SWR VCO

$$\beta_m \cong \frac{coupled-magnetic\quad energy}{stored-energy\ of\ uncoupled-resonator} \cong \frac{f_{em}^2 - f_{mm}^2}{f_{em}^2 + f_{mm}^2} \cong \frac{L_{mm}}{L} \quad (8)$$

$$\beta_h \cong \frac{coupled-electro-magnetic\quad energy}{stored-energy\ of\ uncoupled-resonator} \cong \frac{f_{eh}^2 - f_{mh}^2}{f_{eh}^2 + f_{mh}^2} \cong \frac{CL_{mh} + LC_{mh}}{LC + L_{mh}C_{mh}} \quad (9)$$

where

$f_{ee} = \frac{1}{2\pi\sqrt{L(C+C_{me})}}$, $f_{me} = \frac{1}{2\pi\sqrt{L(C-C_{me})}}$, C_{me}: Mutual Capacitance

$f_{em} = \frac{1}{2\pi\sqrt{C(L-L_m)}}$, $f_{mm} = \frac{1}{2\pi\sqrt{C(L+L_m)}}$, L_{mm}: Mutual Inductance

$f_{eh} = \frac{1}{2\pi\sqrt{(L-L_{mh})(C-C_{mh})}}$, $f_{mh} = \frac{1}{2\pi\sqrt{(L+L_{mh})(C+C_{mh})}}$, L_{mh}: Hybrid Inductance

$f_0 = \frac{1}{2\pi\sqrt{LC}}$, f_0: Fundamental resonance frequency of uncoupled resonator.

The loaded quality factor Q_L of the coupled resonator network is given in terms of unloaded quality factor Q_o as

$$Q_L(\omega_0) = \frac{\omega_0}{2}\left[\frac{\partial\phi}{\partial\omega}\right] \quad (10)$$

$$[Q_L(\omega_0)]_{electrical-coupling} \cong 2\left[\frac{Q_0}{(1+\beta_e)}\right]_{\beta_e<<1} \cong 2Q_0 \quad (11)$$

$$[Q_L(\omega_0)]_{magnetical-coupling} \cong 2[Q_0(1+\beta_m)]_{\beta\to 1} \cong 2Q_0 \quad (12)$$

$$[Q_L(\omega_0)]_{hybrid-coupling} \cong 2\left[Q_0\frac{(1+\beta_{mh})}{(1+\beta_{eh})}\right]_{\beta_e<<1,\ \beta_m\to 1} \cong 2Q_0 \quad (13)$$

where $\frac{\partial\phi}{\partial\omega}$ is the rate of change of the phase, and Q_0 is the unloaded Q-factor of the uncoupled single open loop microstrip line resonator (Fig. 13). From (11), (12), and (13), there is trade-off between improving the Q factor and the permissible attenuation required (which is compensated by active device for oscillation build up). The coupling mechanism described in Figure (13) shows improvement in quality factor in comparison to single uncoupled planar resonator but drawback is limited tuning range (less than 1%). From (11), (12) and (13), loaded quality factor (Q_L) can be maximized by either lowering the value of mutual capacitance (C_m) and inductance (L_m) or maximizing the self-capacitance (C) and inductance (L), therefore, upper limit of the loaded Q- factor is dependent on the coupling β that can optimized by controlling the width of the transmission line (w), gap of the open line resonator (p), and spacing between the two open line resonators (d).

However, dynamic controlling and tuning the parameters w, p, and d at high frequency in IC technology is challenging task. For wideband tunability, the coupling factor β_j has to be dynamically tuned for low phase noise performances over the operating frequency band.

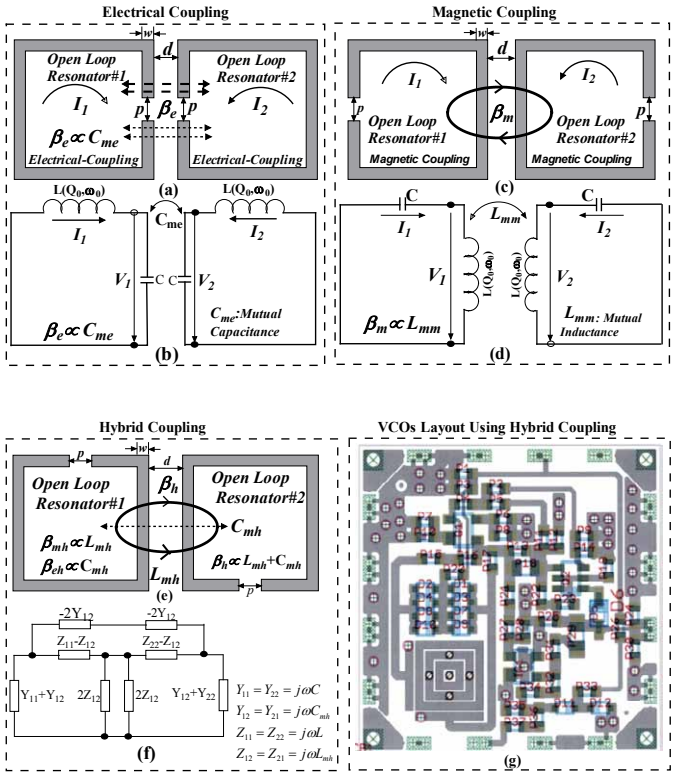


Fig.6. Typical simplified structure of open loop microstrip line coupled resonator networks:(a) Electrical coupling, (b) Equivalent lumped model of electrical coupling, (c) Magnetic coupling, (d) Equivalent lumped model of magnetic coupling, (e) Hybrid coupling, (f) Equivalent lumped model of hybrid coupling and (g) Layout of VCO using electric and magnetic coupling.

C. Phase Noise

The expression of phase noise can be given by [6, pp. 332]

$$\pounds(f_m) = 10\log\left\{\left[1+\frac{f_0^2}{(2f_mQ_L)^2}\right]\left(1+\frac{f_c}{f_m}\right)\frac{FkT}{2P_0}+\frac{2kTRK_0^2}{f_m^2}\right\} \quad (14)$$

$$\pounds(f_m) = 10\log\left\{\left[1+\frac{f_0^2}{(2f_mQ_0)^2m^2(1-m)^2}\right]\left(1+\frac{f_c}{f_m}\right)\frac{FkT}{2P_0}+\frac{2kTRK_0^2}{f_m^2}\right\} \quad (15)$$

where $\pounds(f_m)$, f_m, f_0, f_c, Q_L, Q_0, F, k, T, P_o, R, m and K_0 are the ratio of the sideband power in a 1Hz bandwidth at f_m to total power in dB, offset frequency, flicker corner frequency, loaded Q, unloaded Q, noise factor, Boltzman's constant, temperature in degree Kelvin, average output power, equivalent noise resistance of tuning diode, ratio of the loaded and unloaded quality factor and voltage gain. From (6), (7) and (8), m is given in terms of coupling coefficient as

$$m = \frac{Q_L}{Q_0} \cong \left[\frac{2}{(1+\beta_e)}\right]_{electrical} \cong [2(1+\beta_m)]_{magnetic} \cong \left[\frac{2(1+\beta_{mh})}{(1+\beta_{eh})}\right]_{hybrid} \quad (16)$$

By differentiating (16), with respect to m and equating to zero, local minimum value of phase noise for a given resonator and oscillator topology can be given by

$$\frac{\partial^2}{\partial m^2}[\pounds(f_m)]_{m=m_{opt}} = 0 \Rightarrow$$

$$\frac{d^2}{dm^2}\left[10\log\left\{\left[1+\frac{f_0^{\,2}}{(2f_mQ_0)^2m^2(1-m)^2}\right]\left(1+\frac{f_c}{f_m}\right)\frac{FkT}{2P_0}+\frac{2kTRK_0^2}{f_m^2}\right\}\right]=0\Rightarrow m_{opt}=0.5 \quad (17)$$

$$m_{opt}\to 0.5 \;\Rightarrow\; [\beta_e]_{opt}\ll 1,\quad [\beta_m]_{opt}\to 1,\quad 0<[\beta_h]_{opt}<1 \quad (18)$$

From (18), for low phase noise applications, m_{opt} and β_{opt} should be dynamically tuned and must converge in the vicinity of $m_{opt}\cong 0.5$ and $0<\beta_{opt}<1$ respectively for best phase noise performances.

From (18), broadband tunability and low phase noise performance is opposing requirements due to the problem of controlling the loop parameters, and the optimization of the time average loaded Q of the resonator over the band simultaneously. The novel mode-coupled SWR network allows for a substantial reduction in phase noise by dynamically optimizing impedance transfer function and coupling factor across the progressive medium [6].

The new approach to designing tunable oscillators with mode-coupled SWR yields compact VCOs that tune either optimized or wide bandwidths with excellent phase-noise performance, rivaling performance from high Q resonator oscillators with excellent phase noise performance, and in configurations that can be readily adapted to modern RF integrated circuit (RFIC) and MMIC semiconductor manufacturing processes. As an example for validation of this approach, Figures (7), (8), (9), and (10) shows the phase noise, tuning range, RF out power, and harmonics characteristics for tiny SWR VCO (4200-9400 MHz) shown in Fig. 6(g), fabricated on Rogers substrate material with a dielectric constant of 3.38 and thickness of 22 mils (PCB size: 0.3 x 0.3 x 0. 08 inches).

As shown in Figure 7, SWR VCO offers typical measured phase noise –105dBc/Hz at 100 kHz offset from the carrier frequency of 9 GHz with 3V, 25mA DC bias (power consumption: 75milli Watt). The reported novel SWR VCOs offer more than 2:1 tuning bandwidths (Fig. 8).

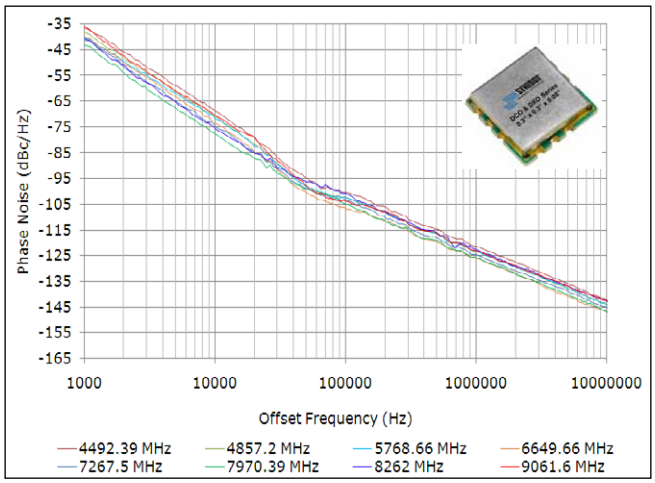


Fig. 7. This plot shows the measured phase noise of the SWR VCO tunes from 4500 to 9000 MHz for coupling dynamics $m_{opt}\cong 0.5$ and $\beta_{opt}\cong 1$.

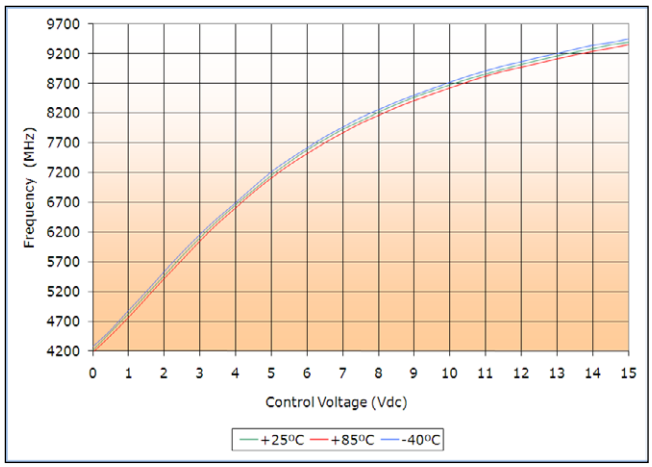


Fig.8. This plot shows the measured tuning characteristics of the SWR VCO tunes from 4500 to 9000 MHz for coupling factor $\beta_{opt}\cong 1$.

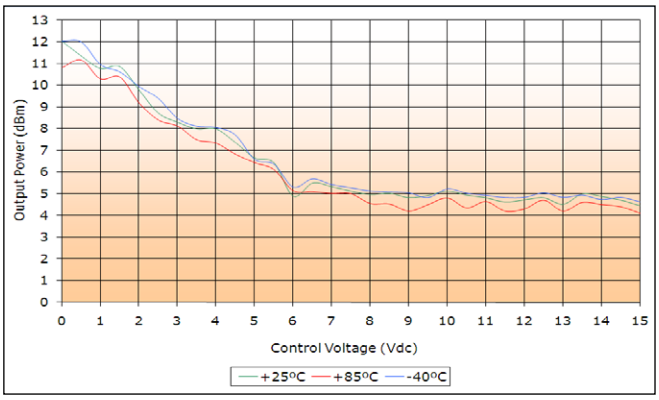


Fig.9 This plot shows the measured output power characteristics of the SWR VCO tunes from 4500 to 9000 MHz for coupling factor $\beta_{opt}\cong 1$.

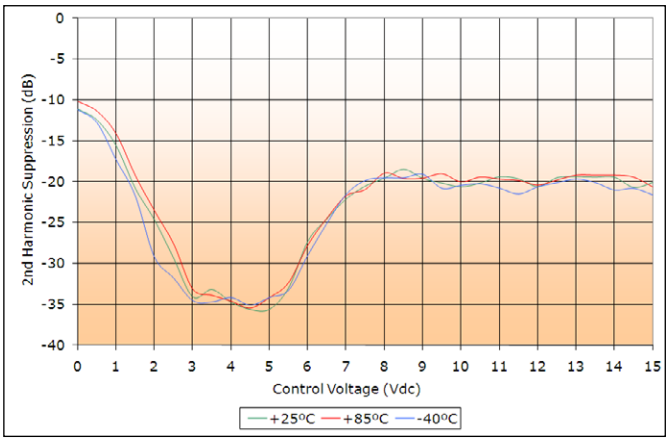


Fig. 10. This plot shows the measured second harmonics characteristics of the SWR VCO tunes from 4500 to 9000 MHz for coupling factor $\beta_{opt}\cong 1$.

These bandwidths support a wide variety of portable and fixed communications applications.

The new SWR VCOs can also be designed into compact housings with low phase noise and optimized bandwidths to challenge the performance of high-Q ceramic-resonator, surface-acoustic-wave (SAW) resonator, Dielectric resonator

and YIG resonator oscillators as well as low-Q resonators based on lumped circuit elements and printed transmission lines but at about one-third the size.

The investigation perform includes the study of the frequency reference sources for minimum jitter caused due to mechanical and thermal stress, thereby less prone to phase hits over the tuning range. The additional feature of this topology is user defined frequency band by adjusting the length/spacing and optimizing the phase velocity and group delay of the SWR networks.

III. Ultra Low Phase Noise SWR VCO

The scope of this paper is to design ultra low phase noise reference SWR VCO for PLL applications that supports the fast convergence by using dynamically tuned SWR network for maximum group delay, thereby, improved dynamic loaded Q, and hence low noise performance over the band.

Fig. 11 shows the block diagram of novel SWR oscillator topology, capable of operating over a wide tuning (6000-12000 MHz) and temperature range (–40°C to +85°C) with minimal frequency drift (less than 10 MHz) due to the temperature. As shown in Fig.11, multi-mode coupled-SWR network is connected across the base and collector of the three terminal active device (bipolar transistor), can be characterized as high Q mode coupling resonator that eventually improves the time average loaded Q over the multi-octave-band operation. As illustrated in Fig.11, the active impedance created by the 3-terminal active device has a negative real part with a real magnitude, and an imaginary part with an imaginary magnitude. The real magnitude is a function of the imaginary magnitude and the imaginary magnitude is selected such that the real magnitude compensates the loss of the SWR as well coincides with a maximum-slope inflection point of the phase characteristic curve for improved group delay. In addition to this, the mode coupling approach includes a methodology for optimum dynamic coupling to enhance the dynamic loaded Q, and to reduce or eliminate phase hits, while reducing the susceptibility to microphonics to an extremely low level, and retaining low phase noise and broadband tunability.

In addition to this, the noise feedback and DC bias circuit is coupled across the base and collector terminals of the active device. A multi-mode coupled resonator network as shown in Fig.11 is capacitively coupled across the base and collector terminals. A slow wave coupled resonator and a progressive wave-coupled resonator described in Fig.11 are coupled in distributed domain at the collector terminal. A phase compensating network is capacitively coupled between the base terminal and the slow-wave and progressive-wave coupled resonators. An RF output signal is coupled through a distributed coupled medium, which is coupled across the slow wave and progressive wave resonator networks.

As shown in the Fig. 12, the lay of the coupled SWR network is symmetrically structured to minimize the stray reactance that would otherwise limit or reduce the circuit tuning range or bandwidth.

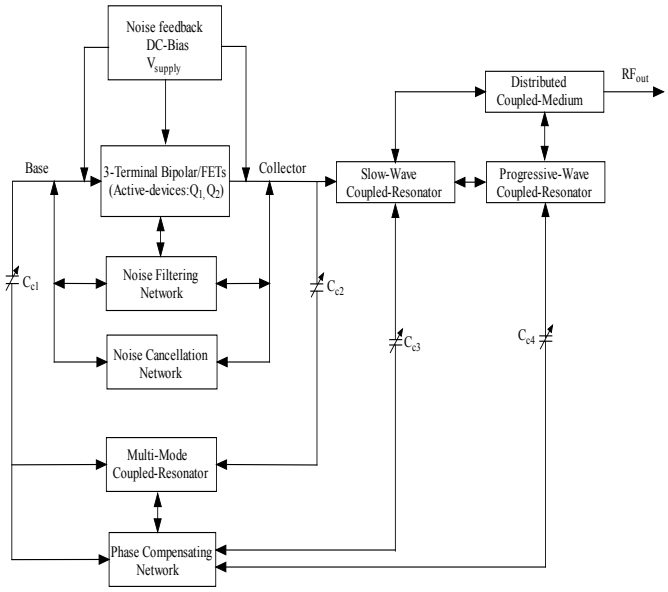


Fig. 11: Block diagram of configurable SWR VCO (6000 MHz-12000 MHz)

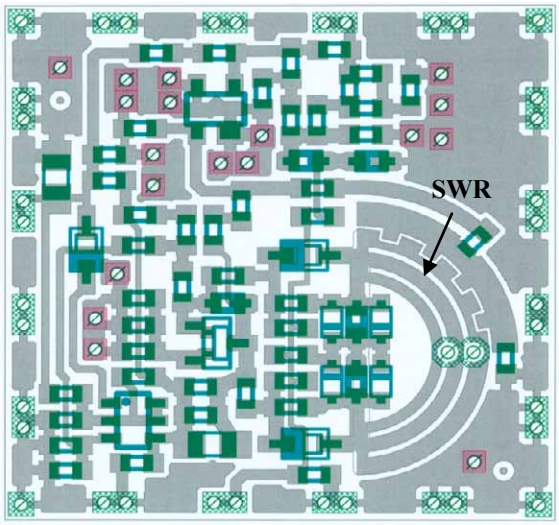


Fig. 12: Layout of the SWR VCO (6-12 GHz) with 5 Volt, 30 mA

An objective of this paper was to improve the Q value of the SWR by cascading a progressive wave coupled resonating network having multiple asymmetric coupled line to increase the phase derivative and group delay [$\delta\varphi/\delta\omega$] of the resonator in a given cycle. Therefore, VCO circuit is designed such that the phase condition occurs in the steepest part of the phase characteristics curve of the SWR network. In this way, coupled mode resonance may substantially coincide with a maximum-slope of the phase characteristic curve of the oscillator circuit, and the operation of the oscillator circuit about this point results in the circuit operating within the lowest phase noise domain of the phase characteristic curve, thereby improved effective dynamic loaded Q.

The layout shown in Fig. 12 is fabricated on 22 mil thick PCB of dielectric constant 3.38. The choice of substrate depends on size, higher-order modes, surface wave effects, implementations (couplings, line length, width, spacing, and spacing tolerances), dielectric loss, temperature stability, and power handling (dielectric strength and thermal conductivity). In many applications, keeping resonator structures to a minimum size and weight is very important. It would seem that planar resonator structures which can be fabricated using printed-circuit technologies would be preferred whenever they are available and are suitable because of smaller sizes and lighter weight. In this paper, we introduce a new class of planar resonators based on coupled slow-wave structure, depicted like saw-tooth/gear structure, which is not only compact size but also have a wider tuning range. Furthermore, SWR is loaded with variable capacitor for uniform loaded Q over the tuning range. As we know that if there were no dispersion, the phase velocity would be constant but this is true only for the unloaded line. However, for the periodically loaded line the phase velocity is frequency dependent, therefore, there is trade-off between constant phase velocity and dispersion.

Fig. 13 illustrates the measured Q of the uncoupled, coupled and MCSW resonator. The quality factor (Q) is given by

$$Q = \frac{\omega_0}{2}\left[\frac{\partial\varphi}{\partial\omega}\right]; \quad \omega_0 = 2\pi f_0 \tag{19}$$

where f_0 and φ are the resonant frequency and phase of the oscillator signal.

The dynamic time average Q factor of the resonator, as well as the tuning diode noise contribution, sets the noise performance of the VCO, and in general, the dynamic loaded Q is inversely proportional to the frequency range of the VCO. The phase noise can vary dramatically over the tuning range of a multi-octave-band VCOs; therefore it is a major challenge to explore and find ways to realize uniform loaded Q over the tuning range.

As shown in the Fig. 13, MCSWR (Multi-coupled slow wave resonator) network acts like a Q-multiplier over the 6-12 GHz, peaking at 9 GHz due to evanescent mode- coupling. The variable coupling capacitors C_c (C_{c1}, C_{c2}, C_{c3}, C_{c4}), as shown in Fig. 11 are designed for the optimum loading of the coupled resonator network across the active device, and dynamically tuned for optimum loaded Q ($m_{opt} \cong 0.5$) for minimum phase noise performance across the band. The values of the coupling capacitor C_c are derived from the input stability circle, and it should be within the input stability circle so that the circuit will oscillate at a particular frequency for the given lowest possible value of the C_c over the band.

The noise factor F of the oscillator is minimized by noise feedback circuit [2] and dynamically controlling the conduction angle of the active device by adjusting optimum drive level corresponding to the desire oscillation frequency. By doing so, the control parameters such as negative resistance, loop gain, and phase shift can be adjusted for improving the phase noise performance over the band.

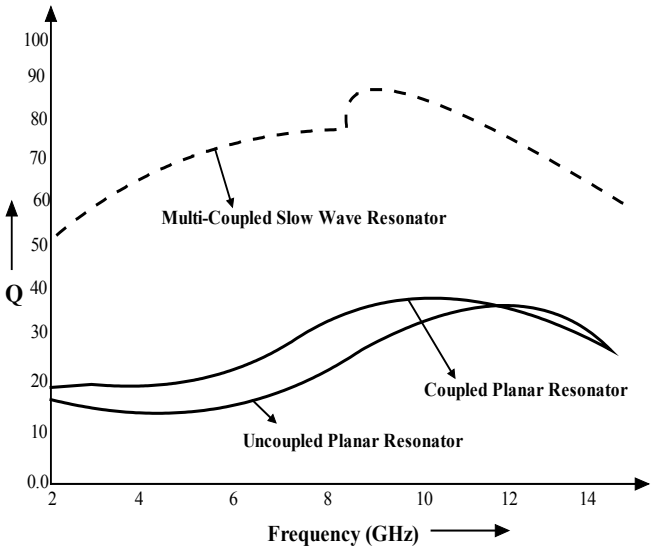


Fig. 13: Measured quality factor (Q) of the slow wave resonator

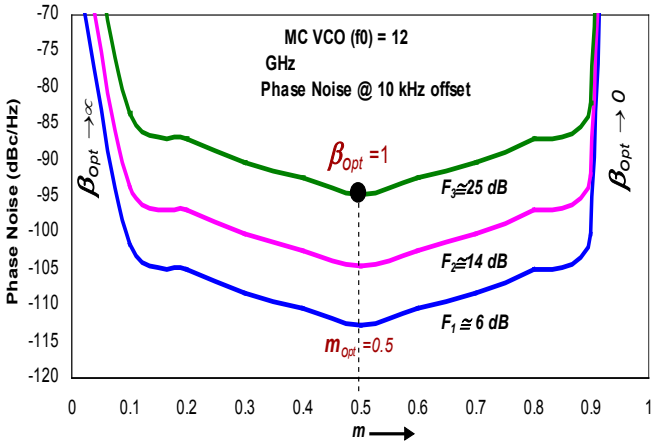


Fig. 14. Simulated phase noise plot of MCSWR VCO with respect to m_{opt},

Figure (14) shows the CAD simulated phase noise plot for a 12 GHz MCSWR VCO (Fig. 13) at offset 10 kHz from the carrier with respect to m_{opt} and β_{opt}.

Fig. 15 shows the typical measured phase noise plots at for 12 GHz carrier frequency. The typical measured phase noise @ 10 kHz offset from the 12 GHz carrier is -110 dBc/ Hz, which agrees within 3-4 dB of the CAD simulated results and to author's knowledge, the reported resonator topology shows first oscillator using slow wave dynamics in multi-mode regimes.

The theoretical treatment and experimental validation of a new class of slow wave resonator suggests the possibility of cost-effective alternative of high Q dielectric resonator based signal sources. Other features and advantages of the present technology include power efficiency, low phase noise, and amenability to integration in chip form.

IV. Synthesizer Using Multi-Band Tunable SWR VCO

Technologies for creating frequency synthesizers are diverse, from traditional analog methods using PLLs to direct digital synthesizers (DDS) that rely on high-speed digital–to-

analog converters (DACs) to transform digital input words into analog output signals. Frequency synthesizers can be categorized into mainly three groups: analogue, digital or mixed signal (hybrid). The frequency synthesizer described in this Figure (16) falls into the hybrid category [7]. As shown in Figure (16), the block diagram for a typical synthesizer includes a SWR VCO, phase-lock-loop (PLL) IC, charge pump, loop filter, amplifier, and voltage regulator. A typical unit tunes from 4000-8000 MHz in 1-MHz steps with settling time of less than 4 ms and typical phase noise of -104 dBc/Hz at 100 kHz offset from the carrier. Figure (17) shows the measured phase noise plot for miniaturized wideband VCO (0.2-1GHz/1-4GHz/4-8GHz/8-12 GHz) with limited power supply (3.1V, 10mA) that enables power-efficient frequency synthesizers (Figure 16) in a SMD package measuring just 0.6 x 0.6 inches.

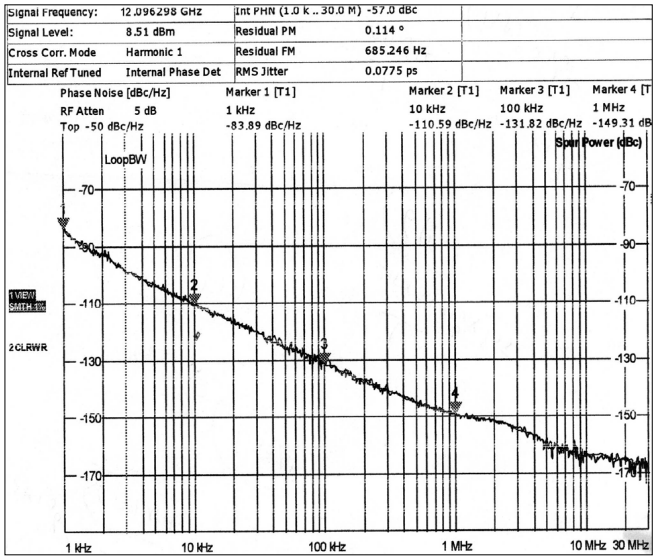


Fig.15 A typical measured phase noise plot of 12 GHz MCSWR VCO

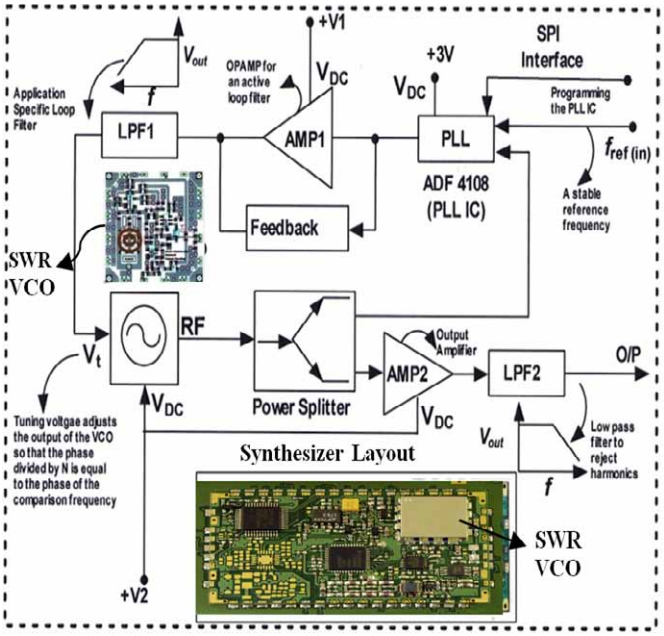


Fig.16. Block diagram of configurable synthesizer using Tunable SWR VCO

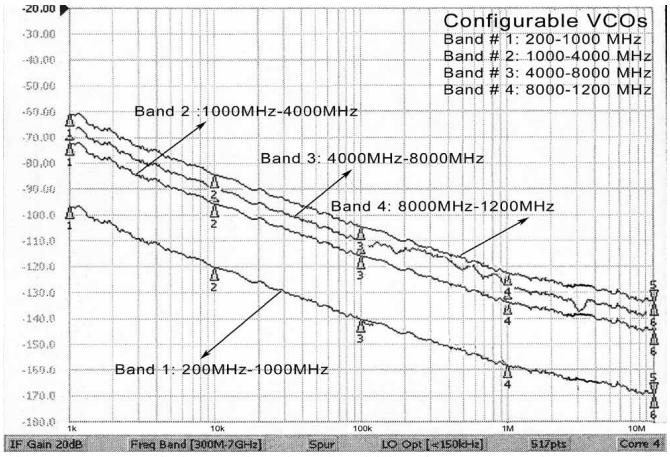


Fig. 17. Measured phase noise plot of multi-band SWR VCO for $m_{opt} \cong 0.50$

Figures (18) and (19) show the measured data of phase noise for 4000-8000 MHz configurable synthesizer. The synthesizer draws typically 30 mA current from a 5-V supply and can operate as low as 100 MHz and can be extended to 12 GHz using multi-band/multi-mode VCOs resulting low cost and power-efficient configurable synthesizer in 0.6x0.6 inches with excellent performances in terms of cost and phase noise.

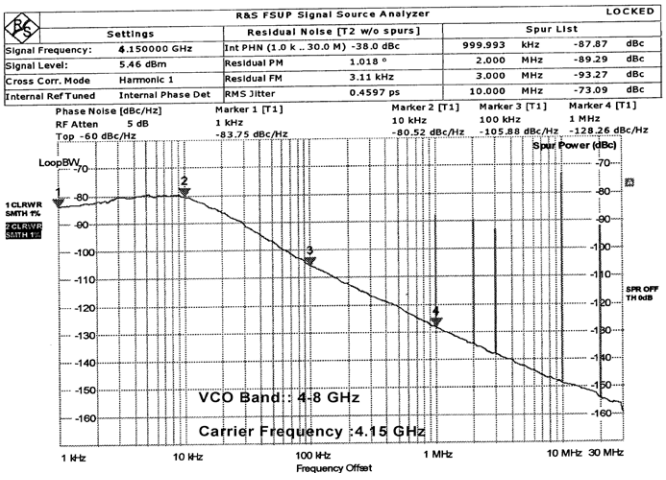


Fig.18. Measured phase noise plot at 4.15 GHz of synthesizer (4-8 GHz)

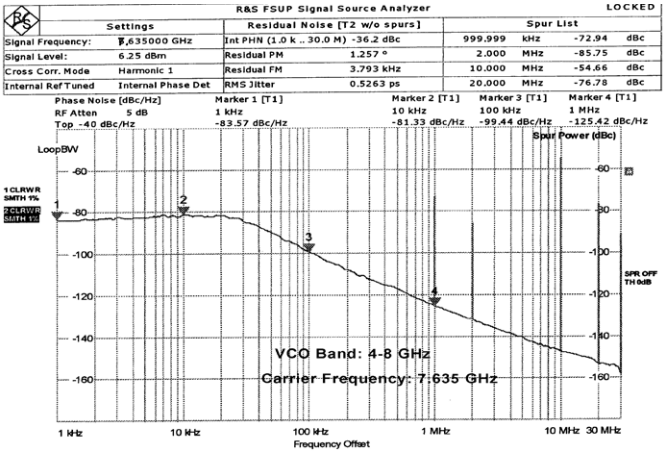


Fig.19. Measured phase noise plot at 7.635 GHz of synthesizer (4-8 GHz)

V. CONCLUSION

We have reported details of multi-band SWR VCO providing concurrent and seamless tuning over a wide tuning range for feeding tiny configurable synthesizer circuit. The freedom of selection of the frequency, low phase noise, low phase hits, low power consumption, compact size, and stable over temperature will make SWR VCO technology promising and attractive for next generation high frequency mobile communication systems and test and measuring instruments.

REFERENCES

[1] Jia-Sheng Hong and M. J. Lancaster," Theory and Experiment of Novel Microstrip Slow-Wave Open-Loop Resonator Filters", IEEE Trans. on MTT, Vol. 45, No. 12, Dec 1997.

[2] US Patent No. 7,605,670 B2, October 20, 2009, User-Definable Low Cost Low noise and Phase Hit Insensitive Multi-Octave-Band Tunable Oscillator, Ajay K. Poddar, Klaus J. Schoepf, Parimal Patel, Ulrich L. Rohde.

[3] A. K. Poddar, J. K. Bansal,K. N. Pandey, "Millimeter Wave Evanescent Mode Power Combiner-Gunn Oscillator in Suspended Stripline Configuration", IEEE MMET'98 pp. 384-386, June 1998.

[4] Z. A. Shaik, P.N. Shastry, and S.N Prasad," A Novel Distributed Voltage-Controlled Oscillator for Wireless Systems", *IEEE RWS,* San Diego, California, USA, 14 -20, January 2006.

[5] R. Leier, "SiGe Silences YIG Oscillatotr Phase Noise", Microwave & RF, pp. 79-82, January 2006.

[6] U. L. Rohde, A. K. Poddar, and G. Boeck, Modern Microwave Oscillators for Wireless Applications: Theory and Optimization, John Wiley & Sons Inc., 2005.

[7] U. L. Rohde, and A. K. Poddar, "Cost-Effective, Power-Efficient and Configurable YIG Replacement Signal Source," German Microwave Conference-GeMiC 2006, 28-29 March 2006, Germany.

[8] Wei-Shin Chang, Cheng-Hsien Liang, and Chi-Yang Chang," Compact Microstrip Bandpass Filters Using Miniaturized Slow-Wave Quarter-Wavelength Resonators", Proceedings of Asia-Pacific Microwave Conference 2010, pp. 1958-1961

[9] Lung-Hwa Hsieh and Kai Chang, "Slow-Wave Bandpass Filters Using Ring or Stepped-Impedance Hairpin Resonators," *IEEE Transactions on Microwave Theory & Techniques*, Vol. 50, No. 7, July 2002, pp. 1795-1800.

[10] E. Pistono, M. Robert, L. Duvillaret, J.-M. Duchamp, A. Vilcot, and P. Ferrari, "Compact Fixed and Tune-All Bandpass Filters Based on Coupled Slow-Wave Resonators," *IEEE Transactions on Microwave Theory & Techniques,* Vol. 54, No. 6, June 2006, pp. 2790-2798.

[11] A. P. S. (Paul) Khanna, "Microwave Oscillators: The State of The Thechnology," *Microwave Journal,* April 2006, pp. 22-42.

[12] J. Everard and K. Theodoropoulos, "Ultra-Low Phase Noise Ceramic Based Dielectric Resonator Oscillators," Proceedings of the IEEE International Frequency Control Symposium, June 4-7, 2006, pp. 869-874.

[13] V. Walkar and I. C. Hunter, "Design of triple mode TE_{01} resonator transmission filters," *IEEE Microwave and Wireless Components Letters*, Vol. 12, June 2002, pp. 215-217.

[14] U. L. Rohde and A. K. Poddar, "Low Cost Signal Source for Multi-Band Multi-Mode Wireless Systems," *Microwave Journal,* July 2007.

[15] J. S. Kim, W. Wu, J. Lin, A. Verma, S. Jang, F. Ren, S. Pearton, and J. Gillespie, "A High-Efficiency GaN/AlGaN HEMT Oscillator Operating at L-Band," Proceedings of the 2006 Asia Pacific Microwave Conference, December 12-15, 2006, Yokohama, Japan.

[16] S. Hamano, K. Kawakami, and T. Takagi, "A Low Phase Noise 19 GHz-band VCO using Two Different Frequency Resonators," 2003 IEEE MTT-S Symposium Digest, pp. 2189-2192.

[17] A. Grebennikov, *RF and Microwave Transistor Oscillator Design*, John Wiley & Sons Ltd., 2007, West Sussex, UK, 2007.

[18] J. K. A. Everard and C. Broomfield, "Reduced Transposed Flicker Noise in Microwave Oscillators using GaAs based Feedforward Amplifiers," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control,* Vol. 54, No. 6, June 2007, pp. 1108–1117.

[19] J.K.A. Everard and L. Zhou, "Non Linear Effects in varactor tuned resonators," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, Vol. 53, No. 5, May 2006, pp. 853–861.

[20] P.A. Dallas and J.K.A. Everard, "Characterization of flicker noise in GaAs MESFETs for oscillator applications," *IEEE Transactions on Microwave Theory and Techniques,* Vol. 48, No. 2, February 2000, pp. 245-257.

[21 Hsiu-Ying Cho, Tzu-Jin Yeh, Sally Liu, and Chung-Yu Wu, "High-Performance Slow-Wave Transmission Lines With Optimized Slot-Type Floating Shields," *IEEE Transactions on Electron Devices*, Vol. 56, No. 8, August 2009.

[22] T. Masuda, N. Shiramizu, T. Nakamura, and K. Washio," Characterization and Modeling of Microstrip Transmission Lines with Slow-wave Effect," Silicon Monolithic Integrated Circuits in RF Systems, 2008, pp.55-158.

Progress in Rubidium Fountain Clock Research at SIOM

Chun-Yan Shi, Rong Wei, Zi-Chao Zhou, Tang Li, Lin Li and Yu-Zhu Wang
Key Laboratory of Quantum Optics
Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences
Shanghai, China
Email:2003274026@163.com

***Abstract*—A laser-cooled rubidium fountain clock is studied at Shanghai Institute of Optics and Fine Mechanics (SIOM). One important characteristic of the setup was, instead of six laser beams used in normal fountains, a fold optical path (FOP) was used to trap, cool, and launch atoms, which decreased laser power by 60% and the optical system was simplified greatly. This fountain clock had been operated with the center Ramsey fringe of line-width of 1Hz and the signal noise ratio (SNR) of 100. Recently, some evaluations of the fountain clock systematic uncertainty have been finished. By comparing with an H-maser (VCH-1003A), the evaluation of the frequency instability was obtained by $8\times10^{-13}\tau^{-1/2}$ and 6×10^{-15} in one day. The system errors (type-B uncertainty) had been evaluating, and some results, such as second-order Zeeman shift and blackbody radiation shift, had been achieved. Some experiments of ^{85}Rb fountain were also reported in this paper.**

I. INTRODUCTION

As a combination of the separated oscillatory field [1] method and the laser cooling techniques[2], the atomic fountain clock (AFC)s have been developed to be the most accurate running clocks in the world[3]. The combined standard fractional uncertainties of the AFCs[4,5,6,7] have reached to 10^{-16}. The first laser-cooled rubidium fountain clock at SIOM was constructed in 2005, and a center Ramsey fringe with the line-width of 1Hz and the signal noise ratio (SNR) of more than 100 were obtained. Since 2009, the evaluating works of the frequency stability and the fractional uncertainties were started and some improvements for the fountain were advanced.

II. FOUNTAIN SETUP

Schematic of the fountain device was illustrated in Fig. 1, detailed description can be found in [8]. Atoms were trapped and cooled by magneto-optical trap (MOT) and optical molasses (OM), and then were launched by optical moving molasses (OMM) in cooling zone, at the bottom of the device, with its pressure under 2×10^{-7} Pa. The cylindrical state-selection cavity was located above the cooling zone, with the mode of TE011 and its diameter and height of 6.92cm and 3.46cm respectively. The detection zone above state-selection cavity included two parts to detect atoms at state F=2 and F=1, respectively. The Ramsey cavity, the same with the state-selection cavity in cavity mode, diameter and height, was located 59.7cm above the cooling center and inside the magnetic shields made of miu-metal (three internal layers) and soft iron (the external layer), the quality factor was 18 000, the temperature of Ramsey interaction zone was kept at 52℃ to ensure the resonance frequency of the cavity.

The optical system of the fountain was shown in Fig. 2. A diode laser (TOPTICA TA100) was treated as a master, whose frequency was locked to the saturated absorption crossover optical resonance of F=2→F'=3 and F=2→F'=2 transitions, exported 150mW laser power to supply the whole fountain. The laser was divided into three parts known as detecting beams, cooling beams and repumping beams. The detecting beams and cooling beams were obtained by some of optical components (wave plates and polarization beam splitters) and four acousto-optic modulator (AOM)s to tune the frequency and power of the beams. The repumping beams were produced by a method of modulating a fiber electro-optic modulator (FEOM) with 6.7GHz modulating frequency. There

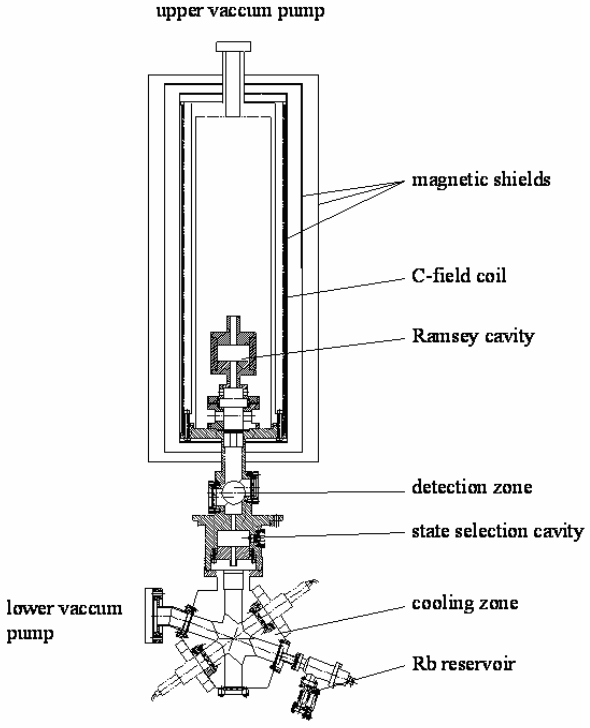


Figure 1. Schematic of the Rb fountain clock at SIOM

were two characteristics at our optical setup: the first one was a fold optical path (FOP) structure used to cool and launch atoms. As shown in Fig.3, two input beams were reflected six times by six mirrors to cool atoms instead of conventional six-beam launching structure in normal (1, 1, 1) fountains. As discussing in [9], FOP could work as a normal six-beam fountain and fit all requirements of fountain, and compared with a normal fountain, which reduced 2/3 of the laser power and decreased a lot of parts for optics and fiber, it was less power dissipation, more stable, compact, and easier to manipulate. The results demonstrated in the experiment showed that the trapping number (~10^8) and temperature (2.5μK) of cooling atoms in FOP were approximate with that in normal six-beam-fountain. The other characteristic was that repumpping beams created by an FEOM instead of a laser enhanced the stability of the optical system as FEOM was more compact, stable and reliable than a laser.

The microwave system of the fountain was consisted of a local oscillator and a microwave synthesis chain. The work shown below was base on a device synthesizer (Agilent E8257C) referring to H-maser or oscillator (OSA, 8607), whose phase noise affected the SNR of Ramsey fringe and the stability of fountain. And now, with the help of Giorgio Santarelli at SYRTE, a new synthesizer was researched, the residual phase noise of two synthesis chains was measured at frequency 7.034GHz and was shown in Fig.4. In the pulsed

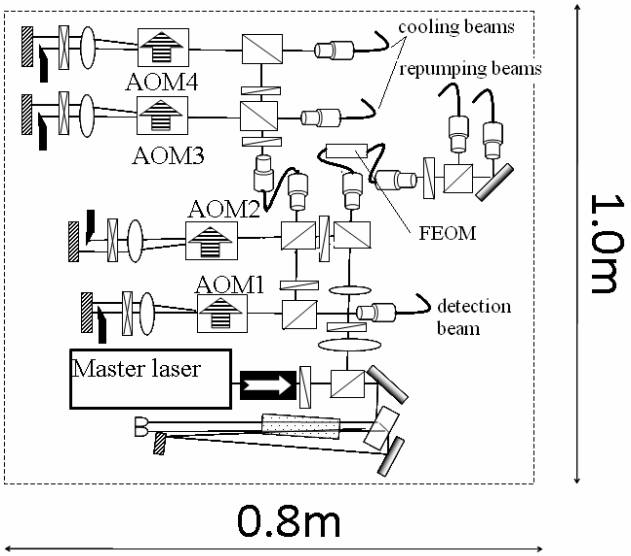


Figure 2. Schematic of the optical setup of the fountain.

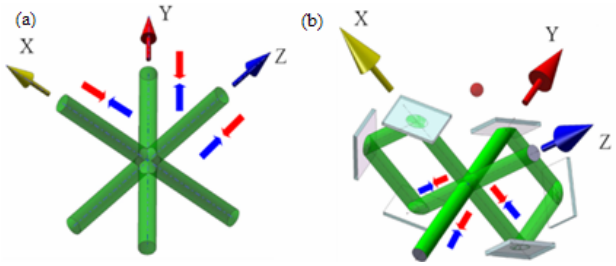


Figure 3. Schematic of input beams in cooling zone of the fountain device. The launching atoms way was (111) in the cooling zone. (a) Normal 6-beam-fountain; (b) FOP-based fountain

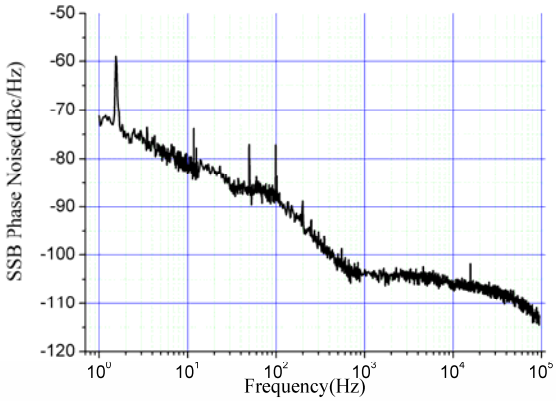


Figure 4. Residual phase noise of two chains at frequency 7.034 GHz.

operating fountain, the phase noise at Fourier frequencies, close to multiples of $1/T_C$, was down-converted, which leaded to degradation of frequency stability[10] named Dick effect[11], the data in Fig.4 means that the contribution of synthesizer to the short-term stability would be less than $1\times10^{-13}\tau^{-1/2}$.

III. PERFORMANCE OF THE FOUNTAIN

The device cycled with a normal fountain procedure, 'trapping and launching – state selecting – interacting - detecting', as description in [8], the falling atoms with launching velocity of 4.188m/s was measured by time of flight method at 10^6 level with the temperature at 2.5μK at detection zone, and the fountain cycle was 2s. The acquired Ramsey fringe was shown in Fig. 5, the interaction time between the atoms and microwave field was 10ms, the free evolution time was 0.5s, the line-width of the Ramsey central fringe was 1Hz, the SNR is more than 100, and the short-term stability was estimated to be $8\times10^{-13}\tau^{-1/2}$ based on SNR.

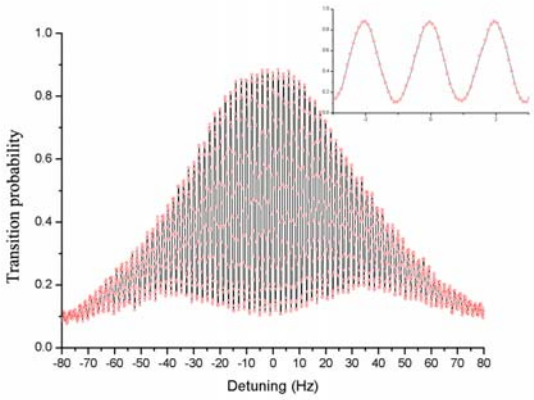


Figure 5. Ramsey interference fringe of ^{87}Rb fountain at SIOM.

The frequency instability (statistical error) of the fountain was measured by comparing the local oscillator (5MHz, BVA-8607) locked to the fountain with H-maser (VCH-1003A) using a commercial phase comparator, the Allan standard deviation was shown in Fig. 6. The short-term stability of the fountain was limited by the Dick effect from previous microwave synthesis chain. The compared result was believable after 10s as every feedback signal to local oscillator had a time delay about two fountain cycles (4s). Another comparing method was used by taking the H-maser as the reference of the commercial synthesis (Agilent E8257C) and the feedback signal was sent to the synthesis by using GPIB-488 card communication, the Allan standard deviation was obtained from the error signal shown in Fig. 7, the evaluation of the frequency instability was obtained by $9\times10^{-13}\tau^{-1/2}$ and 6×10^{-15} in one day. The comparing time was about 10 days.

A series of interested experiments correlative with laser-cooled ^{85}Rb atomic fountain were performed in our fountain device[12]. ^{85}Rb atoms were trapped and cooled by MOT and OM and were then launched by OMM when changing the lock-in peak of the saturated absorption of laser and adjusting the modulating frequencies of AOMs and FEOM, the cooling temperature was about 1.5μK and the detected atomic number in the detection zone was at 10^7 level. In our fountain device, the cavities were designed for ^{87}Rb with the resonance frequency of 6.835GHz which was far away from the clock frequency of ^{85}Rb at 3.036GHz, so the microwave field for ^{85}Rb was built by antenna-coupling, an antenna was located above the top window of the fountain, emitting two microwave pulses for Ramsey interaction. The Ramsey fringe, averaged by 7 times, was obtained and shown in Fig. 8 when the launching velocity was 4.188m/s, and the line-width was 2.1Hz. The frequency stability of ^{85}Rb atomic fountain was evaluated, and the Allan standard deviation was shown in Fig. 9 with stability of $4\times10^{-12}\tau^{-1/2}$ and 5×10^{-14} in 8000s. A +5.8Hz deviation from 3 035 732 439.0(60) Hz was measured in the experiments, however, as space microwave field brought many large uncertainties, the accuracy was bad. A special fountain device for ^{85}Rb is necessary for measuring the clock frequency of ^{85}Rb with uncertainty at 10^{-15} level, which has a potential application on measuring the variation of the fine structure constant[13,14].

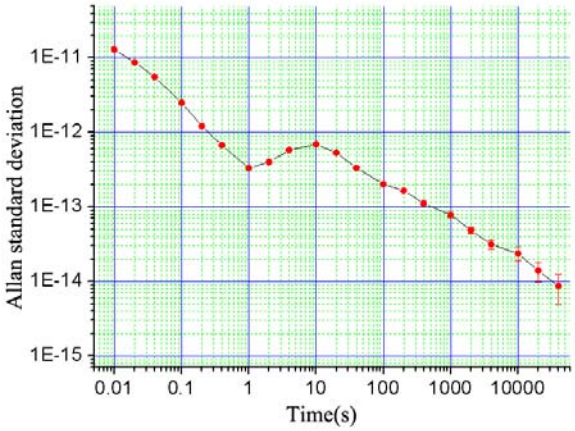


Figure 6. Allan standard deviation of the fountain measured against the

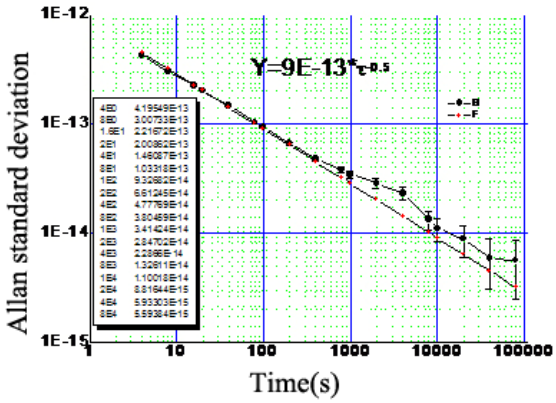


Figure 7. Allan standard deviation of the fountain measured against the compare between the locked synthesis and the H-maser (VCH-1003A).

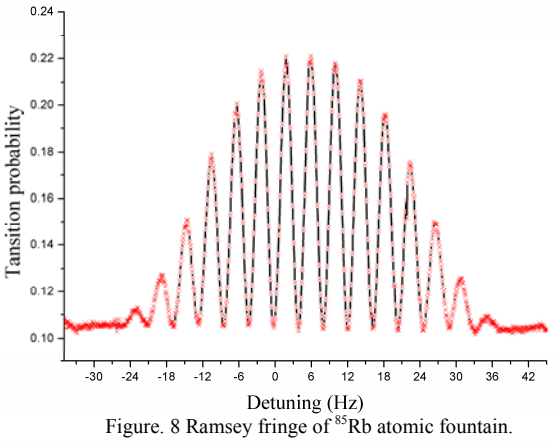


Figure. 8 Ramsey fringe of ^{85}Rb atomic fountain.

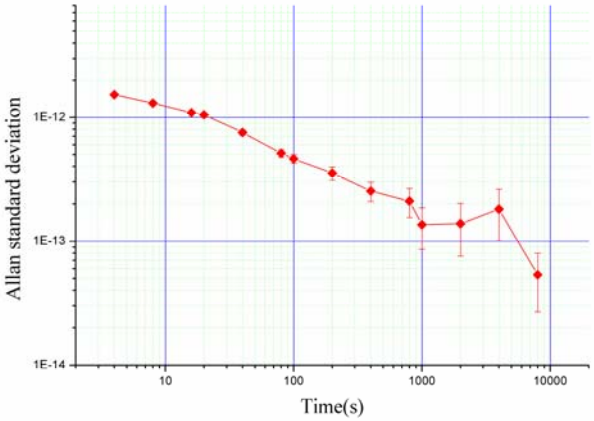


Figure 9. Allan standard deviation of ^{85}Rb atomic fountain.

IV. CORRECTED SYSTEMATIC FREQUENCY BIASES

The frequency biases and frequency uncertainties due to the systematic errors in the fountain were estimated in experiments, and the corrected biases included second-order Zeeman shift, blackbody radiation shift and gravitational redshift.

The clock transition of the rubidium fountain occurred in hyperfine ground state $|F=2,m_F=0\rangle$ to $|F=1,m_F=0\rangle$, and the first-order Zeeman shift was zero and the second-order Zeeman shift must be considered due to the presence of the C-field which afforded the quantized axis for the atoms. This field brought clock frequency shift by $\Delta\nu = 0.0575\text{Hz}/\mu\text{T}^2$. In our fountain, the magnetic map was carried out by measuring the Rabi transition envelope, two approaches were employed by microwave antenna coupling[15] and simulated Raman transition[16] respectively, and the data of the methods coincided well. The magnetic field from lower cut-off waveguide of the Ramsey cavity to the apogee of the atomic trajectory was illustrated in Fig.10 by using the simulated Raman transition approach. The second-order Zeeman relative shift was estimated as 143.6×10^{-15} in the case, the current of the C-field had been monitored for a long time (more than one day) and the fluctuation was less than 0.015%, the corresponding frequency uncertainty was less than 5×10^{-16}.

The blackbody radiation shift was a little big as the microwave cavity temperature was kept at 52.0℃ to adjust the resonance frequency of the cavity to 6.835GHz. Four PT100 thermal resistances were adhered to the vacuum tube of interaction zone to monitor the temperature of the zone, three of them were monitors and the other one was employed as a reference for the control system in the fountain. The temperature fluctuation in the fountain was shown in Fig.11, which indicated that the temperature gradient was about 2K and the fluctuation was ±0.1K, the blackbody shift was calculated to be -17.6×10^{-15} and the uncertainty was less than 5×10^{-17} by the quoting coefficient[17].

The clock frequency also changed with the gravitational potential. The gravitational redshift was measured to be 1.0×10^{-15} and the uncertainty was about 1×10^{-16} by measuring the altitude of the fountain from GPS.

V. UNCORRECTED SYSTEMATIC FREQUENCY BIASES

The cold atom collision shift for rubidium was very small compared with cesium. As the shift was estimated to be less than 0.8×10^{-16} according to the experimental result in BNM-SYRTE[18] and the uncertainty was less than 0.4×10^{-16}, the shift was not corrected. And direct measurement by method of high/low density is one of our following works.

The Majorana transition due to nonadiabatic change of the magnetic field encountered by the atoms resulted in frequency shift of the fountain[19,20]. To estimate the effect, the magnetic map in whole vacuum device was measured by the simulated Raman transition shown in Fig. 12 and the zero field was not found. The atomic numbers in state $m_F\neq0$ after the state-selection process were not found with in SNR at about 100, and the shift was conservatively evaluated to be less than 4×10^{-16}.

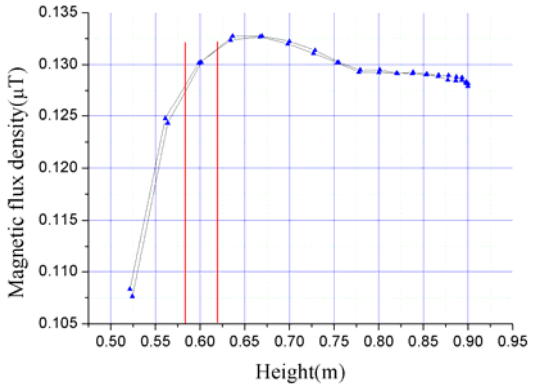


Figure 10. Magnetic map above the lower cut-off waveguide. The horizontal axis is height above cooling zone central, the area between the two red line is Ramsey cavity, the magnetic flux density in the different position is measured by the simulated Raman transition approach.

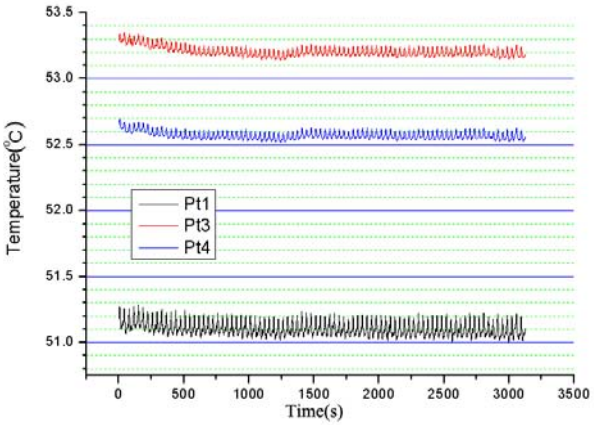


Figure 11. Temperature fluctuation of the Ramsey interaction in the fountain. The different color line represents the thermal resistor in the different position.

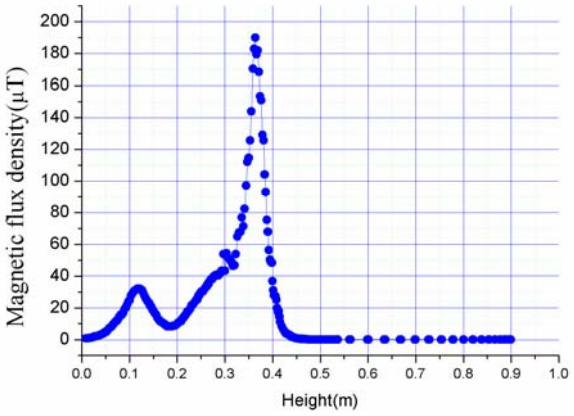


Figure 12. Magnetic map above cooling zone.

Other uncorrected frequency biases are to be measured, such as the correlative shift with microwave and cavity, light shift, Ramsey and Rabi pulling and so on.

VI. CONCLUSIONS

A rubidium fountain clock, continuously worked for more than 10 days, was developed in SIOM, the statistical error and the systematic error were estimated. The short-term frequency stability of the fountain, limited by the microwave synthesis chain, was $9\times10^{-13}\tau^{-1/2}$ and the long-term stability had reached 6×10^{-15} in one day.

To improve the performance of the fountain, a better simplified and stable optical system, with new microwave synthesis chain, effective prepared method of the atomic state and a set of the precise control system are in process. In the near future, a new microwave synthesis chain will be used, and the short-term stability is expected to be less than $3\times10^{-13}\tau^{-1/2}$.

ACKNOWLEDGEMENTS

The authors would like to thank Giorgio Santarelli and Hai-feng Jiang for their help on microwave synthesizer, and the work is supported by National Natural Science Foundation of China No. 10974215 and the Open Research Fund of State Key Laboratory of Precision Spectroscope.

REFERENCES

[1] N. F. Ramsey, “History of atomic clocks”, J. Res. NBS 88301, 1983.
[2] S. Chu, C. Wieman, “Laser cooling and trapping of atoms”, J. Opt. Soc. Am. B, 1989,Vol. 6, pp. 2020-2020.
[3] A. Bauch, “The PTB primary clocks CS1 and CS2”, Metrologia,2005, Vol.42, pp. 43-54.
[4] T. P. Heavner, S. R. Jefferts, E. A. Donley, J. H. Shirley and T. E. Parker, “NIST-F1: recent improvements and accuracy evaluations”, Metrologia, 2005,Vol.42, pp. 411─422.
[5] V. Gerginov, N. Nemitz, S.Weyers, R. Schröder, D. Griebsch and R. Wynands, “Uncertainty evaluation of the caesium fountain clock PTB–CSF2”, Metrologia, 2010, Vol. 47, pp. 65–79.
[6] K. Szymaniec, S. Eon Park1, G. Marra and W. Chałupczak. “First accuracy evaluation of the NPL-CsF2 primary frequency standard”. Metrologia, 2010, Vol. 47, pp. 363–376.
[7] C. Vian, P. Rosenbusch, H. Marion, S. Bize, L. Cacciapuoti, S. Zhang, M. Abgrall et al. “BNM-SYRTE Fountains: Recent Results”. IEEE Trans. Instrum. Meas. 2005, Vol. 54, no. 2.
[8] Z. C. Zhou, R.Wei, C.Y. Shi, D.S. Lv, T. Li, Y. Z. Wang, “Progress of the ^{87}Rb Fountain Clock”, Chin. Phys. Lett. 2009, Vol.26, 123201.
[9] R. Wei, et al. “Fold optics path: an improvement for atomic fountain”, Chin Phys. B. 2011,Vol.20(in press)
[10] G. Santarelli, C. Audoin, “Frequency Stability Degradation of an Oscillator Slaved to a Periodically Interrogated Atomic Resonator”, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 1995, Vol. 45, no. 4.
[11] G. J. Dick, “Local oscillator induced instabilities in trapped ion frequency standards", in Proc. Precise Time and Time Interval, Redondo Beach, 1987, pp. 133 – 147.
[12] T. Li, R. Wei, C. Y. Shi, Z. C. Zhou, Y.Z.Wang, “Realization of an ^{85}Rb Atomic Fountain”. Chin. Phys. Lett. 2010, Vol. 27, 083201.
[13] J. D. Prestage, R. L. Tjoelker, L. Maleki, “Atomic clock and variations of the fine structure constant”, Phys. Rev. Lett. 1995, Vol. 74, 18.
[14] H. Marion, F. Pereira Dos Santos, M. Abgrall, S. Zhang, Y. Sortais, S. Bize et al. “Search for variations of fundamental constants using atomic fountain clocks”, Phys. Rev. Lett. 2003, Vol. 90, pp. 15.
[15] C. Y. Shi, R. Wei, Z. C. Zhou, D. S. Lv, T. Li, and Y. Z. Wang, “Magnetic field measurement on ^{87}Rb atomic fountain clock”, Chin. Opt. Lett., 2010, Vol. 6, pp.549.
[16] Z. C. Zhou, R. Wei, C. Y. Shi, T. Li, and Y. Z. Wang, “Magnetic field measurement based on a stimulated two-photon Raman transition”. Chin. Phys. B, 2011, Vol. 20, pp. 034206.
[17] E. J. Angstmann, V. A. Dzuba, and V. V. Flambaum, “Frequency shift of hyperfine transitions due to blackbody radiation”, Phys. Rev. A, 2006, Vol. 74, no. 023405.
[18] Y. Sortais, S. Bize, C. Nicolas, A. Clairon, C. Salomon, C. Williams et al. “Cold Collision Frequency Shifts in a ^{87}Rb Atomic Fountain”, Phys. Rev. Lett.,2000, Vol. 85, pp.15.
[19] E. Majorana, “Atomi orientati in campo magnetico variabile,” Nuovo Cim.N.S., 1932, Vol. 9, pp. 43–50.
[20] A. Bauch and R. Schröder, “Frequency shifts in a cesium atomic clock due to Majorana transitions”, Ann.Physik. 1993, Vol. 2, pp.421-449.

The 2nd Harmonic Signal in Vapor-Cell Atomic Clocks

Gilda Fathi and James Camparo

Physical Sciences Laboratories
The Aerospace Corporation
2310 E. El Segundo Blvd., El Segundo, CA
james.c.camparo@aero.org

***Abstract*—The 2nd harmonic signal has been recognized for some time as a "status-of-health" indicator for Rb clocks. Unfortunately, while the 2nd harmonic signal can provide *qualitative* assessments of clock-signal health, at present there is no analytic theory capable of connecting the 2nd harmonic to various parameters of interest. In large part, the difficulty lies in the fact that clocks operate in a regime where none of the standard approximations for solving field/matter interaction problems apply, in particular the Quasi-Static Approximation (QSA) which arguably has the greatest intuitive appeal. Here, we describe our first set of studies examining the validity of the QSA under conditions where the Rabi frequency, the field's modulation frequency, and the atom's dephasing rate are all nearly equal. Specifically, we examine the QSA's prediction of the 2nd harmonic signal amplitude as a function of microwave power. Our results suggest that the QSA can have utility for understanding the *relative* change in the 2nd harmonic signal with microwave power, and therefore potential *quantitative* utility for understanding a Rb clock's status-of-health.**

I. INTRODUCTION

In the standard frequency-feedback loop of atomic clocks, a field's frequency is modulated at f_m; the atomic dynamics are detected and demodulated at the modulation's fundamental frequency (i.e., the 1st harmonic signal), and this demodulated signal is then used to generate a correction voltage for a crystal oscillator. Typically, the intuitive understanding of this process is obtained by viewing the atomic dynamics in terms of the quantum system's static response to the resonant field:

1st harmonic dynamics ⇔ Static lineshape 1st derivative,

2nd harmonic dynamics ⇔ Static lineshape 2nd derivative.

Consequently, the Quasi-Static Approximation (QSA) to the atomic dynamics suggests that a 2nd harmonic signal can be employed as a measure of the atoms' frequency discrimination capability (i.e., the slope of the static lineshape's 1st derivative). In a phrase, the 2nd harmonic signal can be employed as a vapor-cell clock's status-of-health indicator.

However, in order to obtain large signal-to-noise ratios and narrow atomic linewidths in atomic clocks, it is typical to operate under conditions such that $f_m \sim \Omega \sim \gamma$, where Ω and γ are the Rabi frequency and dephasing rate, respectively. Unfortunately, this regime of operation is neither quasi-static (large γ), adiabatic (large Ω), nor sudden (large f_m). Consequently, none of the standard approximations of atomic dynamics legitimately apply, notably the QSA.

In order to develop a better intuitive understanding of atomic dynamics in the regime of clock operation, and in particular the nature of the atoms' 2nd harmonic response to a frequency modulated field, we have begun a series of experiments investigating 2nd harmonic signals under clock operating conditions (i.e., $f_m \sim \Omega \sim \gamma$). Here, we discuss our first set of experiments examining the amplitude of 2nd harmonic signals as a function of microwave power, and the comparison of our results to the predictions of a QSA model. In the next section, we outline the QSA model, and discuss the conditions of clock operation that would lead one to doubt its validity. Then in Section III we discuss our experimental system. Our results are presented in Section IV, where they are compared with the QSA model as well as a 2-level density matrix model of the 2nd harmonic signal. Our conclusions are presented in Section V.

II. THE QUASI-STATIC APPROXIMATION

A. Overview of the QSA Model

As is well known, the signal of a Rb atomic clock is obtained by passing the light from a lamp through an absorption cell contained within a microwave cavity as shown in Fig. 1 [1]. In the absence of microwaves, the light intensity transmitted by the vapor is high as a consequence of optical pumping [2], which removes atoms from the optically absorbing state. However, when resonant microwaves are applied to the vapor atoms return to the optically absorbing state, and the transmitted light intensity decreases. Thus, the light intensity transmitted by the vapor carries information on the atoms' response to the microwave field, which is the origin of the Rb atomic clock signal.

Under static conditions, the light intensity transmitted by a Rb vapor of length L, I(L), is given by the Beer-Lambert law of exponential attenuation [3]:

$$I(L) = I_o e^{-\eta N \sigma L}. \quad (1)$$

Here, I_o is the light intensity at the entrance to the vapor, N is the number density of Rb atoms in the vapor, σ is the optical absorption cross section, and η corresponds to the fraction of

This work was funded by U.S. Air Force Space and Missile Systems Center under Contract No. FA8802-04-C-0001.

atoms in the absorbing state. According to the Generalized Vanier Theory of the alkali atom's 0-0 hyperfine transition [4], η is given by

$$\eta = \frac{g_p \gamma_c}{g_u B + g\gamma_c}\left[\frac{\gamma^2 + \Delta^2 + \left(\gamma/\Gamma_\alpha\right)\Omega^2}{\gamma^2 + \Delta^2 + \left(\gamma/\Gamma_\beta\right)\Omega^2}\right], \qquad (2)$$

where Δ is the microwave field detuning,

g_p = degeneracy of the optically excited ground-state hyperfine level, $F = \widetilde{F}$,

g_u = degeneracy of the non-optically excited ground-state hyperfine level,

γ_c = dephasing rate due to "collisions," [a]

B = optical photon absorption rate,

$$g = g_p + g_u,$$

$$\gamma = \frac{B}{2} + \gamma_c,$$

$$\Gamma_\alpha = \frac{g_p \gamma_c (B + \gamma_c)}{g_p (B + \gamma_c) - \widetilde{F} B},$$

$$\Gamma_\beta = \frac{2\gamma_c (B + \gamma_c)(g_u B + g\gamma_c)}{(g_u - 1)B^2 + (g + 2g_u)B\gamma_c + 2g\gamma_c^2}.$$

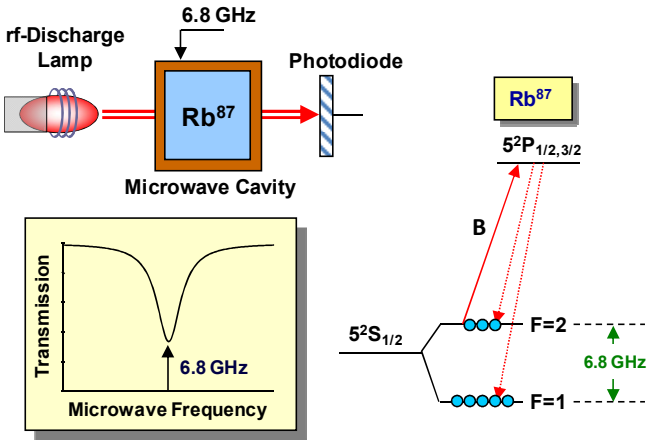


Figure 1: Generic diagram of a Rb atomic clock. In this figure, the lamplight excites Rb^{87} atoms out of the F = 2 ground-state hyperfine level, so that $\widetilde{F} = 2$, $g_p = 5$, and $g_u = 3$. The atoms are excited out of the F = 2 hyperfine manifold at the rate B.

As employed here, *the QSA model assumes that the dephasing rate of the atomic system is larger than all other parameters describing the atomic dynamics* [5]. In particular, if the field's frequency, ν, is modulated in some fashion:

$$\Delta(t) \equiv \nu(t) - \nu_o = (\bar{\nu} - \nu_o) + \delta_m g(t, f_m), \qquad (3)$$

where $\bar{\nu}$ is the field's average frequency, δ_m is the amplitude of the periodic frequency modulation, and $g(t,f_m)$ is an arbitrary periodic function of time with period $1/f_m$; then the QSA assumes that $\gamma >> f_m, \delta_m$, along with the assumption that the system is not in the regime of power broadening: $\gamma >> \Omega$.

We now expand I(L) in a Taylor series about $\Delta_o \equiv (\bar{\nu} - \nu_o)$:

$$I(t) = I(L, \Delta_o) + \delta_m g(t, f_m)\left.\frac{\partial I}{\partial \Delta}\right|_{\Delta_o} + \frac{\delta_m^2 g^2(t, f_m)}{2}\left.\frac{\partial^2 I}{\partial \Delta^2}\right|_{\Delta_o} + \ldots, \qquad (4a)$$

or, since the transmitted light intensity only depends on the microwave detuning through η,

$$I(t) = I(L, \Delta_o) + \delta_m g(t, f_m)\left(\frac{\partial I}{\partial \eta} + \frac{\delta_m g(t, f_m)}{2}\frac{\partial^2 I}{\partial \eta^2}\frac{\partial \eta}{\partial \Delta}\right)\left.\frac{\partial \eta}{\partial \Delta}\right|_{\Delta_o} + \frac{\delta_m^2 g^2(t, f_m)}{2}\frac{\partial I}{\partial \eta}\left.\frac{\partial^2 \eta}{\partial \Delta^2}\right|_{\Delta_o} + \ldots. \qquad (4b)$$

Since we're interested in on-resonance conditions (i.e., $\Delta_o = 0 \Rightarrow \eta = \eta_o$), the first derivative of η with respect to Δ is zero, and so we have to second-order in the Taylor-series expansion

$$\frac{\delta I(t)}{\langle I \rangle} \equiv \frac{I(t)}{I_o e^{-\eta_o N\sigma L}} - 1 = -N\sigma L \frac{\delta_m^2 g^2(t, f_m)}{2}\left.\frac{\partial^2 \eta}{\partial \Delta^2}\right|_{\Delta_o = 0}. \qquad (5)$$

Finally, computing the derivatives of Eq. (2) we obtain

$$\frac{\delta I(t)}{\langle I \rangle} = -\frac{N\sigma L g_p \gamma_c}{(g_u B + g\gamma_c)}\frac{\gamma(\Gamma_\alpha - \Gamma_\beta)}{\Gamma_\alpha \Gamma_\beta}\frac{\delta_m^2 g^2(t, f_m)\Omega^2}{\left(\gamma^2 + \left(\gamma/\Gamma_\beta\right)\Omega^2\right)^2}. \qquad (6)$$

B. The Atom's 2nd Harmonic Response

From the photodiode shown in Fig. 1, the signal will typically pass to a splitter, one branch of which will go to the RF port of the "fundamental-signal" mixer. The local oscillator of this mixer will be derived from $g(t,f_m)$ (e.g., $\sin(2\pi f_m t)$), and this mixer's output, S_1, will provide the correction signal that is used to lock the microwave field's frequency to the 0-0 hyperfine resonance frequency (i.e., $dS_1/d\Delta_o$).

The other branch of the split signal will also go to the RF port of a mixer. However, in this case the LO signal will be derived from $g^2(t,f_m)$ (e.g., $\cos(4\pi f_m t)$). The output of this mixer, S_2, yields the 2nd harmonic signal of interest for status-of-health monitoring, since it provides a measure of $d^2S_1/d\Delta_o^2$, and after passing through a low pass filter the amplitude of S_2 will be given by

$$S_2 = \frac{N\sigma L g_p \gamma_c}{(g_u B + g\gamma_c)}\frac{\gamma(\Gamma_\alpha - \Gamma_\beta)}{\Gamma_\alpha \Gamma_\beta}\frac{\delta_m^2 \Omega^2}{\left(\gamma^2 + \left(\gamma/\Gamma_\beta\right)\Omega^2\right)^2}. \qquad (7)$$

[a] By collisions, we mean all processes other than light absorption that lead to dephasing (e.g., diffusion out of the light field). Additionally, to simplify the equations, we have assumed that the longitudinal "collisional" relaxation rate equals the collisional dephasing rate.

There are at least two things worth noting about Eq. (7). First, independent of the form of the modulation (e.g., sinewave or squarewave), the QSA predicts a 2nd harmonic signal. This is somewhat surprising in the case of squarewave modulation, since the microwave field is never truly on resonance; it simply samples the field at two specific frequencies symmetrically placed about resonance. Nevertheless, the QSA recognizes that the nonlinear nature of the atom's response to the field is captured by the atomic dynamics, and that a 2nd harmonic signal is produced.

Additionally, the QSA indicates that S_2 will have a resonance-like response to the Rabi frequency. This is perhaps best illustrated by writing $\Omega^2 = \alpha P$, where P is the microwave power, and then transforming to a new variable $y \equiv \ln(\alpha P/\gamma\Gamma_\beta)$. In effect, the variable "y" is akin to the microwave power measured in dB. With this transformation, Eq. (7) becomes

$$S_2 = \left(\frac{N\sigma L g_p \gamma_c}{(g_u B + g\gamma_c)} \frac{(\Gamma_\alpha - \Gamma_\beta)}{4\Gamma_\alpha} \frac{\delta_m^2}{\gamma^2} \right) \operatorname{sech}^2\left(\frac{y}{2}\right), \quad (8)$$

where $\operatorname{sech}^2(y/2)$ has a well-known resonance-like shape: resonance occurs at $y = 0$ (i.e., $\alpha P = \gamma\Gamma_\beta$) and the symmetric lineshape has a FWHM of $\Delta y_{1/2} \cong 3.526 \Rightarrow 15.31$ dB (i.e., $\Omega^2_{high} = \alpha P_{high} \cong 5.83\gamma\Gamma_\beta$ and $\Omega^2_{low} = \alpha P_{low} \cong 0.17\gamma\Gamma_\beta$.)

C. *Fundamental Problems for the QSA*

There are at least three serious problems with the quasi-static approximation in terms of its ability to describe the atomic dynamics occurring in vapor cell clocks. First, and perhaps most serious, vapor-cell atomic clocks *do not* operate in a regime where $\gamma \gg f_m, \delta_m$, and Ω. In order to ensure a high signal-to-noise ratio without losing atomic line-Q via power broadening, vapor-cell atomic clocks operate with $\Omega \sim \gamma$ (i.e., clocks operate at the onset of saturation). Further, since the linewidth of the 0-0 hyperfine resonance is on the order of γ, S_1 can be made largest by choosing $\delta_m \sim \gamma$ [6]. Finally, the signal amplitude can be further enhanced through a process akin to adiabatic fast passage if $f_m\delta_m \sim \Omega^2$ (and as a consequence of $\Omega \sim \gamma$, $f_m\delta_m \sim \gamma^2$) [7]. Therefore, for reasons of maximum signal-to-noise ratio and atomic line-Q, vapor-cell atomic clocks typically operate under conditions where $\gamma \sim f_m \sim \delta_m \sim \Omega$.

In addition to the above considerations, squarewave frequency modulation posses an additional problem. When the frequency of a reasonably strong resonant field changes abruptly, a quantum system will undergo transient nutation [8]. In the atomic dynamics, this will appear as a damped oscillation: the oscillation frequency will be given by $\sqrt{\Omega^2 + \Delta_o^2}$, and the decay time will be $\sim \gamma^{-1}$. Of course, none of this is captured by the quasi-static approximation, and we are left to wonder how the transient nutation would influence a 2nd harmonic signal.

Finally, though not as serious as the previous issues but nonetheless problematic, the QSA truncates the Taylor series expansion at 2nd order. While this yields a relatively simple expression for the 2nd harmonic signal, it is not necessarily true that the higher-order terms are negligible. In particular, as noted above $\delta_m \sim \gamma$, suggesting that terms containing $(\delta_m/\gamma)^\kappa$ will not be negligible for $\kappa > 2$.

For all the reasons cited in the preceding paragraphs, we must expect the QSA to fail in predicting the behavior of the 2nd harmonic signal. The question we want to (partially) address in the following sections concerns the relative magnitude of this failure. Said differently, to what extent can the QSA provide us with *any* insight into S_2?

III. EXPERIMENT

Figure 2 shows a block diagram of our experimental arrangement. We stabilize our VCSEL diode laser wavelength using the ^{87}Rb resonance cell on the left of the figure (cell A); this cell contains 10 torr of N_2 as a buffer gas, and is maintained at a temperature of 37 ^{o}C. Though the Rb atomic clocks employed in GNSS and satellite communication systems use an rf-discharge lamp [9] rather than a diode laser as their light source [10-12], we use a laser to give us better control of the light field. With the laser, we not only have the ability to vary the intensity of the light and see its effect on the 2nd harmonic signal, we can also easily vary the laser frequency. Varying the laser center frequency gives us an opportunity to mimic spectral shifts in an rf-discharge lamp's output [13], and to thereby study the effect of such variations on the 2nd harmonic signal, which might arise as a consequence of the light shift [14].

The light intensity passing through cell A is kept very low, so that optical pumping does not occur. The laser light also passes through a second ^{87}Rb resonance cell (cell B), which again contains 10 torr of N_2 as a buffer gas. The laser light optically pumps the rubidium atoms in cell B, which is our atomic-clock signal cell, and creates a population imbalance between the atom's ground state hyperfine levels. Since the optical pumping process essentially removes atoms from the optically-absorbing hyperfine level, the laser light passing through this second cell is relatively high when the optical pumping process completes. (The optical pumping process is relatively fast, taking roughly a millisecond to complete.) If microwaves of the appropriate frequency impinge on the vapor of atoms, however, atoms are forced to return to the optically absorbing hyperfine level with a concomitant decrease in the transmitted light intensity. Consequently, as illustrated in the top graph on the right, plotting the vapor's absorption as a function of microwave frequency yields a Lorentzian absorption profile. In addition, we modulate the frequency of the microwaves in our experiment at $f_m \sim 100$ Hz. We can then monitor the atoms' first harmonic response to the microwaves or the atoms' second harmonic response to the microwaves using a lock-in amplifier and the appropriate reference frequency for the lock-in. This is shown in the lower graph on the right in Fig. 2. In this work, we are primarily concerned with the amplitude of the 2nd harmonic signal when the microwaves are on resonance, corresponding to locked vapor-cell clock conditions.

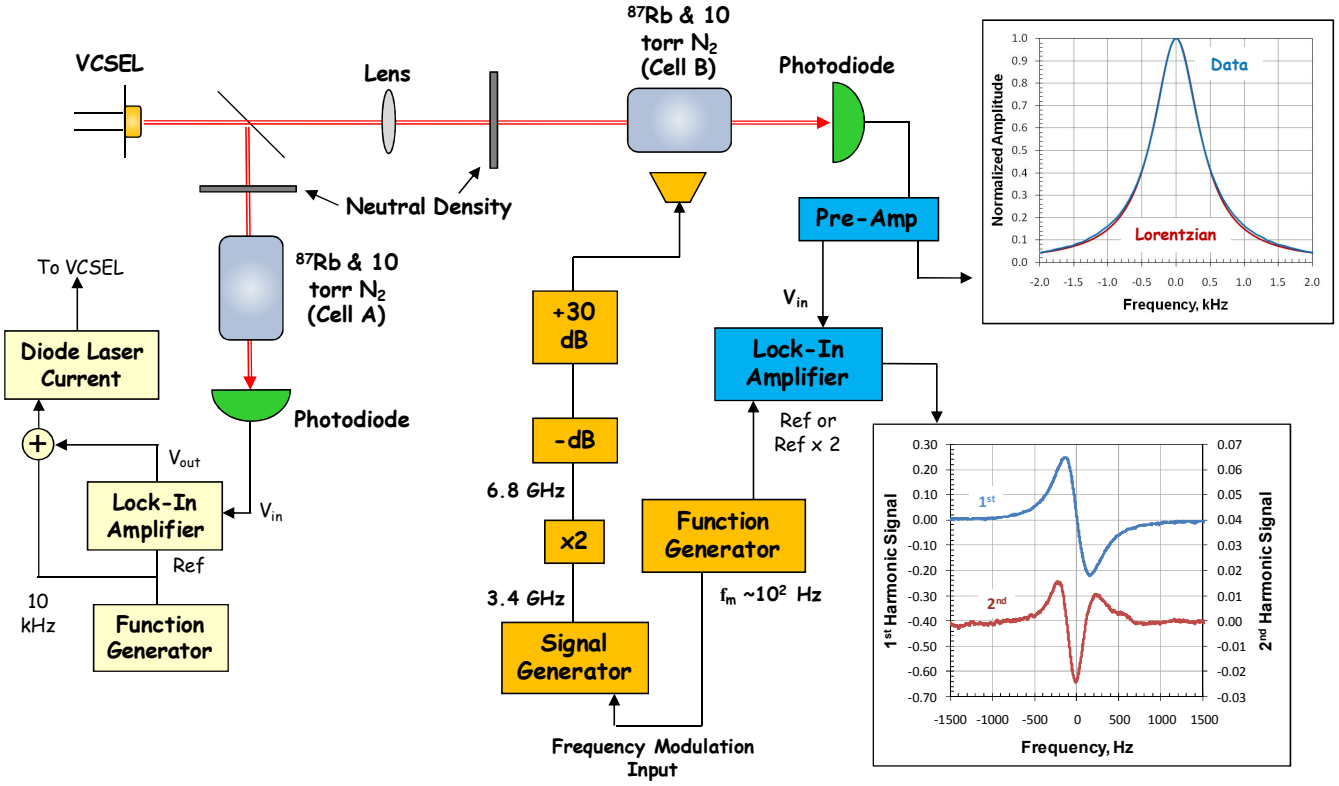


Figure 2: Block diagram of our experimental arrangement.

IV. Results

A. Experiment vs. Theory

The curve of the 2nd harmonic amplitude as a function of Rabi frequency is known as a "Rabi resonance" [15], and this is shown for our experiment in Fig. 3: $\gamma = 136$ Hz, $f_m = 118$ Hz, and $\delta_m = 120$ Hz. In the figure, the blue curve (rightmost) corresponds to a 2-level density matrix calculation, and the reddish curve (leftmost) to the predictions of the QSA. Note that the density matrix calculations and the experimental results agree reasonably well, indicating that the overall nature of the atomic dynamics (at least with regard to S_2's dependence on Ω) is not overly influenced by the multi-level nature of the Rb system. At present, we believe that the slight disagreement between the density matrix theory and experiment may be due to inhomogeneous broadening [16],[a] and in future work we plan to examine this possibility more carefully.

If we don't concern ourselves with the center position of the Rabi resonance, then Fig. 4 shows S_2 as a function of microwave power in dB, where the normalizing power is set equal to that corresponding to the peak of the Rabi resonance in the three separate cases. Again the reddish curve (now the wider curve) corresponds to the QSA, and the blue curve (i.e., the narrower curve) corresponds to the 2-level density matrix calculation. In the case of Fig. 4, the QSA does very well in predicting the shape of the Rabi resonance, to some extent better than that of the 2-level density matrix calculations. However, we should not infer too much from the QSA's better agreement with the experimental data, since this is very likely coincidental, and due to the nature of our (hypothesized) inhomogeneously-broadened transition lineshape.

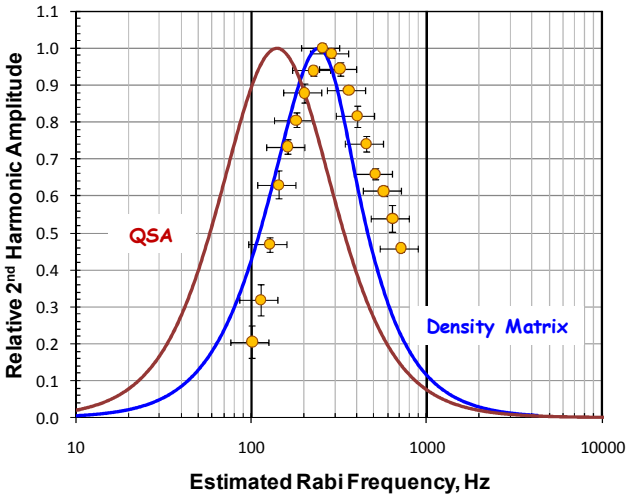


Figure 3: The Rabi resonance: S_2 vs. Ω. For these results, the resonance cell temperature was 33 °C, corresponding to an optical attenuation coefficient, $N\sigma L$, of 0.8.

[a] Nevertheless, it seems strange that inhomogeneous broadening should play a significant role in our experiments. As illustrated in Fig. 2, the 87Rb 0-0 transition lineshape is very well described by a Lorentzian. Typically, inhomogeneous broadening leads to clear non-Lorentzian character in a transition lineshape.

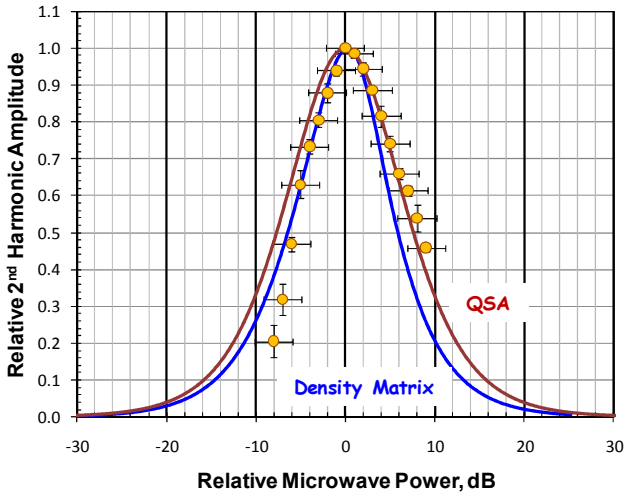


Figure 4: The Rabi resonance: S_2 vs. $20\text{Log}_{10}[\Omega/\Omega_{peak}]$. Here, Ω_{peak} is the microwave Rabi frequency that maximizes the Rabi resonance. The data is the same as that of Fig. 3.

The results of Figs. 3 and 4 suggest that the QSA is likely fairly good in predicting the *shape* of the Rabi resonance, but not so much its peak position. However, from a practical standpoint, it is the shape of the Rabi resonance that has the greatest value in the 2nd harmonic signal's employment as a status-of-health indicator: given a change in the 2nd harmonic signal, we want to know the relative microwave power change that could have occurred. Therefore, given the agreement between the 2-level density matrix calculations and experiment, we decided to more fully examine the QSA's utility for status-of-health monitoring by comparing the predictions of the QSA to those of the 2-level density matrix under a wide variety of modulation conditions.

B. QSA vs. 2-Level Density Matrix

From Eq. (7), it is straightforward to show that $\Omega_{peak} = (\gamma\Gamma_\beta)^{1/2}$ for the QSA. Note, however, that in the limit of large ground-state degeneracy g,

$$\frac{\Gamma_\beta}{\gamma} \cong \frac{4\left(1+B/\gamma_c\right)}{\left(2+B/\gamma_c\right)^2}. \qquad (9)$$

Thus, under the conditions of most interest for vapor phase clocks (i.e., $B \le 3\gamma_c$), we have $0.8 \le (\Gamma_\beta/\gamma)^{1/2} \le 1$; so that for vapor-phase clocks the QSA yields $\Omega_{peak} \cong \gamma$.

Of course based on Fig. 3, we know that this prediction of the QSA is in error when the modulation parameters are not small. However, we also know from previous work on Rabi resonances that when the modulation parameters are large we expect $\Omega_{peak} \cong 2f_m$ [15]. Thus, a reasonable empirical expression to hypothesize for the Rabi frequency that maximizes S_2 is

$$\Omega_{peak} = \sqrt{\gamma^2 + 4f_m^2}. \qquad (10)$$

Figure 5 shows the peak of the Rabi resonance, Ω_{peak}, (relative to the expression given by Eq. (10)) as a function of the normalized modulation frequency: circles and diamonds correspond to two different values of the modulation depth, while the dashed line is the prediction of the QSA. The first thing to note from the figure is that in the limit of small modulation frequency and small modulation depth the 2-level density matrix calculations appear to asymptote to the QSA. More importantly, however, under conditions of vapor cell clock operation: $f_m \sim \gamma$ and $\delta_m \sim \gamma$, the empirical expression of Eq. (10) appears to be a reasonably good estimator of the Rabi frequency that yields the maximum 2nd harmonic signal.

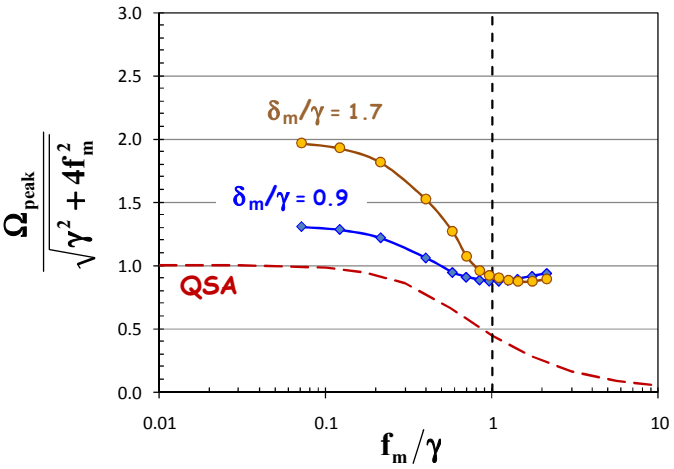


Figure 5: This figure shows the Rabi frequency that maximizes S_2, Ω_{peak}, as a function of the relative modulation frequency. Note that Ω_{peak} is plotted relative to the empirical expression of Eq. (10). In the figure, circles and triangles correspond to two different values of the frequency modulation amplitude, δ_m, with the results obtained from the 2-level density matrix calculations. The dashed line is Ω_{peak} as obtained from the QSA: $\Omega_{peak} = \gamma$.

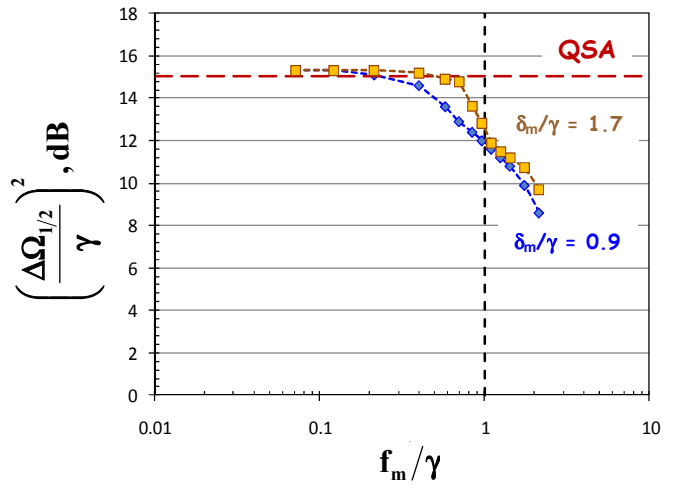


Figure 6: This figure shows the FWHM of the Rabi resonance, $\Delta\Omega_{1/2}$, as a function of the relative modulation frequency; the circles and triangles correspond to two different values of the frequency modulation amplitude, δ_m, with the results obtained from the 2-level density matrix calculations. The dashed line is $\Delta\Omega_{1/2}$ as obtained from the QSA.

Figure 6 shows the FWHM of the Rabi resonance in dB (i.e., Fig. 4 and Eq. (8)), as a function of the relative modulation frequency. Again, the 2-level density matrix calculations agree with the prediction of the QSA in the limit

of small modulation frequency. However, even under conditions of typical clock operation (i.e., $f_m \sim \gamma$), the agreement between the QSA and the 2-level density matrix calculations is not too bad: the QSA only overestimates the FWHM by about 20%, and for engineering status-of-health purposes this is likely a tolerable error.

V. CONCLUSIONS

In this work, we have examined the quasi-static approximation (QSA) and its ability to predict the magnitude of the 2nd harmonic signal in vapor-cell atomic clocks as a function of microwave power. More specifically, we examined the QSA when the microwaves are resonant with the 0-0 hyperfine transition frequency, since this is the only condition of relevance for the 2nd harmonic signal's use as a status-of-health indicator. In general, the QSA *does not* predict the position of the Rabi resonance very well. However, this short-coming of the QSA is easily overcome through the use of the empirical formula given by Eq. (10): $\Omega_{peak} = \sqrt{\gamma^2 + 4f_m^2}$.

Under vapor-cell clock operating conditions, the QSA *does* appear to do an adequate job in predicting the FWHM of the Rabi resonance (in dB), though it overestimates the width by about 20%. For the engineering application of status-of-health monitoring, this is likely not too large an error. Specifically, if a change in microwave power produces a significant change in the 2nd harmonic signal, we should be able to use the QSA to estimate approximately how large a microwave power change occurred.

REFERENCES

[1] J. Camparo, "The rubidium atomic clock and basic research," Phys. Today, vol. 60(11), pp. 33-39, 2007.

[2] A. L. Bloom, "Optical pumping," Sci. Am., vol. 203, pp. 72-80, 1960; L. C. Balling, "Optcal pumping," Adv. Quantum Electron., vol. 3, pp. 1-167, 1975; W. Happer, "Optical pumping," Rev. Mod. Phys., vol. 44(2), pp. 169-249, 1972.

[3] P. Atkins, Physical Chemistry, W. H. Freeman & Co: New York, 1998, Ch. 16.

[4] J. C. Camparo and R. P. Frueholz, "Linewidths of the 0-0 hyperfine transition in optically pumped alkali-metal vapors," Phys. Rev. A, vol. 31(3), pp. 1440-1448, 1985; J. C. Camparo and R. P. Frueholz, "Saturation of the 0-0 hyperfine-transition linewidth enhancement factor in optically pumped alkali-metal vapors," Phys. Rev. A, vol. 32(3), pp. 1888-1889, 1985.

[5] O. E. Myers and E. J. Putzer, "Measurement broadening in magnetic resonance," J. Appl. Phys., vol. 30(12), pp. 1987-1991, 1959.

[6] J. Vanier and L.-G. Bernier, "On the signal-to-noise ratio and short-term stability of passive rubidium frequency standards," IEEE Trans. Instrum. Meas., vol. 30(4), pp. 277-282, 1981.

[7] J. Camparo and R. Frueholz, "Parameters of adiabatic rapid passage in the 0-0 hyperfine transition of ^{87}Rb," Phys. Rev. A, vol. 30(2), pp. 803-811, 1984.

[8] H. C. Torrey, "Transient nutations in nuclear magnetic resonance," Phys. Rev., vol. 76(8), pp. 1059-1068, 1949.

[9] W. E. Bell, A. L. Bloom, and J. Lynch, "Alkali metal vapor spectral lamps," Rev. Sci. Instrum., vol. 32(6), pp. 688-692, 1961; R. G. Brewer, "High intensity low noise rubidium light source," Rev. Sci. Instrum., vol. 32(12), pp. 1356-1358, 1961; V. B. Gerard, "Laboratory alkali metal vapour lamps for optical pumping experiments," J. Sci. Instrum., vol. 39, pp. 217-218, 1962.

[10] T. McClelland, I. Pascaru, and M. Meirs, "Development of a rubidium frequency standard for the Milstar satellite system," in *Proc. 41st Annual Frequency Control Symposium*. Piscataway, NJ: IEEE Press, 1987, pp. 66-74.

[11] A. Jeanmaire, P. Rochat, F. Emma, "Rubidium atomic clock for Galileo," in Proc. 31st Annual Precise Time and Time Interval (PTTI) Meeting. Washington, DC: US Naval Observatory, 2000, pp. 627-636.

[12] R. Dupuis, T. J. Lynch, and J. R. Vaccaro, "Rubidium frequency standard for the GPS IIF program and modifications for the RAFSMOD program," in *Proc. 2008 IEEE International Frequency Control Symposium*. Piscataway, NJ: IEEE Press, 2008, pp. 655-660.

[13] N. Kuramochi, T. Matsuo, I. Matsuda, and H. Fukuyo, "Spectral profiles of the 87 Rb D1 line emitted from a spherical electrodeless lamp," Jap. J. Appl. Phys., vol. 16(5), pp. 673-679, 1977.

[14] S. Pancharatnam, "Light shifts in semiclassical dispersion theory," J. Opt. Soc. Am., vol. 56(11), pg. 1636, 1966; B. S. Mathur, H. Tang, and W. Happer, "Light shifts in the alkali atoms," Phys. Rev., vol. 171(1), pp. 11-19, 1968; J. C. Camparo, R. P. Frueholz, and C. H. Volk, "Inhomogeneous light shift in alkali-metal atoms," Phys. Rev. A, vol. 27(4), pp. 1914-1924, 1983.

[15] J. G. Coffer, B. Sickmiller, A. Presser, and J. C. Camparo, "Line shapes of atomic-candle type Rabi resonance," Phys. Rev. A, vol. 66, 023806, 2002.

[16] P. J. Oreto, Y.-Y. Jau, A. B. Post, N. N. Kuzma, and W. Happer, "Buffer-gas-induced shift and broadening of hyperfine resonances in alkali-metal vapors," Phys. Rev. A, vol. 69, 042716, 2004.

Ramsey Narrowing of the Coherent Population Trapping Resonance Induced by Motion of the Atoms in Wall-Coated Cells.

G.A. Kazakov, A.N. Litvinov, B.G. Matisov
Theoretical physics department, St-Petersburg State Polytechnical University, Politekhnicheskaya 29, St-Peterburg, 195251, Russia
e-mail: andrey.litvinov@mail.ru

E.Breschi
Fribourg Atomic Physics, University of Fribourg, Chemin du Musée 3, Fribourg 1700 Switzerland

***Abstract*—In this paper we consider the theory of coherent population trapping resonance formation in cell with anti-relaxation wall coating illuminated by the narrow laser beam. We study the dependence of resonance amplitude and width on radiation intensity and illuminated zone diameter.**

I. INTRODUCTION

In Ramsey spectroscopy method [1] the ansamble of quantum systems (atoms or molecules) is interrogated by series (at least by two) of coherent field applications separated in time. The Ramsey interrogation can be realized either with an atomic (or molecular) beam passing through two spatially separated fields, or with a pulsed electromagnetic field (see [2] and references therein). The effect of atomic-motion induced Ramsey narrowing of Coherent Copulation Trapping (CPT) resonance has been studied both experimentally and theoretically in buffered and coated cells, see [3-8] and references therein. The signature of this effect is the spectral narrowing of the dark resonance due to motion of oriented atomic spin in-and-out of the optical interaction region, see Figure 1 (a). In [4] the authors identify the main feature of the phenomenon in buffer gas cell while in [5] they extensively explain the model developed for the interpretation. In [6, 8] the atomic-motion narrowing effect was studied experimentally. It was found [6] that in weak fields the light shift and light broadening increase linearly with the square of illuminated zone, i.e. it behaves only on the total power of laser beam. More detailed theory of the CPT resonance formation in partially illuminated wall coated cell without buffer gas was developed in [7]. In [8] the similar theory was developed independently and demonstrates good quantitative agreement with experiment. At the same time, divergences between theory and experiment in [8] increases with the laser intensity and beam diameter.

In present work we consider the theory [7] and the limit case of this theory in weak laser fields. In present work we consider a model based on a simplified three-level atom, see Fig.1.a. We suppose all the main parameters of atom (such as energies of levels, atomic mass, et.c.) close to the parameters of D1-transition in ^{87}Rb atom, where the transition between levels *1* and *2* is microwave and to level *3* is optical.

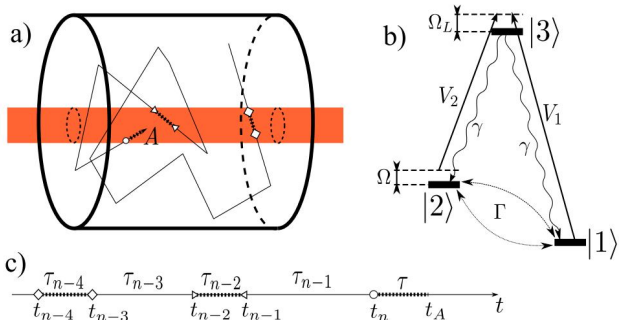


Figure 1. (a): Example of indiviual atom trajectory. Segments of trajectory are shown by fringed lines inside and by solid lines outside the beam (red). (b) Λ system describing the spin state of atom V_1 and V_2 are the laser Rabi frequencies of resonance laser field components, Γ and 2γ are the ground and exited state relaxation rates respectively, Ω_L and Ω_R are the optical and two-photon (Raman) detunings. (c): time axis illustrating the individual "history" of atom.b)

II. THEORETICAL MODEL

A. *Density matrix equations*

The density matrix evolution for this system can be inscribed in common form as:

$$\dot{\rho}_{ij} = -\frac{i}{\hbar}\sum_k \left[H_{ik}\,\rho_{kj} - \rho_{ik}\,H_{kj} \right] + \sum_{k,l} \Gamma_{ij,kl}\,\rho_{kl}\,, \qquad (1)$$

where H_{ik} is the matrix element of the atomic Hamiltonian, $\hat{H}$. Following the standard theory is $\hat{H} = \hat{H}_0 + \hbar\hat{V}(v_z,t)$ where $\hat{H}_0$ is the Hamiltonian of the unperturbed atom including static perturbation and $\hbar\hat{V}(v_z,t)$ represents the atom-laser interaction which depends on the time, *t*, and on the atomic velocity components along the quantization axis, v_z; $\Gamma_{ij,kl}$ is the relaxation matrix element describing the relaxation processess. In the case of atomic system contained in evacuated wall coated cell the relaxation processes are: the collisions between the atom and the coating material, the spin-exchange collisions and the reservoir losses. Also we suppose

that the cell length is small in comparison with wavelength λ_{21} of microwave $|1\rangle \leftrightarrow |2\rangle$ transition. In such a case we can neglect the Doppler broadening of this transition (Dicke narrowing, see [9]).

Let us firstly consider the density matrix equations for atoms inside the laser beam. After goto the rotating frame, adiabatic elimnation of exited state $|3\rangle$, and using the normalization condition $\rho_{11} + \rho_{22} = 1$, one can obtain the equation for spin state density matrix in the beam zone as:

$$\begin{aligned} \dot{f} &= G\frac{V_2^2 - V_1^2}{\gamma'} - (W+\Gamma) f - 4F\frac{V_1 V_2}{\gamma'} J, \\ \dot{R} &= -G\frac{V_1 V_2}{\gamma'} - (W+\Gamma) R - (\Omega - \Delta) J, \\ \dot{J} &= F\frac{V_1 V_2}{\gamma'} f + (\Omega - \Delta) R - (W+\Gamma) J, \end{aligned} \tag{2}$$

where $f = \rho_{11} - \rho_{22}$, $R = \mathrm{Re}(\rho_{12})$, $J = \mathrm{Im}(\rho_{12})$, $\gamma' \simeq \gamma + \Gamma_L / 2$ is the optical coherence relaxation rate, $2\gamma \simeq 3.6 \cdot 10^7$ s^{-1} is the spontaneous decay rate of excited state $|3\rangle$, Γ_L is a laser spectrum line width, $\Delta = F(V_1^2 - V_2^2)/\gamma'$ is a light shift, $W = G(V_1^2 + V_2^2)/\gamma'$ is an optical pumping rate. $F = F(v_z)$ and $G = G(v_z)$ are the real and imaginary parts of $\gamma'/(\gamma' - i(\Omega_L - k v_z))$ respectively, $k = \omega / c$ is the wave number of optical radiation. The variation of the transmitted light is proportional to the exited state population $\rho_{exc} \equiv \rho_{33}$ which can be expressed via ground state density matrix $\hat{\rho} = (f, R, J)$ as

$$\rho_{33} = \frac{W}{2\gamma} + \frac{G}{2\gamma\gamma'}\left(\left[V_1^2 - V_2^2\right] \cdot f + 4 V_1 V_2 \cdot R\right). \tag{3}$$

The set of equations (2) can be symbolically inscribed as

$$\dot{\hat{\rho}}(v_z, t) = \hat{\hat{A}}(v_z)\hat{\rho}(v_z, t) + \hat{B}(v_z), \tag{4}$$

and (4) as

$$\rho_{33}(v_z, t) = \hat{U}^T(v_z)\hat{\rho}(v_z, t) + D(v_z). \tag{5}$$

When atom is outside the laser beam, $\rho_{33} = 0$ and its ground state density matrix $\hat{\rho}$ evolves according to the set of equation similar to (4) but with zero values of V_1, V_2 and Δ:

$$\dot{\hat{\rho}}(v_z, t) = \hat{\hat{A}}'\hat{\rho}(v_z, t). \tag{6}$$

B. Motion of atom inside the cell

In this work we consider the cylindrical cell illuminated by the co-axial cylindrical laser beam. We suppose that the laser intensity is distributed homogeneously over the cross-section of the beam. So we can subdivide all the cell volume by "beam zone" inside the beam and "dark zone" outside the beam. In beam zone the density matrix describing the internal state of certain atom is governed by the equation (4) whereas in dark zone it is governed by equation (6). Let us consider an atom moving by the trajectory showed at Figure 1 (b). Suppose that at observation time t_A atom is in the point A in the beam. During time T between t_n and t_A the density matrix of this atom is governed by equation (4). The solution is:

$$\hat{\rho}(t_A) = \left(e^{\hat{\hat{A}} \cdot T} - 1\right)\hat{\hat{A}}^{-1}\hat{B} + e^{\hat{\hat{A}} \cdot T}\,\hat{\rho}(t_n). \tag{7}$$

In one's turn, during time τ'_{n-1} between t_{n-1} and t_n (see Figure 1(c)) the density matrix evolution is governed by (6).

$$\hat{\rho}(t_n) = e^{\hat{\hat{A}}' \cdot \tau'_{n-1}}\,\hat{\rho}(t_{n-1}), \tag{8}$$

et.c. Therefore we obtain an infinite sequence. After averaging over atomic velocities and times spent in beam and dark zones and summation one obtain [7]:

$$\begin{aligned} \rho_{33} = \overline{D} + \overline{\hat{U}^T (e^{\hat{\hat{A}} \cdot T} - 1)\hat{\hat{A}}^{-1}\hat{B}} + \overline{\hat{U}^T e^{\hat{\hat{A}} \cdot T}} \times \\ \overline{e^{\hat{\hat{A}}' \cdot \tau'}}\left(1 - \overline{e^{\hat{\hat{A}} \cdot \tau}}\,\overline{e^{\hat{\hat{A}}' \cdot \tau'}}\right)^{-1} \overline{\left(e^{\hat{\hat{A}} \cdot \tau} - 1\right)\hat{\hat{A}}^{-1}\hat{B}}. \end{aligned} \tag{9}$$

The averaging procedure will be given below.

If the characteristic velocity of the density matrix evolution is small in comparison with transition rate between illuminated and dark zone, i.e., if

$$\overline{\tau} \cdot (W + \Gamma) << 1, \quad \overline{\tau}' \cdot \Gamma << 1, \tag{10}$$

where $\overline{\tau}$ and $\overline{\tau}'$ are the average times what atom spent in beam and dark zones continuously, we can expand the exponents in (9) linearly. Keeping the zero power term of $\tau \cdot \hat{\hat{A}}$, $\tau' \cdot \hat{\hat{A}}'$ we obtain:

$$\rho_{33} = \overline{D} - \overline{\hat{U}^T} \cdot \left(\overline{\hat{\hat{A}}'\tau'} + \overline{\hat{\hat{A}}\tau}\right)^{-1} \cdot \overline{\hat{B}\tau}. \tag{11}$$

This is an expression for ρ_{33} in the beam zone whereas $\rho_{33} = 0$ in the dark zone. Averaging ρ_{33} over the atomic trajectory or, the same, over the cell volume (this averaging is denoted by brackets), we obtain:

$$\langle \rho_{33} \rangle = \frac{\overline{\tau}}{\overline{\tau} + \overline{\tau}'} \left(\overline{D} - \overline{\hat{U}^T} \left(\overline{\hat{\hat{A}}} + \frac{\overline{\tau}'}{\overline{\tau}} \overline{\hat{\hat{A}}'} \right)^{-1} \overline{\hat{B}} \right). \quad (12)$$

Let us consider an expression (15) more detailed. It is easy to see from (2) — (5) that $\overline{D}$, $\overline{\hat{U}}$ and $\overline{\hat{B}}$ are proportional to the squares of Rabi frequencies V_1 and V_2 and therefore $\overline{D}$, $\overline{\hat{U}}$ and $\overline{\hat{B}}$ linear over the radiation intensity. The matrix $\hat{\hat{A}}$ can be written as $\hat{\hat{A}} = \hat{\hat{A}}' + \hat{\hat{A}}''$, where $\hat{\hat{A}}'$ is independent, and $\hat{\hat{A}}''$ is linear on the radiation intensity. It allows us to rewrite (12) as:

$$\langle \rho_{33} \rangle = \langle \overline{D} \rangle - \left\langle \overline{\hat{U}^T} \right\rangle \left\langle \overline{\hat{\hat{A}}} \right\rangle^{-1} \langle \overline{B} \rangle. \quad (13)$$

Therefore if (11) is fulfilled, the CPT resonance parameters are determined by the total power by the whole cross-section of the cell independently on the laser beam diameter.

It should be noted that the calculation method is applicable to the homogeneously illuminated wall-coated cell with presence of buffer gas. In this case we should put $\overline{\tau}' = 0$ in (11) and threat $\overline{\tau}$ as average time between collisions of active and buffer atoms. It is easy to see that increasing of the buffer gas pressure leads to decreasing of $\overline{\tau}$. It leads to meet the condition (11) of applicability of expression (13).

C. Averaging

Here we describe the details of numerical calculation of (9). Let us denote the eigenvectors and eigenvalues of $\hat{\hat{A}}$ as $\hat{X}^{(i)}$ and λ_i respectively:

$$\hat{\hat{A}} \hat{X}^{(i)} = \lambda_i \hat{X}^{(i)}. \quad (14)$$

Let us introduce the matrixes $\hat{\hat{X}}$ and $\hat{\hat{\Lambda}}$:

$$\hat{\hat{X}} = \left(\hat{X}^{(1)}, \hat{X}^{(2)}, \hat{X}^{(3)} \right), \quad (15)$$

$$\hat{\hat{\Lambda}} = \hat{\hat{X}}^{-1} \hat{\hat{A}} \hat{\hat{X}} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \quad (16)$$

Similarly we build the matrixes $\hat{\hat{X}}'$ and $\hat{\hat{\Lambda}}'$ by the eigenvectors and eigenvalues of $\hat{\hat{A}}'$. Vector $\hat{B}$ can be easily decomposed by $\hat{X}^{(i)}$:

$$\hat{B} = \sum_i \hat{X}^{(i)} b_i = \hat{\hat{X}} \hat{b}, \text{ where } \hat{b} = \hat{\hat{X}}^{-1} \hat{B}. \quad (17)$$

Expressions like $e^{\hat{\hat{A}} \cdot t}$ in (11) can be written as

$$e^{\hat{\hat{A}} \cdot t} = \hat{\hat{X}} e^{\hat{\hat{\Lambda}} \cdot t} \hat{\hat{X}}^{-1} = \hat{\hat{X}} \begin{pmatrix} e^{\lambda_1 t} & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 \\ 0 & 0 & e^{\lambda_3 t} \end{pmatrix} \hat{\hat{X}}^{-1}. \quad (18)$$

Using (14) — (18) we can express (9) as

$$\begin{aligned} \rho_{33} = \overline{D} + \overline{\hat{U}^T \hat{\hat{X}} (e^{\hat{\hat{\Lambda}} \cdot T} - 1) \hat{\hat{\Lambda}}^{-1} \hat{b}} + \overline{\hat{U}^T \hat{\hat{X}} e^{\hat{\hat{\Lambda}} \cdot T} \hat{\hat{X}}^{-1}} \\ \times \overline{\hat{\hat{X}}' e^{\hat{\hat{\Lambda}}' \cdot \tau'} \hat{\hat{X}}'^{-1}} \left(1 - \overline{\hat{\hat{X}} e^{\hat{\hat{\Lambda}} \cdot \tau} \hat{\hat{X}}^{-1}} \; \overline{\hat{\hat{X}}' e^{\hat{\hat{\Lambda}}' \cdot \tau'} \hat{\hat{X}}^{-1}} \right)^{-1} \\ \times \overline{\hat{\hat{X}} \left(e^{\hat{\hat{\Lambda}} \cdot \tau} - 1 \right) \hat{\hat{\Lambda}}^{-1} \hat{b}}. \end{aligned} \quad (19)$$

Here the overline denotes the averaging over v_z, T, τ and τ'. As an example, let us consider an expression

$$\begin{aligned} \overline{\hat{\hat{X}} e^{\hat{\hat{\Lambda}} \cdot \tau} \hat{\hat{X}}^{-1}} = \int_{-\infty}^{+\infty} M_1(v_z) \times \\ \int_0^{+\infty} \hat{\hat{X}}(v_z) e^{\hat{\hat{\Lambda}}(v_z) \cdot \tau} \hat{\hat{X}}^{-1}(v_z) f(\tau) \, d\tau \, dv_z. \end{aligned} \quad (20)$$

Here $f(\tau)$ is the probability density function of random value τ, $M_1(v_z) = (\pi v_T)^{-1/2} \exp\left(-v_z^2 / v_T^2\right)$ is the one-dimensional Maxwell distribution function for v_z, $v_T = \sqrt{2 k_B T / m}$ is the most probable velocity, k_B — is a Boltzmann constant, T is a temperature, m is atomic mass. It is easy to see that the dependence on τ is appeared only in arguments of $e^{\hat{\hat{\Lambda}}(v_z) \cdot \tau}$ and is reduced to the form of $e^{\lambda' \tau'}$. Therefore it is convinient to introduce the function

$$g(\lambda) = \int_0^{\infty} \exp(\lambda \tau) f(\tau) d\tau \quad (21)$$

and to write down the *il*-th matrix element of (20) as

$$\begin{aligned} \left(\overline{\hat{\hat{X}}' e^{\hat{\hat{\Lambda}}' \cdot \tau'} \hat{\hat{X}}'^{-1}} \right)_{il} = \sum_k \int_{-\infty}^{+\infty} M_1(v_z) \times \\ X'_{ik}(v_z) \cdot g'\left(\lambda'_k(v_z) \right) \cdot \hat{\hat{X}}^{-1}_{kl}(v_z) dv_z. \end{aligned} \quad (22)$$

This expression contains only one integral over v_z.

Now let us find the probability distribution functions $f(\tau)$, $f'(\tau')$ and $f_T(T)$. The time τ what atom spent in the beam is

$$\tau = \frac{2\sqrt{r^2 - R^2 \sin^2 \varphi}}{v_\perp}, \tag{23}$$

where φ is an angle between the atomic velocity projection $\vec{v}_\perp$ on the plane orthogonal to the cell axis and the vector normal to the cell vall, R and r are the cell and beam radius respectively, $v_\perp = |\vec{v}_\perp|$. The probability density function of $v_\perp$ is a two-dimensional Maxwell distribution function $M_2(v_\perp) = 2\exp\left(-v_\perp^2 / v_T^2\right) v_\perp / v_T^2$. The probability density function $\Phi(\varphi)$ of random value φ is proportional to $\cos\varphi$ [8]. Note that the value of φ varies from 0 to $\arcsin(r/R)$ if atom passes the beam. Therefore $\Phi(\varphi) = R\cos\varphi / r$. It gives us the cumulative distribution function $F_\tau(\tau)$ of random value τ as

$$\begin{aligned} F_\tau(\tau) &= \int_0^{\arcsin(r/R)} \Phi(\varphi) \int_{\frac{2\sqrt{r^2 - R^2\sin^2\varphi}}{\tau}}^{\infty} M_2(v_\perp) dv_\perp d\varphi \\ &= \frac{\tau v_T \sqrt{\pi}}{4ir} e^{-\frac{4r^2}{\tau^2 v_T^2}} \operatorname{erf}\left(\frac{2ir}{\tau v_T}\right), \end{aligned} \tag{24}$$

what gives us $f(\tau)$ as

$$f_\tau(\tau) = \frac{dF_\tau(\tau)}{d\tau} = -\frac{1}{\tau} + \frac{v_T\sqrt{\pi}}{4ir} e^{-\frac{4r^2}{\tau^2 v_T^2}} \operatorname{erf}\left(\frac{2ir}{\tau v_T}\right)\left[1 + \frac{8r^2}{\tau^2 v_T^2}\right]. \tag{25}$$

The average time $\overline{\tau}$ what atom spend in the beam continuously is

$$\overline{\tau} = \int_0^\infty f(\tau)\,\tau\, d\tau = \int_0^\infty \left(1 - F(\tau)\right) d\tau = \frac{\pi^{3/2}}{2}\frac{r}{v_T}. \tag{26}$$

Now let us find the probability density function f_t of time t what atom continuously spent in the beam before the observation time t_A. Let dn/dt is a number of atoms coming into the beam per unit of time. Total number of atoms in the beam is:

$$N_{tot} = \int_0^\infty \left(1 - F_\tau(t')\right)\frac{dn}{dt'} dt', \tag{27}$$

(because $1 - F_\tau(t')$ is the probability for atom to survive in the beam during time t'). Therefore the number of atoms what already spent time t or more is

$$N_{t+} = \int_t^\infty \left(1 - F_\tau(t')\right)\frac{dn}{dt'} dt'. \tag{28}$$

It gives us the cumulative distribution function $F_t(t)$:

$$F_t(t) = 1 - \frac{\int_t^\infty \left(1 - F_\tau(t')\right)\frac{dn}{dt'} dt'}{\int_0^\infty \left(1 - F_\tau(t')\right)\frac{dn}{dt'} dt'}, \tag{29}$$

and the probability density function f_t:

$$f_t(t) = \frac{1 - F_\tau(t)}{\int_0^\infty \left(1 - F_\tau(t')\right) dt'} = \frac{1 - F_\tau(t)}{\overline{\tau}}. \tag{30}$$

It gives:

$$g_\tau(\lambda) = 1 + \lambda\,\overline{\tau}\, g_t(\lambda). \tag{31}$$

We should note that the function $g_t(\lambda) = \int_0^\infty e^{\lambda x} \cdot f_t(t)\, dt$ can not be expressed via elementary functions. So we approximated the exact expression (30) by

$$\begin{aligned} f_t(t) &\to \frac{v_T}{r}\tilde{f}_t\left(t\frac{v_T}{r}\right), \\ \tilde{f}_t(x) &= \frac{2}{\pi^{3/2}} e^{-bx} + \frac{16\left(1 - e^{-ax}\right)^3}{3\pi^{3/2} x^2} + A x e^{-cx}. \end{aligned} \tag{32}$$

The coefficients $A = 0.70839$, $a = 0.49$, $b = 2.286$ and $c = 1.272$ were chosen so as to functions $f_t(t)$ and $\tilde{f}_t\left(t\, v_T / r\right) v_T / r$ have the same asymptotes at infinity coincides near zero to a second-order terms by t inclusive, and that the function $\tilde{f}_t(x)$ has the correct norm $\int_0^\infty \tilde{f}_t(x)\, dx = 1$. It gives us

$$\begin{aligned} g_t(\lambda) &= \tilde{g}_t(\lambda r / v_T), \\ \tilde{g}_t(\Lambda) &= \frac{2}{\pi^{3/2}}\frac{1}{b - \lambda} + \frac{A}{(c - \lambda)^2} + \frac{16}{3\pi^{3/2}}\big(p(-\Lambda) \\ &\quad - 3p(a - \Lambda) + 3p(2a - \Lambda) - p(3a - \Lambda)\big), \end{aligned} \tag{33}$$

where $p(\beta) = \beta \ln \beta$. It is easy to find $g_\tau(\lambda)$ using (33) and (31).

At last we must found the probability density function $f'(\tau')$ for time τ' what atom spent in dark zone before go to beam. There are two different possibilities for atom when it is collide with the cell wall: it can collide elastically, or stick to the wall and after fly out in arbitrary direction. If atom is sticking to the wall after collision, we should note that the probability for atom to reach the highlighted zone between some pair of consecutive wall collisions is much less than to move through the cell in dark zone. In any wall collision act atom obtain some random value of velocity, and the probability to reach the highlighted zone is not depend on the number of collisions what atom experienced being in dark zone continuously. Therefore the probability density function $f'(\tau')$ can be approximated by the exponential distribution:

$$f'(\tau') = \frac{1}{\overline{\tau}'} \exp\left[-\frac{\tau'}{\overline{\tau}'}\right], \tag{34}$$

what gives

$$g'(\lambda') = \frac{1}{1-\lambda'\overline{\tau}'}. \tag{35}$$

Average time $\overline{\tau}'$ what atom spent in dark zone continuously can be calculated as

$$\overline{\tau}' = \overline{\tau}\frac{R^2 - r^2}{r^2}. \tag{36}$$

III. Results and Discussion

We made the simulations of the CPT resonance line shape and dark resonance parameters in Λ-atomic vapors in cylindrical cell with $R = 0.5$ cm. Atomic mass Macca атомов m was taken to be equal of ^{87}Rb mass, $T = 20$ °C, ground state relaxation rate $\Gamma = 300$ s^{-1}, optical coherence relaxation rate $\gamma' = 1{,}8 \cdot 10^7$ s^{-1}, $V_1 = V_2 = V$, $\Omega_L = 0$. Beam radius r was a variable parameter. We took the optical pumping rate $\overline{W}$ averaged over v_z as a main characteristic of laser intensity:

$$\overline{W} = \frac{2V^2}{\gamma'}\overline{G}, \quad \overline{G} = \int_{-\infty}^{+\infty} M_1(v_z)\, G(v_z)\, dv_z. \tag{37}$$

In our case $\overline{G} \approx 0{,}0168$, and $V = 2{,}3 \cdot 10^4 \sqrt{\overline{W}}$, where both V and $\overline{W}$ are measured in inverse seconds ($2\pi \cdot 1$ s^{-1} = 1 Hz).

It is convenient to introduce the value of

$$\left\langle \overline{W} \right\rangle = \frac{\overline{W} \cdot r^2}{R^2} = \frac{\overline{W} \cdot \overline{\tau}}{\overline{\tau}' + \overline{\tau}}. \tag{38}$$

i.e. the optical pumping rate averaged both over v_z and the cell volume. The value of $\left\langle \overline{W} \right\rangle$ is determined by the ratio of the laser power to the cell cross-section square.

It should be noted that the main contribution into the CPT resonance formation comes from the slow atoms whose velocity projection v_z is less or of order of $\gamma' \cdot \omega / c$. Therefore it is useful to introduce the characteristic optical pumping rate for slow atoms:

$$W_{slow} = V^2 / \gamma' = \overline{W} / (2\overline{G}). \tag{39}$$

Typical dark resonance line shape calculated according eq. (9) is presented on Figure 3. The resonance has characteristic "dual" structure: it consists on the broad "pedestal" with several tenth of kHz width, and a narrow central peak whose width is about 100 Hz. The broad peak is formed by atoms whose passed the beam once whereas the narrow peak is formed by multiply passage. This result is in agreement with experimental data presented in [3,5,6,8].

We studied the dependence of the central peak parameters on the laser power and beam diameter and verified the applicability of eq. (13). The $\left\langle \overline{W} \right\rangle$- dependences of central peak amplitude and FWHM are presented on Figure 3 and Figure 4 respectively. These dependences were calculated according (13) (solid line) and according (9) for various values of r. It should be outlined that $\left\langle \overline{W} \right\rangle$ is proportional to the laser intensity what is inverse proportional to r^2 at fixed laser power. Therefore the narrowing of the beam (at fixed power) leads to the narrowing and decreasing of amplitude of the CPT resonance.

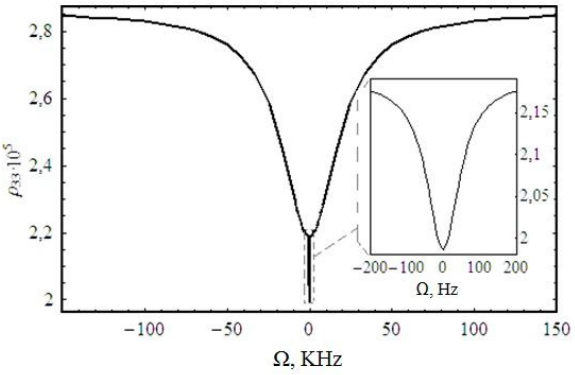


Figure 2. Dual structure of the dark resonance for $\overline{W} = 10^3$ s^{-1} and $r = 1$ mm.

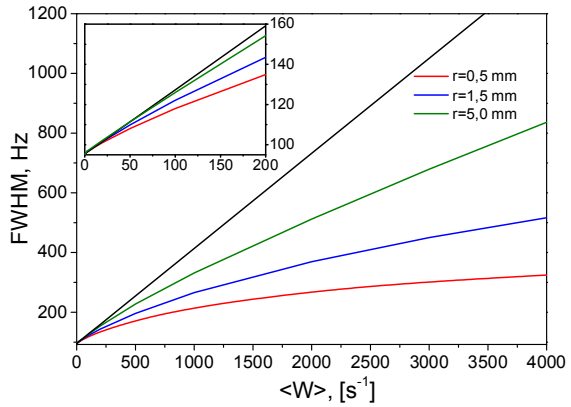


Figure 3. Dependence of the CPT resonance line width on $\langle \overline{W} \rangle$. The resonance line shape was calculated according (13) (solid line), and (9) (short dash, dot and dash) for various values of r.

Let us consider now the applicability of eq. (13). Characteristic values of $\overline{\tau}$ are: $\overline{\tau} = 5.9 \cdot 10^{-6}$ s for $r = 0.5$ mm, $\overline{\tau} = 1.75 \cdot 10^{-5}$ s for $r = 1.5$ mm, and $\overline{\tau} = 5.03 \cdot 10^{-5}$ s for $r = 5$ mm (in the last case the time $\overline{\tau}$ is not the time what atom spent in the beam continuously but the time between velocity-changing collisions). Consider the conditions (10) of applicability of eq. (13). In our case $\Gamma = 300$ s^{-1}, and even at $r = 0.5$ mm the value of $\overline{\tau}' = 5,8 \cdot 10^{-4}$ s, the second condition of (10) is fulfilled. The first condition of (13) reads as $\overline{\tau} \cdot W_{slow} << 1$ for slow atoms. It gives the following bounds for $\langle \overline{W} \rangle$: $\langle \overline{W} \rangle << 56$ s^{-1} for $r = 0,5$ mm, $\langle \overline{W} \rangle << 170$ s^{-1} for $r = 1,5$ mm, and $\langle \overline{W} \rangle << 660$ s^{-1} for $r = 5$ mm. Figure 4 shows that the resonance linewidth calculated according (13) and (16) are noticeably differs from these values of $\langle \overline{W} \rangle$.

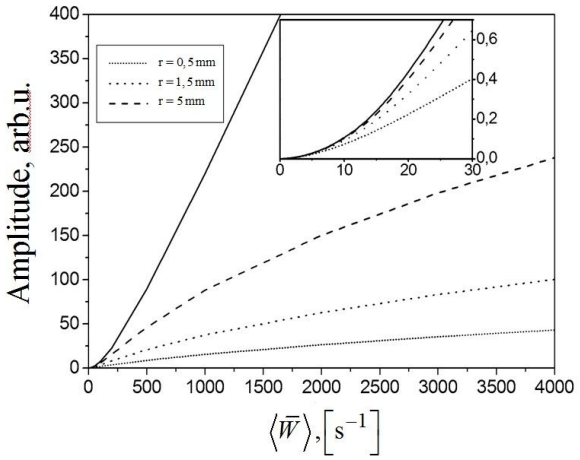


Figure 4. Dependence of the CPT resonance amplitude on $\langle \overline{W} \rangle$. The resonance line shape was calculated according (13) (solid line), and (9) (short dash, dot and dash) for various values of r.

The dependence of central peak amplitude on $\langle \overline{W} \rangle$ is presented on Figure 5. As for FWHM, the formula (13) is applicable if conditions (10) are fulfilled. But now these conditions reads strongly, i.e., $W_{slow} \cdot \overline{\tau} < 0.1$. It should be noted that even if the laser beam illuminate the whole cell volume, the CPT resonance amplitude doesn't reach the values predicted by formula (13). To increase the amplitude it is necessary to increase the velocity-changing collisions rate. It can be attained by adding some amount of buffer gas into wall-coated cell.

IV. Conclusion

In present work we present the theory of CPT resonance formation in partially illuminated wall-coated cell without buffer gas. This method is applicable also to the wall-coated cell with buffer gas below the diffusion approximation density limit, if these cells are illuminated by the laser light homogeneously. We obtain a simple calculation formula (13) what can be used if conditions (10) are fulfilled. It is found that if conditions (10) are broken, both amplitude and linewidth decreases with decreasing of the laser beam diameter, but amplitude decreases more significantly.

Acknowledgment

This work is supported by the grant of President of Russian Federation for young candidates of sciences MK-5318.2010.2, by Russian Foundation for Basic Research (11-02-90426-Ukr_f_a), the Russian Federal Special-purpose Program "Science and Science-pedagogical personnel of Innovation Russia on 2009-2013", by SNF (200020-105624), Association Suisse de la Recherche Horlogère (ASRH). We thank prof. G. Mileti for useful discussion.

References

[1] N.F. Ramsey, "A molecular beam resonanse method with separated oscillating fields," Phys. Rev., vol. 78, pp. 695–699, 1950.

[2] F. Riehle, Frequency Standards: Basics and Applications, Wiley-VCH, New York/Weinheim, 2004.

[3] Klein, I. Novikova, D.F. Phillips, R.L. Walsworth, “Slow light in paraffin-coated Rb vapor cells,”. Journal of Modern Optics, vol. 53, pp. 2583–2591, 2006

[4] Y. Xiao, I. Novikova, D. F. Phillips, R. L. Walsworth, “Diffusion-Induced Ramsey Narrowing,” Phys. Rev. Lett., vol. 96, p. 043601, 2006.

[5] Y. Xiao, I. Novikova, D.F. Phillips, R.L Walsworth, “Repeated interaction model for diffusion-induced Ramsey narrowing,” Optics Express, vol. 16, pp. 14128-14141, 2008.

[6] E. Breschi, G. Kazakov, C. Schori, G. Di Domenico, G. Mileti, A. Litvinov, and B. Matisov, “Light effects in the atomic-motion-induced Ramsey narrowing of dark resonances in wall-coated cells,” Phys. Rev. A, vol. 82, p. 063810 (2010)

[7] G.A. Kazakov, B.G. Matisov, A.N. Litvinov, “Narroving of the coherent population trapping resonance under the zonal pumping in coated cells,” (in Russian), Nauchno-Tekhnicheskie vedomosti SPbSPU, vol. 109, pp. 11-20 (2010)

[8] M. Klein, M. Hohensee, D. F. Phillips, and R.L. Walsworth, “Electromagnetically induced transparency in paraffin-coated vapor cells,” Phys. Rev. A, vol. 83, p. 013826 (2011)

[9] **Dicke, R.H.** The Effect of Collisions upon the Doppler Width of Spectral Lines // Phys. Rev. Vol. 89. – 1953. – P. 472–473

A Digital Servo with Single Frequency Modulation for Passive Hydrogen Maser

YuYing LI, TieXin LIU, ChuanFu LIN, JiaYu DAI, Yong ZHANG
Time & frequency research center
Shanghai Astronomical Observatory ,CAS
Shanghai, China
tiexinl@shao.ac.cn

YuYing LI
Graduate University
CAS
Beijing, China
daliyuying@gmail.com

Abstract: This paper introduces a digital servo using single frequency modulation for passive hydrogen maser. The main parts of this system include ADC, FPGA and DAC which sample error signal, process the signal and output two channels controlling voltage.

The algorithm is used to separate two errors by single frequency modulation principle. Then response curves of two channels are gotten to calculate proper PID controller parameters. In the end, two incremental PID controllers are used to get two channels control voltage to control the resonant cavity and OCXO respectively. The influence caused by the speed and precision of DAC to frequency stability has been analyzed. The experiments prove that the precision of DAC had a little influence on the short and medium term frequency stability at high control speed

The frequency stability of passive hydrogen maser with this system has been tested. It has achieved at $\sigma_y(1000)<3\times10^{-14}$.

I. INTRODUCTION

Shanghai Astronomical Observatory has devoted to development of hydrogen maser for a long period, since 2002 study on passive hydrogen maser had been carried out [1]. Since 2008, the electrical system digitization of signal-frequency passive hydrogen maser has been studying. The research of digital servo is a part of it, which Simplifies parameter regulation. And it has been achieved at $\sigma_y(1000)<3\times10^{-14}$.

II. DESIGN

In the physics package, magnetron cavity is used. The bulb is located in the center of cylindrical radiofrequency cavity. The magnetic shielding of the physics package has 4 layers and its shielding factor is about 50000.The passive Hydrogen maser is subdivided into two principal functional packages, the physics package which provides the actual atomic oscillator, and the electronic package which provides the atomic signal processing circuits and parameter control functions. The gain of physics package maser is about 3.2dB at -80dBm microwave signal input, and the H line-width is about 5Hz, as shown in fig1.

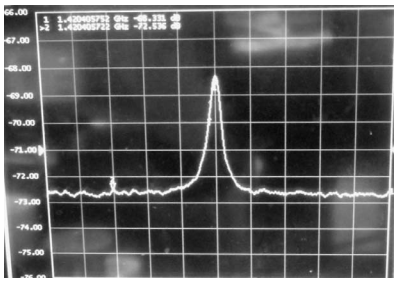


Figure 1. The parameters of the passive maser.

The main units of electronic package are detailed in fig2.These units are in charge of frequency locking of an OCXO (Oven Controlled Crystal Oscillator) and ACT (Automatic Cavity Turning) varactor diodes which compensates the frequency variation of the cavity.

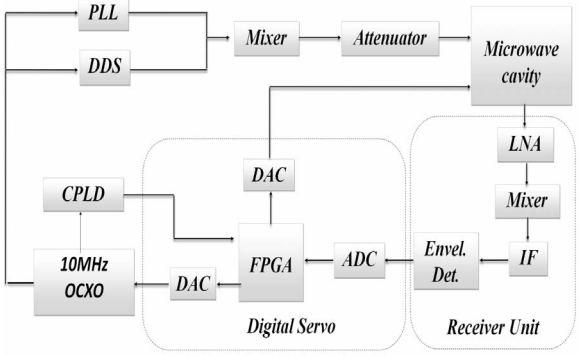


Figure 2. The functional block diagram of the passive H maser

Supported by:
National Natural Science Foundation of China, No. 110910778704

The 10MHz of the OCXO is directly multiplied by 140 to 1.4GHz, at the same time, it is multiplied to provide the clock of DDS to generate 20.405157Hz FSK modulated signal. The modulating signal is square wave frequency at 12.5 KHz provided by CPLD. The 1.4GHz and the modulated signal are mixed to allow double servo system for both the OCXO and ACT. After the interaction with the hydrogen atoms in microwave cavity, the output signal are first amplified by Low Noise Amplifier (LNA), then frequency mix with a free running oscillator to 20 MHz, and AM converted by envelope detector after Intermediate Frequency (IF) amplifier to generate the error signal at 12.5KHz. Finally the error signal is processed by the servo system together with Phase shifter signal from CPLD to get proper control voltage for OCXO and ACT varactor diodes.

FPGA maximizes the flexibility of the controller because it can be varied through programming means, without modification of the controller hardware. This paper introduces a digital servo system based on FPGA.

The main hardware structures of this system include ADC, FPGA and DAC. The block diagram of the digital servo system is shown in Fig. 3.

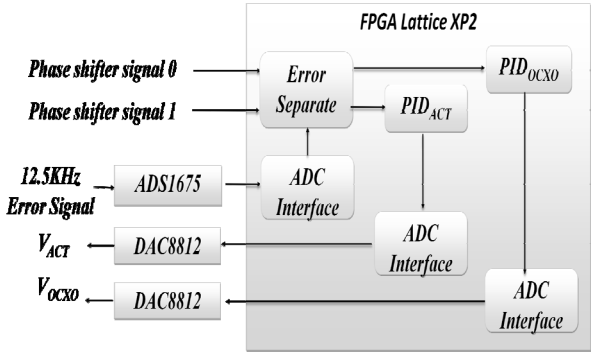


Figure 3. Block diagram of digital servo system

The ADS1675 samples the error signal form receiver unit at high sampling rate 1.28MHz. The ADC interface generates control signals in the FPGA that control the ADS1675 and transfer the 24 bit data to error separate module.

The error separate module separates two errors by single frequency modulation principle [2].

The two increment PID control systems take the errors and computes the correct signal based on the equation respectively [3].

$$\begin{aligned}\Delta u(k) &= u(k) - u(k-1) \\ &= P(\mathrm{e}(k) - \mathrm{e}(k-1)) \\ &+ I\mathrm{e}(k) + D(e(k) - 2e(k-1) + e(k-2))\end{aligned} \tag{1}$$

Δu is correction signal for OCXO and ACT varactor diodes; e is error; P, I and D are the proportional, derivative and integral parameters. PID parameters can be gotten from the response curves of two channels scanning process, which calculate proper PID controller parameters.

The two ADC interfaces generate necessary control signal in FPGA that control the DAC8812s and transfer the 16 bit data from the PIDs to DACs respectively. DAC8812s output DC voltage to control OCXO and ACT varactor diodes.

The state machine of the FPGA coding has 6 states, as shown in fig.4.

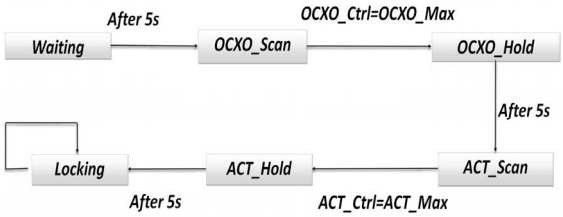


Figure 4. The state machine of the FPGA coding

At first waiting for 5s, In this state, the OCXO_Ctrl (The control variable of OCXO) set as the minimum, the ACT_Ctrl (The control variable of ACT varactor diodes) set as initial value. After 5s, get into OCXO_Scan state, in this state the OCXO_Ctrl increases to maximum.by degrees At the same time the maximum, minimum and square minimum error of the OCXO is ongoing record (fig. 5). In the COXO_Hold state, The Kp, Kd and Ki of PIDOCXO is calculated by response curve of OCXO_Scan , and set the OCXO_Ctrl at error square minimum point. The ACT_Ctrl has been also set as minimum. After 5s, get into ACT_Scan state, in this state, similar with OCXO_Scan to get response curve of ACT varactor diodes (fig. 6). In ACT_Hold state, The Kp, Kd and Ki of PIDACT is calculated by the results of ACT_Scan, and set the ACT_Ctrl at error square minimum point. The system will be in locking state after 5s. In Locking state, two channel of PID control the OCXO and ACT varactor diodes at its own speed.

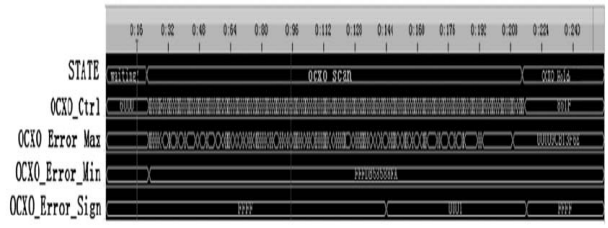


Figure 5. The OCXO_Scan state

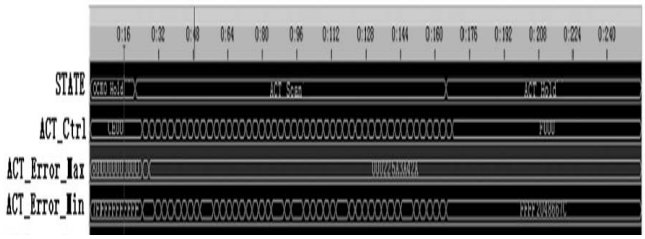


Figure 6. The ACT_Scan state

Form these two figure, we can see that the response curve of OCXO and ACT are linear within the scope of the scanning. The scope of the scanning is gotten from the experiment, and they will lock in this liner scope too.

III. VERIFICATION

Based on the above software and hardware, digital servo circuit is completed. In order to validate digital servo circuit, the frequency stability of the passive H maser with this digital servo is tested by PicoTime (from SpectraTime of Switzerland). The active H maser VCH-1003 provides the reference frequency. Tested results shown that frequency stability of the passive H-maser has been achieved at $\sigma_y(1000)<3\times10^{-14}$.

The experiments in the influence caused by the speed and precision of DA on frequency stability have been done for component selection.

The FPGA controls the DAC to imitate 14 bit, 13bit, 12bit DAC. The frequency stability has been tested with these different DAC at 7ms control speed. The experiments prove that DA had a little influence on the Short and medium term frequency stability at high control speed.

IV. CONCLUSION

This paper details the software and hardware design of a digital servo using single frequency modulation for passive hydrogen maser. Based on this servo system, good short-term stability is achieved. In order to further validate the effectiveness of the designed system, long-term stability will be tested with parameter regulation in future.

REFERENCES

[1] Xie Yonghui, Dai Jiayua, etc. Development of passive hydrogen maser in Shanghai Astronomical Observatory, 24th European Frequency and Time Forum, The NetherlandsApril 2010.

[2] G. Busca, H. Brandenberger. Passive H Maser, 33rd Annual Symposium on Frequency Control. 1979

[3] Zhang Y.S., Zhang J.S. Computer Control Technology . NanJing, Southeast University press. 2008. P62-76.

The Orientation Dependence of the SHF Radio-Optical Resonance Frequency Light Shift in Rubidium Vapors

Alexey Baranov, Sergey Ermak, Vladimir Semenov
Department of Quantum Electronics
St. Petersburg State Polytechnical University
St. Petersburg, Russia
lexusbar@gmail.com

***Abstract*— The results of the theoretical and experimental research of the light and orientation frequency shift effects in laboratory scheme of the ^{87}Rb cell frequency standard are provided. The real orientation frequency shift importance is shown.**

I. INTRODUCTION

The information on dynamics of behavior of so-called light shift has principal significance for utilization of a radio-optical resonance effect. With reference to vapors of the alkaline metals light shift as it is known contains two components [1]:

- Scalar shift $\Delta\nu_0$ which has similar values for all magnetic sublevels of the hyperfine structure F or F^*
- Tensor shift $\Delta\nu_T$ which is connected to the atom polarization dependence on the quantum magnetic number m_F

Light shift effect is a major destabilizing factor in the optical pumped quantum discriminators used in the rubidium vapor cell frequency standards [2]. In the classical scheme which suggests usage of isotopic filtering of the pumping light, the errors in these devices are largely determined by the scalar component of the light shift.

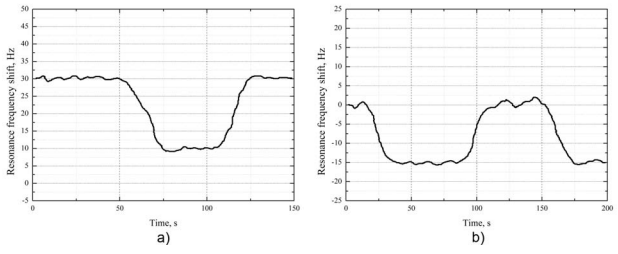


Figure 1. An example of the light(a) and orientation(b) shifts recording

In the case of optical pumping of ^{87}Rb vapor by nonresonant light of the second isotope contribution of the tensor component in the resulting light shift of the *0-0* transition does not exceed *15* and *2%* respectively for the pumping lines D_1 and D_2 [3]. As shown in [4], the situation changes radically if the optical pumping of ^{87}Rb vapor is provided by the spectral lamp with isotopic filtering of resonant light in a ^{85}Rb vapor cell. So, when pumping is provided by only D_1 spectral line $\Delta\nu_T$ values of magnetic-independent microwave transition can be significantly (several times) greater than the scalar frequency shift $\Delta\nu_0$. In this case, a characteristic feature of the light shift tensor component is its proportionality to the factor $1\text{-}3\cos^2\theta$, where θ is an angle between the direction of the working magnetic field and the pumping light beam [5].

The consequence of this dependence is ability to control the light shift in order to minimize it in the rubidium quantum discriminators used in the technique of frequency model measures [2]. This paper presents experimental results that confirm analytical studies [4] for various modes of spectrally selective optical pumping of ^{87}Rb vapors.

II. EXPERIMENT

The experiment was provided on a laboratory prototype of a quantum discriminator, which functional diagram contains the well-known set of composite components used in the gas cell frequency standards.

The discriminator was placed to the center of Helmholtz coils that create magnetic field in the working area of the absorption chamber with a working substance inside. To reduce external magnetic disturbances and irregularities of the laboratory magnetic field the entire system was placed in a dual-layer magnetic shield with a shielding factor about 10^3. Resonant microwave field that induces transitions in the hyperfine structure of rubidium atoms was produced by the ordinary scheme of frequency multiplication of the reference oscillator, followed by amplification. In the detection of the radio-optical resonance signal the standard synchronous detection technique and frequency servo-control of the reference oscillator to the resonance value were used.

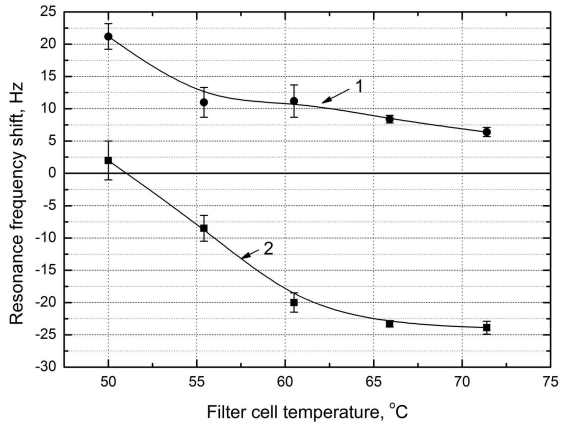


Figure 2. Experimental dependences of experimental dependences of the orientation (1) and light (2) shifts on the filter cell temperature for $I_{D1}=1,5I_{D2}$

Light-shift behavior was observed on the PC monitor at a fixed change in the pumping light intensity by a neutral and interference filters while changing the magnetic field direction angle θ from *0* to *90°* Fig. 1. The corresponding experimental dependences of the light and orientation shifts on the filter cell temperature in the case of optical pumping by the lamp source with spectral components intensity ratio D_1 and D_2 as 2 to 3 are shown on the Fig. 2.

The figure shows that there is no such filter cell temperature that provides a zero orientation shift for noticed above pump regime. This conclusion sorts well with the findings of the analytical studies [4] and illustrates the importance of the tensor component of the D_1 pumping line even in the form of a significant weakening of the D_2 line. To determine their optimum ratio at which the orientation shift of *0-0* transition frequency is minimized, we calculated the corresponding dependence, similar to how it is done in [4].

III. THEORY

The calculation was performed for the pump spectrum approximated by the sum of the Doppler contours with different weight contribution of the components D_1 and D_2 (Fig. 3 (a,b)) at a constant integrated intensity of the source (100μW/cm2). It is clear that requirements for the spectral component composition of the lamp source are not identical for minimization both light and orientation frequency shifts in the resonance transition. This situation causes to meet compromise requirements to maintain a high stability of the quantum discriminator's operating frequency.

In the case of laser pumping of one of the D_2 transition lines, orientation frequency shift is much higher than the value observed at the same integrated intensity of the spectral lamp source Fig. 4. In this case, attention is drawn to the fact that orientation shift takes place even at a zero light frequency shift that corresponds to the laser source tuned on the resonant frequency of the atomic transition.

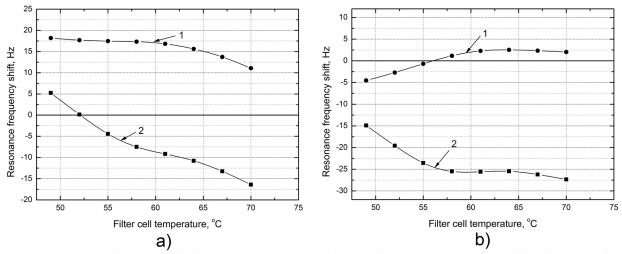


Figure 3. Calculated dependences of the orientation (1) and light (2) shifts on the filter cell temperature for different pumping modes: a) $I_{D1}=1,5I_{D2}$ b) $I_{D1}=10I_{D2}$

IV. CONCLUSION

Our experiments and calculations confirm an essential role of the tensor component for the orientation frequency shift of the *0-0* transition for both spectral lamp and laser pumping. In the experiment with lamp source interference filters usage can only reduce but not eliminate the influence of this component. However, it can significantly reduce the orientation shift of the *0-0* transition. For instance, ratio of intensities of the D_1 and D_2 lines as 1 to 20 leads to a tenfold weakening of the orientation shift in the filter cell temperature range 60-70 °C.

The results of these studies suggest that despite the relatively weak dependence of the frequency of the *0-0* transition on the absolute value of the external magnetic field, the requirement to ensure the constancy of the angle between the pumping light direction and the magnetic field vector is absolutely essential in quantum discriminators for moving and rotating applications.

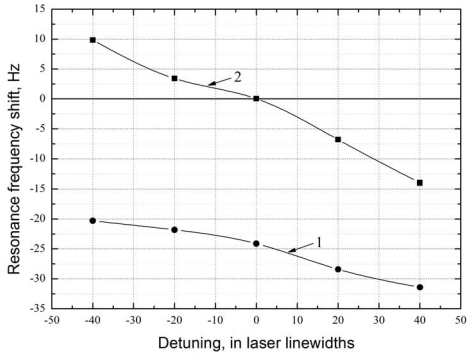


Figure 4. The theoretical dependence of the orientation (1) and light (2) shifts on the laser detuning from the center frequency of the D_1. Δf=1.5 MHz

REFERENCES

[1] W. Happer, B.S. Mathur, "Effective operator formalism in optical pumping" Phys. Rev., vol. 163, №1, 1967, pp. 12-25.

[2] Fritz Riehle, "Frequency standards. Basics and applications", WILEY-VCH Verlag GmbH Co. KGaA, 2005, p.526.

[3] B.S.Mathur, H.Tang, W.Happer "Light shifts in the alkali atoms," Phys. Rev. vol. 171, 1968, pp. 11–19.

[4] V.V. Semenov, "About the contribution of the radio-optical resonance tensor component light shift in rubidium vapors", Izvestia Vuzov, Physics, №2, 1999, pp. 86-90.

[5] E.B. Aleksandrov, G.I. Hvostenko, M.P. Chajka, "Interference of Atomic States", M.-Nauka, 1991

Most recent result of pulsed microwave hydrogen maser

Pengfei Chen, Yonghui Xie, Tiexin Liu,
Yuying Li, Chuanfu Lin
Time & frequency research center
SHAO
Shanghai, China
xyh@shao.ac.cn

Pengfei Chen, Yuying Li
Graduate University
CAS
Beijing, China

Abstract—Adequate coherent pulsed microwave input into microwave cavity of atomic clock can induce Ramsey fringes of the atomic transition, by which the line width of atomic clock can be greatly reduced and the performance might be improved. This method has been applied to the pulsed optically pumped （POP） rubidium clock.

Shanghai Astronomical Observatory has studied the application of this method in passive hydrogen maser. Electrical circuit for pulsed microwave frequency sweeping has been designed, and the Ramsey fringes of hydrogen atom have been observed, which means that this method for passive hydrogen maser is feasible. We plan to optimize the parameters of the system and microwave pulse to further reduce the line width and design full-scale servo circuit to realize the pulsed microwave hydrogen maser. Finally system performance will be measured and optimized. This work is original in China.

I. Introduction

For the property of high stability and narrow width, transition spectrum of specific atom can be utilized as reference signal to control crystal oscillator. This is the basic principle of atomic clock. Line width of the atomic transition spectrum is one of the most critical factors of the atomic clock [1]. All kinds of technologies are taken to ensure the narrow line width. In the passive rubidium clocks, glass cell filled with buffer gas is adopted to keep rubidium atoms in a rather limited space so as to reduce the Doppler broadening; in the hydrogen clocks, the atomic storage bulb technology is introduced, which greatly extends the atoms' storage time and their life time; in the caesium beam clocks, method of separated interaction regions works: caesium atoms pass two separated in-phase microwave fields when flying in the physical package which induces Ramsey interference whose line width at central fringe is greatly reduced.

Recently, a new method to induce Ramsey interference emerges. Successively importing two coherent pulsed microwave signals into the microwave cavity could simulate two coherent microwave fields which are separated in space. Ramsey interference could also be induced in this way. This method has been applied to the pulsed optically pumped (POP) rubidium clock successfully. Ramsey interference fringes have been observed and the POP rubidium clock has been locked on the central Ramsey interference fringe.

Principle of passive hydrogen maser is similar to rubidium clock. Atoms interact with detecting signal. Stimulated transition is induced and error signal is provided, by which voltage-controlled crystal oscillator is locked. And atomic signal gain of the passive hydrogen maser is about 3dB. So the method inducing Ramsey interference with pulsed microwave could also be applied to passive hydrogen maser. Time & frequency research center in Shanghai Astronomical Observatory has carried out research on pulsed microwave atomic clock based on conventional passive hydrogen maser.

II. Theory

When two coherent microwave pulses are imported into the passive hydrogen maser, the transition probability $P(\tau)$ of atom system is [2]

This research is supported by National Natural Science Foundation of China, No.11003037
Also supported by the Knowledge Innovation Program of the Chinese Academy of Sciences

$$P(\tau)=\frac{4b^2}{\Omega^2}\sin^2\frac{1}{2}\Omega\tau \left(\cos\frac{1}{2}\Omega\tau\cos\frac{1}{2}\Omega_0 T-\frac{\Omega_0}{\Omega}\sin\frac{1}{2}\Omega\tau\sin\frac{1}{2}\Omega_0 T\right)^2, \quad (1)$$

where b represents the Rabi angular frequency, Ω_0 is the angular frequency offset of imported microwave signal from atomic resonance frequency, T is the interval between two coherent pulses and Ω is defined as

$$\Omega^2=b^2+\Omega_0^2. \quad (2)$$

When the frequency of inducing microwave fits the atomic resonance frequency, which means $\Omega_0=0$, (1) and (2) give

$$P(\tau)=\sin^2 b\tau. \quad (3)$$

Equation (3) shows that, when $b\times\tau=\pi/2$, the transition probability $P(\tau)$ attains maximum. So if $\tau=10\times10^{-3}s$, we should set $b=157rad/s$. In the microwave cavity of hydrogen clock, b is determined by (4):

$$b=\mu_B\langle B_z\rangle_b/\hbar, \quad (4)$$

where μ_B is Bohr magneton, $\langle B_Z\rangle_b$ is the mean value, over the bulb volume, of the amplitude of the projection of the microwave magnetic induction on the direction z of the static magnetic field. And the microwave power P in the cavity is:

$$P=\varpi_0 V_b\langle B_z\rangle_b^2/2\mu_0 Q_C\eta', \quad (5)$$

where ϖ_0 is the hydrogen atom transition frequency, V_b is the volume of the atomic storage bulb, μ_0 is the permeability of vacuum, Q_C is the cavity quality factor and η' is the filling factor of the microwave cavity. From (4) and (5), the microwave power P in the cavity is given by

$$P=\varpi_0 V_b b^2\hbar^2/2\mu_0\mu_B^2 Q_C\eta'. \quad (6)$$

Assuming $b=157rad/s$, from (6) we have $P=0.91\times10^{-9}W$, about-$59.6dBm$. That is to say, if the width of microwave pulse is $10\times10^{-3}s$, the microwave power in the storage bulb of hydrogen clock should be $-59.6dBm$. Considering that the microwave cavity has energy attenuation of about $-13dB$, imported power of the microwave pulse should be about $-47dBm$.

III. Experimental Setup

To verify the applicability of inducing Ramsey interference with microwave pulse in passive hydrogen maser, we construct a set of test system to detect the Ramsey interference fringes of hydrogen atoms. Fig. 1 shows the schematic of the test system.

The physics package of microwave pulsed hydrogen clock is based on conventional passive hydrogen maser. The physics package contains electrode microwave cavity, whose loaded quality factor is about 10000, energy attenuation is about -13dB and filling factor is about 0.5. The atomic storage bulb has a diameter of 60×

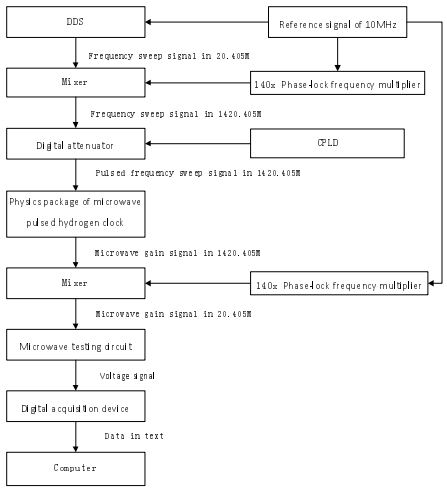


Fig. 1. Schematic setup of the test system for Ramsey fringes of hydrogen maser.

10^{-3}m, a length of 150×10^{-3}m and a volume of 420×10^{-3}L. Storage time of atoms is about 0.8s. The working current of 3×10^{3}V ion pump is 300×10^{-6}A. Relaxation time T1 and T2 of hydrogen atoms are about 200×10^{-3}s. There are 4 layers of magnetic shielding, whose shielding factor is about 50000.

In the test system, a reference signal of 10×10^{6}Hz is multiplied to 1.4×10^{9}Hz by a phase-lock frequency multiplier, then mixed with intermediate frequency signal of 20.405×10^{6}Hz generated by Direct Digital Synthesizer(DDS) to output microwave signal at 1.420405×10^{9}Hz. Pattern control signal generated by Complex Programmable Logic Device (CPLD) has a pulse duration $t_1=10\times10^{-3}$s, a typical pulse interval $t_p=100\times10^{-3}$s and a Ramsey time $T=t_1+t_p=110\times10^{-3}$s, which are shown in Fig. 2. It controls the digital attenuator which attenuates the 1.420405×10^{9}Hz microwave signal to obtain pulsed microwave signal. We should adjust the power of the microwave signal to ensure $b\times t_1=\pi/2$. The pulsed microwave signal is imported into the microwave cavity and interacts with hydrogen atoms, and then Ramsey interference could be induced. Frequency of the intermediate frequency signal is adjusted linearly to implement frequency sweep whose center frequency is $1.420405xxx\times10^{9}$Hz and width is 100Hz. Microwave signal in cavity is exported, mixed to about 20×10^{6}Hz and transformed to voltage signal by a microwave testing circuit composed of AD8307. The voltage signal is collected and stored by the digital acquisition device. Accumulating a period of the data, we can get the Ramsey fringes.

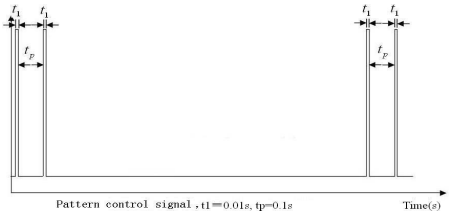


Fig. 2. Pattern control signal generated by CPLD.

IV.Most Recent Result

At present, we have gotten the decay signal of hydrogen atoms induced by the microwave pulses, shown in Fig. 3. Ramsey fringes could be obtained by collecting and processing the data in effective region, shown in Fig. 4. The central fringe with line width of about 1.5Hz and the side fringes are clear in Fig. 4.

The S curve could be acquired by differentiating the data of the Ramsey fringes and be used to lock the hydrogen clock in future, shown in Fig. 5. Changing the power of the pulsed microwave to make $b\times t_1$ deviate from $\pi/2$ will broaden the Ramsey fringes, shown in Fig. 6. This fit the result of the relevant experiments[3].

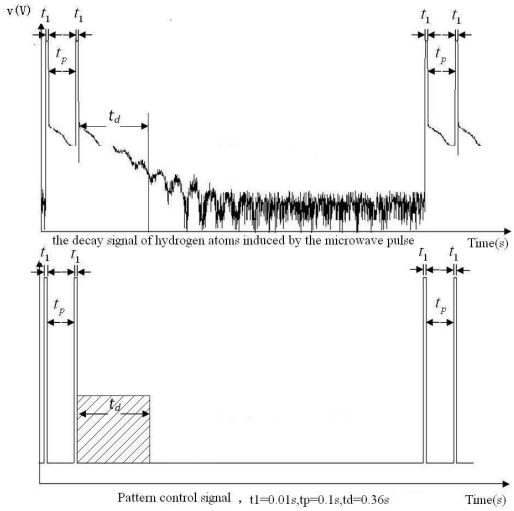


Fig. 3. Relaxation signal of hydrogen atoms stimulated by pulsed microwave signal.

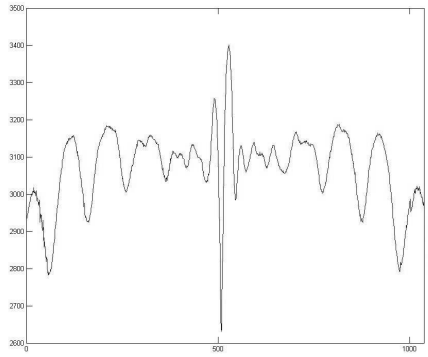


Fig. 4. Ramsey fringes of hydrogen maser.

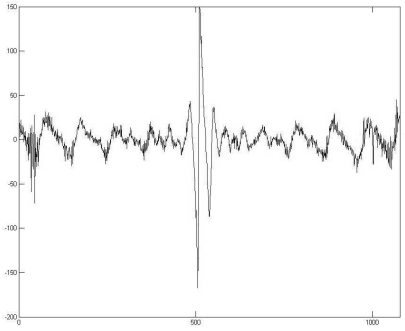


Fig. 5. S curve of pulsed microwave hydrogen maser.

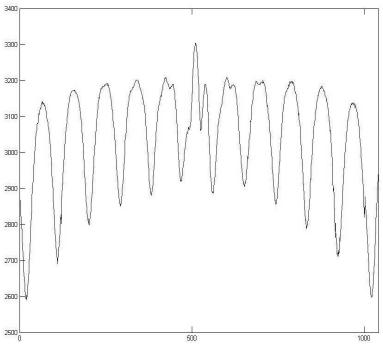


Fig. 6. The Ramsey fringes of hydrogen maser at deflected microwave pulse parameters.

Increasing the pulse interval t_p could compress the line width of the Ramsey fringes. Fig. 7 shows the contrast of the line widths of the Ramsey fringes in different intervals. And when $t_p=160\times10^{-3}$s, the line width is down to about 1.2Hz.

V. Discussion

Time & frequency research center in Shanghai Astronomical Observatory has carried out theoretical analysis and relative researches on pulsed microwave atomic clock based on conventional passive hydrogen maser and has constructed a set of test system. The Ramsey fringe was observed and its line width is about 1.2Hz, smaller than conventional passive hydrogen maser obviously. At present, we are optimizing the pulse parameters to compress the line width of the Ramsey fringes and to improve the Signal Noise Ratio (SNR) of the fringes as far as possible.

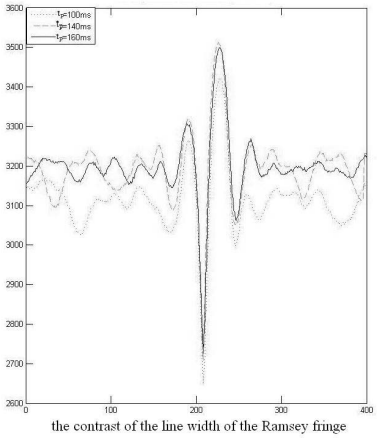


Fig. 7. The contrast of the widths of central wave crests of Ramsey fringes

In future we would deploy digital serving circuit and high stability voltage-controlled crystal oscillator to form a closed control loop based on the existing test system. Assisting with the auxiliary circuit such as temperature control system, C field system and dissociator, a whole pulsed microwave hydrogen clock would be built up. We will optimize the system parameters to obtain good electricity performance including frequency stability and frequency drift.

REFERENCES

[1] Y. Wang, Q. Wang, J. Fu, and T. Dong, The Principle of Quantum Frequency Standards. Beijing: Science Press, 1986, pp.11~12.

[2] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards, Vol. 1 (l). Bristol UK: Institute of Physics Publishing, 1989, pp.627~629.

[3] Aldo Godone, Salvatore Micalizio, and Filippo Levi, "Pulsed optically pumped frequency standard," Physical Review A: Atomic, Molecular, and Optical Physics, 2004, 70, 023409.

Proposal of Rubidium Atomic Beam Clock Based on Lamp Pumped and Fluorescence Detection

Y. H. Wang, Y. Gu, J. Q. Huang, S. Q. Liu, T. Q. Dong
School of Electronics Engineering and Computer Science
Peking University
Beijing, China
wangyanhui@pku.edu.cn

Z. H. Lu
School of Physics
Huazhong University of Science and Technology
Wuhan, China

***Abstract*—A compact, portable rubidium atomic beam clock based on lamp pumped and fluorescence detection scheme is proposed. The expected short-term frequency stability can be more than two orders of magnitude better than previous experimental results. The usages of lamp pumping, fluorescence detection and microwave slow-wave resonance structures make this design robust and compact. The shortcomings will be discussed too.**

I. INTRODUCTION

With the development of optical frequency standards based on cold atoms or ions, improved accuracy and stabilities have been achieved during the last few years. However, most of them are still running in scientific laboratories and require professional scientific researchers for day-to-day maintenance. The traditional microwave frequency standards, such as rubidium gas cell clock, cesium beam clock and hydrogen maser are still the most popular commercial frequency standards. Especially the most accurate cesium beam clock, like Agilent 5071A, is widely used in time keeping, satellite navigation and deep space exploration.

Here we propose a rubidium beam clock scheme similar as cesium beam clock to improve the performance. The initial idea and preliminary experimental results of an optically pumped rubidium beam clock was reported by Arditi *et al.* around thirty years ago [1]. Later, a detailed theoretical analysis of the operation was developed [2]. From the analysis, it was concluded that the performance limitation of this kind of beam clock was dominated by the poor signal to noise ratio of the optical detection signal.

With the development of the diode lasers, a few groups in the world started working on the laser-pumped cesium and rubidium beam clocks [3-7]. The main advantages are the higher utilization rates of the atoms and the more compact design. Experimental results also demonstrated the improved results. However, due to the frequency drift of the diode lasers, the long-term performance is not so great. This limits the further improvements and the marketing of the products. Up to now, as far as we know, there are still no commercial laser-pumped cesium or rubidium beam clocks available.

This work is supported by "the Fundamental Research Funds for the Central Universities".

II. OPERATION AS AN ATOMIC CLOCK

In ref. [8], we proposed the scheme based on lamp-pumping and fluorescence-detection. Here we briefly summarize the main changes in comparison with Arditi *et al.* design.

A. Lamp-pumping

In our scheme, we use ^{87}Rb lamp with ^{85}Rb filter cell to produce pumping light, which can improve the pumping rate of rubidium atoms by a factor of 2.

B. Fluorescence-detection

The fluorescence detection instead of transmission detection can provide higher signal to noise ratio. The improvement is around two orders, whose calculation can be seen in ref. [8].

C. Multiple photo- detectors

Since the time constant of the servo-loop can be in the range of second, multiple photo-detectors can be used to detect the fluorescence signal, and the photo-detectors' output can be combined to further improve the signal to ratio. Taking into account the actual available vacuum space, we estimate that a reasonable number for the detectors is around 10, which means the signal to noise ratio is improved by a factor of 3.

D. Microwave slow-wave structure

The microwave slow-wave structure is a widely used helical type, where the microwave field propagates along the wire of the helix. The other end of the helix is grounded. The microwave is completely reflected and a standing wave field forms inside the structure. For the Ramsey type beam clock the linewidth is inversely proportional to the distance between the two microwave interaction regions. For the microwave slow-wave structure the linewidth depends on the total effective length of the wire. This design has been widely used in traditional rubidium gas cell clocks, and is much more compact.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

E. Multi-reflecting mirrors design

Since the light intensity of rubidium lamp is quite low, the interaction length of light and atoms can be extended by multi-reflected to increase the fluorescence signal in the optical detection region, which is similar as the design of Arditi *et al.* The structure is shown is Fig. 1.

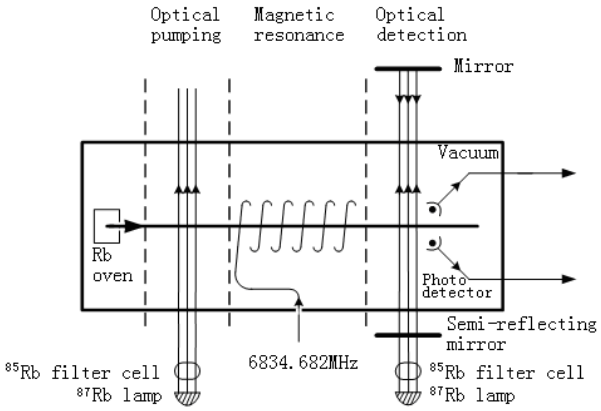


Fig.1. Structure of lamp pumping and fluorescence detection rubidium beam clock

III. DISCUSSION

The proposed scheme of a lamp-pumping fluorescence detection rubidium beam clock has the advantage of a high pumping efficiency, compact structure, light weight, and easy magnet field shielding in comparison with magnetic field selected type cesium beam clocks. It is also more practical than laser pumping beam clocks. The theoretical performance level potentially can be more than two orders of magnitude higher than the former design, and is comparable with commercial cesium beam clocks. However, consider the same temperature of oven (110℃) of rubidium and cesium, on one hand, the velocity of ^{133}Cs and ^{87}Rb are 247m/s and 305m/s, respectively. On the other hand, the saturated vapor pressure of rubidium is only one-third of cesium. In order to get the similar linewidth, we have to extend the length of microwave slow-wave structure, which should be around half meter and larger than that of cesium beam clock.

REFERENCES

[1] M. Arditi and P. Cerez, "Hyperfine Structure Separation of the Ground State of 87Rb Measured With an Optically Pumped Atomic Beam," IEEE Trans. Instrum. Meas. IM21(4), 391-395, 1972.

[2] P. Cerez and F. Hartmann, "Theoretical Analysis of the Opteration of an Optically Pumped Rubidium Atomic Beam Clock," IEEE J. Quantum Electr. QE13(5), 344-351, 1977.

[3] C. Sallot, M. Baldy, D Gin, and R. Petit, "$3 \cdot 10^{-12} \cdot \tau^{-1/2}$ on Industrial Prototype Optically Pumped Cesium Beam Frequency Standard," IEEE Freq. Control Symp., 100-104, 2003.

[4] K. Hagimoto, Y. Koga, and T. Ikegami, "Reevaluation of the Optically Pumped Cesium Frequency Standard NRLM-4 with an H-Bend Ring Cavity," IEEE Trans. Instrum. Meas. 57(10), 2212-2217, 2008.

[5] I. Pascaru, "Atomic frequency standard using optical pumping for state preparation and magnetic state selection of atoms." United States Patent 5107226, 1992.

[6] S. Lecomte, M. Haldimann, R. Ruffieux, P. Berthoud, and P. Thomann, "Performance demonstration of a compact, single optical frequency Cesium beam clock for space applications," IEEE Freq. Control Symp. 1127-1131, 2007.

[7] A. Besedina, A. Gevorkyan, G. Mileti, V. Zholnerov, and A. Bassevich, "Preliminary results of investigation of the high-stable Rubidium atomic beam frequency standard with laser pumping/detection for space application," Proc. 20th EFTF, 270-276, 2006.

[8] Y. H. Wang, J. Q. Huang, Y. Gu, S. Q. Liu, T. Q. Dong, Z. H. Lu, "Improved rubidium atomic beam clock based on lamp-pumping and fluorescence-detection scheme", Journal of the European Optical Society-Rapid Publication, 6, pp. 11005, 2011.

Integrating Sphere Cold Atom Clock with Cylindrical Microwave Cavity

Xucheng Wang, Huadong Cheng, Benchang Zheng, Yanling Meng, Ling Xiao, Liang Liu*, and Yuzhu Wang
Key Laboratory of Quantum Optics
Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences
Shanghai 201800, China
*email: liang.liu@siom.ac.cn

***Abstract*— The integrating sphere is used to generate the diffuse light for cooling atoms. The microwave field is produced in a cylindrical cavity on a TE011 mode. We have realized the microwave interrogation with the pulsed optically pumped (POP) scheme. The Ramsey fringe is demonstrated and some factors to affect the signal to noise ratio (SNR) and the contrast of the fringe are discussed. Optimization of the experimental parameters is still in progress.**

I. Introduction

In 1989, Chu's group built the first cold sodium atomic fountain frequency standard [1]. And then, many other groups built their own atomic fountain clocks. The cold atomic fountain clock has an outstanding performance, but its size is usually larger than 1 m^3, which restricts its applications in some aspects, such as for space application. In making a well performed compact cold atom clock, the integrating sphere has its special features. It uses the diffuse light cooling technique, which does not need any light beam collimation and power balance, and the atoms can be cooled, prepared, and microwave interrogated at the same place [2], [3]. This technique was first realized by Ketterle in 1992 [4]. Cooling cesium atoms directly from the background vapor with the diffuse light was realized in 2001 [5] and Rubidium in 2009 [6]. Recently, Esnault *et al.* have developed the HORACE Cesium clock [7], whose microwave cavity is spherical, and the short term frequency stability has reached to 2.2×10^{-13}, and the long term performance is around 3×10^{-15} in 10^4 s of integration.

Here we develop an integrating sphere rubidium cold atom clock with the POP scheme, which has characters of all optical cooling, simple structure with small volume, and low power consumption, and has potential features in space application. Different from the HORACE clock, we apply the cylindrical microwave cavity, which is easy to be fabricated and sophisticated. The cylindrical cavity has better structure of magnetic field, and we expect that the integrating sphere atom clock with cylindrical cavity will have a better performance.

In this paper, firstly, we will describe the experimental setup and the experimental procedure. Then we will demonstrate the Ramsey fringe and discuss some factors to affect its SNR and contrast.

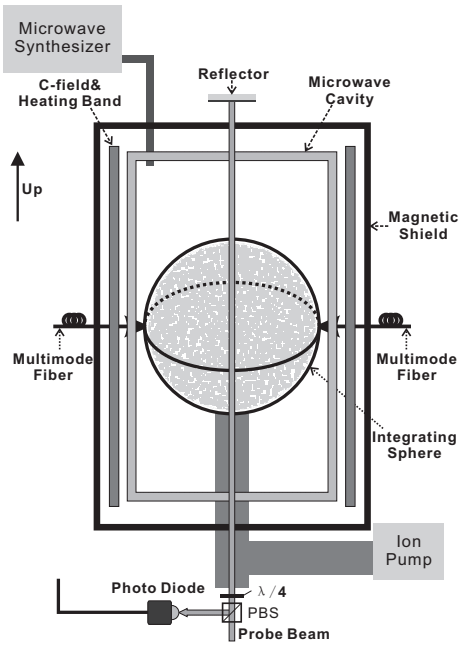


Fig. 1. The schematic of the experimental setup

II. Experimental setup

In our experiment, the integrating sphere is used to generate the diffuse light [6], and the diffuse reflection coefficient at the inner surface is about 98 % at 780 nm. The experimental setup is shown in Fig. 1. A glass cell that is used to contain the ^{87}Rb vapor atoms is located inside the sphere. A cylindrical microwave cavity on the TE011 mode, is tuned to be resonant at 6.835 GHz, and the quality factor of the cavity is about 10, 000. Outside of the cavity, the temperature control system is applied for the cavity to stabilize the resonant frequency. The μ - metal is used to shield the geomagnetic field. C - magnetic field is generated by a solenoid with a pair of compensating coils, and it is around 20 mG in our experiment. The vacuum is about 10^{-6} Pa, which is maintained by a 8 l/s ion pump. Quartz oscillator is the local oscillator (LO) and the 6.835 GHz signal is produced by the microwave synthesizer.

Here, an external cavity diode laser is divided into three

978-1-61284-110-6/11/$26.00 ©2011 IEEE

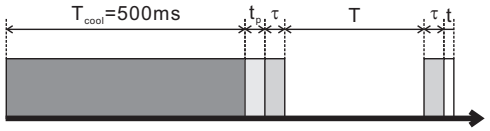


Fig. 2. The time sequence of experiment. T_{cool} - time of cooling light, t_p - time of pumping light, τ - time of the microwave pulse, T - separated time of the two microwave pulses, t - detecting time.

beams (cooling, pumping, and probe laser beams) whose frequency are shifted by three different acousto - optic modulators (AOM). The cooling, pumping, and repumping laser beams (repumping laser is generated by another external cavity diode laser) are combined and injected into the integrating sphere through two multimode fibers.

As shown in Fig. 1, the probe laser beam crosses the cold atomic cloud vertically, and it is retroreflected by a mirror to avoid heating the cold atoms. Finally, we use a photodiode to detect the absorption signal at the bottom. And the power of the probe beam is 0.7 μW, the radius is about 1 mm.

III. Cooling and Interrogation

In order to capture more cold atoms, the loading time of cold atoms in the integrating sphere is about 1 s [8]. In fact, the cycle frequency is another important factor affecting the instability of the clock. We can utilize the pulsed diffuse laser cooling method [9] to shorten the cooling time to 500 ms, and the dead time is 13.5 ms. Making further efforts to reduce the cycle time is necessary in the next work.

Figure 2 is the experimental sequence. Firstly, we turn on the cooling and repumping laser lights to cool the atoms. The cooling laser is detuned -3Γ (natural linewidth) of the transition of $5^2S_{1/2}$, $F = 2 \rightarrow 5^2P_{3/2}$, $F' = 3$. The total power of the cooling laser is about 100 mW, the repumping laser is 6 mW. 500 milliseconds later, about 10^9 atoms are captured successfully, and all of the cold atoms are populated in $F = 2$ states. Secondly, we apply a pumping laser light for $t_p = 1$ ms, whose frequency is a little red shift from the resonant frequency of $F = 2 \longrightarrow F' = 2$. Nearly all of the atoms can be pumped into the ground states of $F = 1$. Then, two $\pi/2$ microwave pulses are applied to make the atoms accomplish the clock transition of $F = 1, m_F = 0 \longrightarrow F = 2, m_F = 0$. The time of each microwave pulse duration is 1 ms, i. e. $\tau = 1$ ms, the separated time between the two interactions is $T = 10$ ms. Finally, after the clock transition, we turn on the probe beam for 0.5 ms to detect the atoms which are populated at the states of $F = 2$, $m_F = 0$, i. e. $t = 0.5$ ms. Using this sequence, there is about 30 % cold atoms experience the clock transition.

IV. Results and Discussion

With the above sequence, the Ramsey fringe can be determined when we sweep the microwave frequency around the resonance frequency, as shown in Fig. 3. The linewidth of the central fringe is about 45 Hz, which is determined by $T + \tau$ [10].

We average the experimental data over five measurements. The contrast is more than 70 %, and the SNR is about

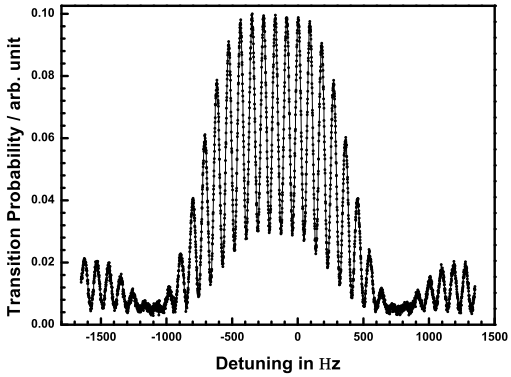


Fig. 3. The observed Ramsey fringe

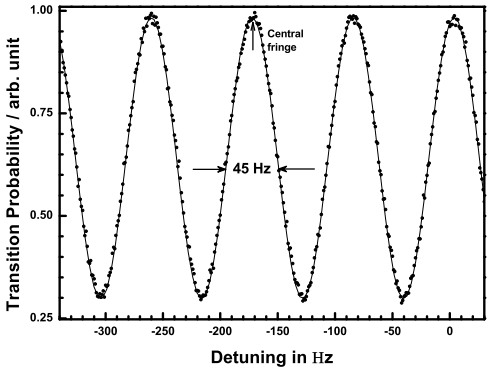


Fig. 4. In the vicinity of the resonance frequency for 10 ms Ramsey pulse separation time

100. Here there are some factors to affect the contrast. First, collisions from the background gas will break the coherence of the cold atoms. This effect can be reduced by improving the vacuum of the system. Second, a relaxation caused by the inhomogeneity of the magnetic field inside the magnetic shield is another factor in breaking the coherence. This is mainly caused by the imperfect of the magnetic shield, the C - field and even the stray magnetic field caused by the temperature control system. So we should put the temperature control system outside the magnetic shield in the future work. Third, the unbalance of the microwave pulses will also reduce the contrast of Ramsey fringe, which may be caused by the nonuniform of the microwave cavity along the probe beam direction and the power shift of the microwave synthesizer. In our system, the above three factors may be the main reasons for the losing contrast. SNR is mainly determined by the noise from the intensity and frequency of probe laser, and the noise from the electrical system. In addition, variation of the atomic population in $F = 2$ will also decrease the SNR.

Figure 4 is the Ramsey fringe in the vicinity of the resonance frequency, from which we can see that there is about

170 Hz shift from the resonance frequency. It may be caused by two main factors. First, the frequency deviation of the LO generates an obvious frequency shift. In the future work, we will use H - maser clock as the LO. Second, the residual magnetic field inside the magnetic shield, as mentioned above, will generate a second order Zeeman shift.

Now, we have locked the cold atom clock, but optimizing the performance of the stability is still under progressing.

V. Conclusion

We demonstrate the Ramsey fringe of the cold ^{87}Rb atoms in the integrating sphere with the cylindrical microwave cavity, and have discussed the factors to affect its contrast and SNR. Now, we are improving the vacuum of the system and performance of the magnetic shield and the homogeneity of the C - field.

Our aim is to make a compact cold atom clock for space application. In the micro - gravity environment, the loss of the cold atoms is only determined by the thermal diffusion [11] and the collisions from the background, so the lifetime of the cold atoms in space is much longer than on the earth, and the linewidth of the central fringe will be extremely narrowed. Besides that, we can also recapture the cold atoms much quicker in the micro - gravity environment, so the cooling time can be shortened. The integrating sphere cold atom clock will have better performance in space.

Acknowledgment

Supports from the National Natural Science Foundation of China under Grant No. 11034008, 11074262, and 10874193 are gratefully acknowledged.

References

[1] M. A. Kasevich, E. Riis, S. Chu, and R. G. DeVoe, "rf spectroscopy in an atomic fountain," *Phys. Rev. Lett.*, vol. 63, no. 6, pp. 612–615, Aug 1989.

[2] E. Guillot, P. E. Pottie, C. Valentin, P. Petit, and N. Dimarcq, "Horace: atomic clock with cooled atoms in a cell," in *Frequency and Time Forum, 1999 and the IEEE International Frequency Control Symposium, 1999., Proceedings of the 1999 Joint Meeting of the European*, vol. 1, 1999, pp. 81 –84 vol.1.

[3] S. Tremine, S. Guerandel, D. Holleville, A. Clairon, and N. Dimarcq, "Development of a compact cold atom clock," in *Frequency Control Symposium and Exposition, 2004. Proceedings of the 2004 IEEE International*, aug. 2004, pp. 65 – 70.

[4] W. Ketterle, A. Martin, M. A. Joffe, and D. E. Pritchard, "Slowing and cooling atoms in isotropic laser light," *Phys. Rev. Lett.*, vol. 69, p. 2483, 1992.

[5] E. Guillot, P. Pottie, and N. Dimarcq, "Three-dimensional cooling of cesium atoms in a reflecting copper cylinder," *Opt. Lett.*, vol. 26, no. 21, pp. 1639–1641, 2001.

[6] H.-D. Cheng, W.-Z. Zhang, H.-Y. Ma, L. Liu, and Y.-Z. Wang, "Laser cooling of rubidium atoms from background vapor in diffuse light," *Phys. Rev. A*, vol. 79, no. 2, p. 023407, Feb 2009.

[7] "Horace: A compact cold atom clock for galileo," *Advances in Space Research*, vol. 47, no. 5, pp. 854 – 858, 2011, scientific applications of Galileo and other Global Navigation Satellite Systems - II.

[8] L. Xiao and X. C. Wang and H. D. Cheng and W. Z. Zhang and L. Liu and Y. Z. Wang, "Loading of cold ^{87}rb atom with diffuse light in an integrating sphere," *Chin. Opt. Lett.*, vol. 8, no. 3, pp. 253–255, 2010.

[9] H. D. Cheng, X. C. Wang, L. Xiao, W. Z. Zhang, L. Liu, and Y. Z. Wang, "Cooling of rubidium atoms in pulsed diffuse laser light," *Chin. Phys. B*, vol. 20, no. 2, p. 023701, 2011.

[10] A. Godone, S. Micalizio, and F. Levi, "Pulsed optically pumped frequency standard," *Phys. Rev. A*, vol. 70, no. 2, p. 023409, Aug 2004.

[11] W. Z. Zhang, X. C. Wang, H. D. Cheng, L. Xiao, L. Liu, and Y. Z. Wang, "Lifetime measurement of cold atoms in an integrating sphere," *Chin. Phys. Lett.*, vol. 26, p. 083703, 2009.

Nonlinear Pressure Shifts of Alkali-Metal Atoms in Xenon

Bart H. McGuyer, Tian Xia*, Yuan-Yu Jau†, and William Happer
Department of Physics, Princeton University, Princeton, NJ 08544, USA
*Present address: Department of Physics, University of Wisconsin, Madison, WI 53706, USA
†Present address: Sandia National Laboratories, Albuquerque, NM 87123, USA
Email: bmcguyer@princeton.edu

***Abstract*— We demonstrate that the microwave resonant frequencies of ground-state ^{87}Rb and ^{133}Cs atoms have a nonlinear dependence on the pressure of the buffer gas Xe. Surprisingly, though the frequency shifts of Rb and Cs in the heavy noble gases Ar, Kr, and Xe are all negative, the nonlinearity in Xe is of the opposite sign to the nonlinearities in Ar and Kr. This discrepancy suggests that the frequency shifts due to alkali-noble van der Waals molecules in Xe are opposite in sign to the those in Ar and Kr. The nonlinearity in Xe shows a reproducible deviation from a simple model used to describe the shifts in Ar and Kr, which suggests that the spin-rotation interaction plays a significant role in the frequency shifts due to molecules. No nonlinearities were observed with the gases He and N_2, as expected, and also Ne to within experimental error, which suggests that molecules do not form in Ne. Additionally, we present improved measurements of the shifts of Rb in Ar and Kr and of Rb and Cs in He and N_2.**

I. Introduction

Compact, portable atomic frequency standards are based on the microwave resonance frequencies of alkali-metal atoms in inert buffer gas. The frequency shift of these resonances due to the buffer gas is known as the pressure shift. The size of the shift depends on the temperature and is remarkably linear with the buffer gas pressure p. Previous work has uncovered a nonlinear dependence of the pressure shifts of ^{87}Rb and ^{133}Cs in Ar and Kr, but not He and N_2, which is thought to be due to the loosely-bound van der Waals molecules that are known to form between alkali-metal and buffer-gas atoms in Ar, Kr, and Xe, but not He and N_2 [1], [2]. The precise measurement of nonlinear pressure shifts is a means to study van der Waals molecules and poorly-known interatomic potentials, and is important for determining optimal buffer gas mixtures. Here we present the first measurements of the nonlinear pressure shifts of ^{87}Rb and ^{133}Cs in natural abundance Xe.

II. Experimental Setup

We measured the resonant frequencies ν of ^{87}Rb and ^{133}Cs as a function of the pressure p of a pure buffer gas for the 0–0 hyperfine (or clock) transition between the higher- and lower-energy ground-state hyperfine sublevels $|\alpha\rangle$ and $|\beta\rangle$ with azimuthal angular momentum quantum numbers $m = 0$. Fig. 1 is a sketch of the experimental setup for both the Rb and Cs measurements, which is a laser-pumped, vapor-cell atomic clock that closely follows the description in Ref. [1]. To isolate the pressure shift, we suppress the frequency shift due to the laser, known as the light shift, by using feedback to lock the laser to a (pressure-dependent) zero-shift optical frequency that produces no light shift.

The Rb and Cs vapor cells are cylindrical Pyrex cells, 17 mm in diameter and 25 mm or 13 mm long, respectively, filled with a small amount of ^{87}Rb or ^{133}Cs metal. Both cells have an external buffer-gas reservoir, pressure gauge, and vacuum port, which permit convenient changes of the buffer-gas species and pressure. Section VI discusses the uncertainty in pressure measurement.

An air-heated, non-magnetic oven (O) holds the cell at a constant temperature of 40.0°C for Rb and 35.0°C for Cs. For this work, we extended our measurements to higher pressures than in Ref. [1]. We found that stabilization of the temperature gradient along the section of the vapor cell that exits the oven was necessary to reduce the additional clock noise and drift at higher pressures that are due to temperature-induced changes in the cell volume. Using the Rb cell and setup, we were unable to detect any pressure difference between the heated and unheated portions of the vapor cells due to thermal transpiration.

Helmholtz coils (HC) cancel ambient magnetic fields and provide a static longitudinal field B_0 of 1.0 G for Rb and 0.2 G for Cs. A Toptica DL100 diode laser (DL) provides 795 nm D1 optical pumping light for Rb and 895 nm D1 light

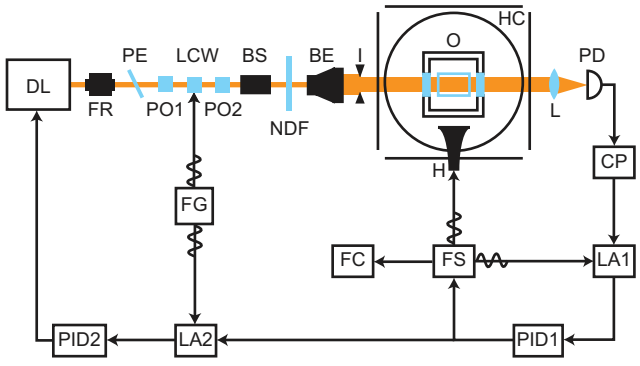


Fig. 1. Experimental setup. DL, diode laser; FR, Faraday rotator; PE, pellicle; PO, polarizer; LCW, liquid crystal wave plate; BS, beam shaper; NDF, neutral density filter; BE, beam expander; I, iris; O, oven; H, horn; HC, Helmholtz coils; L, lens; PD, photodetector; CP, current preamplifier; LA, lock-in amplifier; PID, PID controller; FS, frequency synthesizer; FC, frequency counter; FG, function generator.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

for Cs. A Faraday rotator (FR) isolates the laser from back-reflected light. A pellicle (PE) skims off light for analysis with a wavemeter and a Fabry-Perot interferometer (not shown). Polarizers (PO) ensure that the pumping light is linearly polarized. A liquid crystal wave plate (LCW) driven by a function generator (FG) provides about 30% intensity modulation of the pumping light at a rate of 2 Hz. A beam shaper (BS), a beam expander (BE), and an iris (I) ensure that the pumping light fills the cell uniformly. A rotatable neutral density filter (NDF) adjusts the pumping light intensity. A lens (L) collects the transmitted pumping light onto a photodetector (PD).

Microwaves from a frequency synthesizer (FS) are transmitted towards the cell by a horn (H) roughly 10 cm away to drive magnetic resonances. A frequency counter (FC) referenced to a rubidium frequency standard (not shown) measures the microwave frequency to a precision of 1 Hz. To estimate the uncertainty in frequency measurement, we use the sample standard deviation, not the standard deviation of the mean. For Rb, the microwaves are frequency modulated internally by the FS at a rate between 0.2-9.0 kHz, depending on the buffer gas and pressure, and a modulation index of about 1. For Cs, the microwaves are frequency modulated at a rate of roughly 255 Hz by modulating an external 10 MHz quartz oscillator (Stanford Research Systems SC10) used as an external reference for the FS. This method lead to a restricted range of usable modulation parameters and to increased carrier noise for the Cs measurements.

A lock-in amplifier (LA1) with a 10-300 ms time constant provides an error signal for a proportional-integral-derivative (PID) controller (PID1) to lock the microwave carrier frequency to the clock transition. A second lock-in amplifier (LA2) with a time constant of roughly 1-3 s provides an error signal for another PID controller (PID2) to lock the laser to a zero-shift optical frequency using a piezoactuator. As a quick test for light-shift suppression, we use the NDF filter to temporarily adjust the laser intensity by a factor of 2–4 to verify the intensity independence of the clock output.

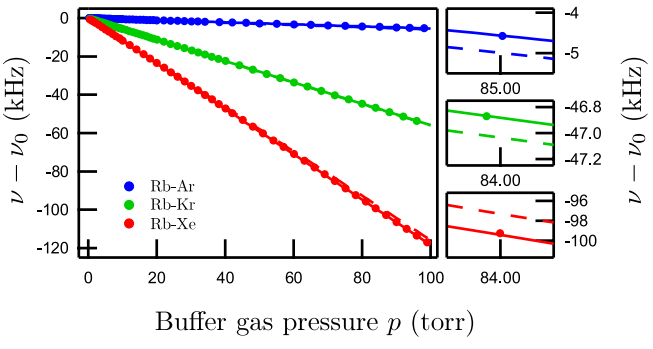


Fig. 2. Resonant frequency ν of ^{87}Rb in pure Ar, Kr, or Xe at a temperature of 40.0°C and magnetic field B_0 of 1 G, with zero-pressure frequency $\nu_0 \approx 6.834683$ GHz removed. Error bars are too small to display. Solid dots are the data, solid lines are the linear-limiting shifts $\nu_0 + sp$, and dashed lines are the binary shifts $\nu_0 + s_b p$, where $s_b = s - s_m$. Insets show that the linear-limiting shift is above the binary shift for Ar and Kr but below for Xe.

III. Nonlinear Pressure Shift of ^{87}Rb in Xe

The linear pressure shifts $\Delta^1\nu = \nu - \nu_0$ of alkali-metal atoms in the heavy noble gases Ar, Kr, and Xe are all negative [3], as measurements show for ^{87}Rb in Fig. 2. Previous work has shown that, after subtraction of the linear, limiting shift sp at high pressures, there is a residual nonlinear pressure shift $\Delta^2\nu = \Delta^1\nu - sp$ for Ar and Kr due to van der Waals molecules [1].

Fig. 3 shows data for ^{87}Rb in natural abundance Xe. Subtraction of the linear, limiting shift reveals a nonlinear pressure shift that is significantly larger than, and of opposite sign to, the nonlinear shifts observed in Ar and Kr. While the shift is described well by the model $\Delta_0^2\nu$ of Ref. [1] (Eq. (3), with parameters in Table I), developed for Ar and Kr, the fit residuals $\Delta^3\nu = \Delta^2\nu - \Delta_0^2\nu$ show a reproducible, nonlinear dependence on pressure not accounted for by that model.

As shown in Fig. 3(c), a new model $\Delta_1^2\nu$ (Eq. (7), with parameters in Table I), described in Section V, is an improved fit to the data. This model includes the effects of the spin-

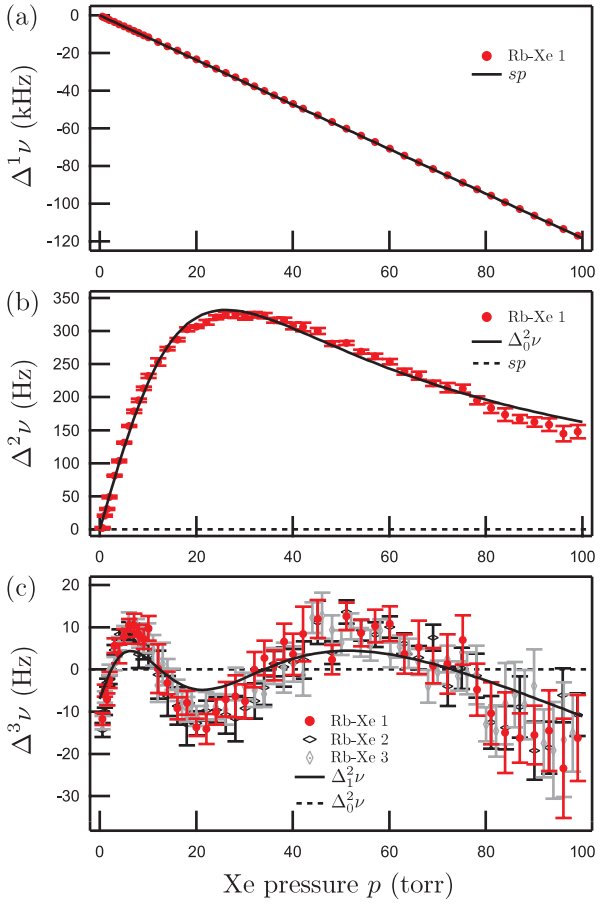


Fig. 3. Resonant frequency ν of ^{87}Rb in Xe at a temperature of 40.0°C and magnetic field B_0 of 1 G. The top panel (a) shows the shift $\Delta^1\nu = \nu - \nu_0$ from the zero-pressure frequency $\nu_0 \approx 6.834683$ GHz. The solid line is the linear, limiting shift sp at high pressures. The middle panel (b) shows the nonlinear shift $\Delta^2\nu = \Delta^1\nu - sp$. The solid curve is from Eq. (3) and Table I. The bottom panel (c) shows fit residuals $\Delta^3\nu = \Delta^2\nu - \Delta_0^2\nu$ for three separate data sets. The solid curve is from Eq. (7) and Table I.

rotation interaction, which were previously ignored. As a result, the data suggest that the spin-rotation interaction plays a significant role in the frequency shifts due to van der Waals molecules.

IV. Physics of Van der Waals Molecular Shifts

Short-lived van der Waals molecules are formed by three-body collisions and destroyed by two-body collisions. During the lifetime τ of a molecule, the alkali-metal atom is perturbed by the hyperfine-shift interaction

$$V_{\text{hfs}} = \delta A\, \mathbf{I} \cdot \mathbf{S}, \tag{1}$$

where δA is a perturbation to the hyperfine coupling of the nuclear spin $\mathbf{I}$ to the electron spin $\mathbf{S}$ for the alkali-metal atom, and by the spin-rotation interaction

$$V_{\text{sr}} = \gamma\, \mathbf{N} \cdot \mathbf{S}, \tag{2}$$

where γ is the coupling coefficient of the electron spin to the rotational angular momentum $\mathbf{N}$ of the colliding alkali-metal atom and buffer-gas atom or molecule [1], [2], [4], [5].

For a hypothetical van der Waals molecule with a single vibration-rotation state and negligible spin-rotation interaction ($\langle \gamma N \rangle = 0$), density matrix modeling predicts the nonlinear shift [1]

$$\Delta_0^2 \nu = -\left(\frac{1}{2\pi T}\right) \frac{\phi^3}{1+\phi^2}. \tag{3}$$

Here, the three-body molecular formation rate is $1/T$ and the effective hyperfine phase shift per molecule is

$$\phi = \langle \delta A \rangle [I] \tau / (2\hbar), \tag{4}$$

where $[I] = 2I+1$ and I is the nuclear spin quantum number of the alkali-metal atom. The molecules also contribute a slope

$$s_{\text{m}} = \phi/(2\pi T p) \tag{5}$$

to the slope s of the linear, limiting shift sp at high pressures.

V. Effects of the Spin-Rotation Interaction

For most alkali-noble molecules, the spin-rotation interaction is a large perturbation which introduces an effective molecular magnetic field

$$B_1 = \frac{\langle \gamma N \rangle}{g_S \mu_B} \tag{6}$$

oriented along $\mathbf{N}$, whose direction is distributed isotropically. Here, $g_S = 2.0023$ is the electronic g factor and μ_B is the Bohr magneton. For Rb and Cs in Xe, B_1 is much larger than the externally applied field B_0 in our measurements, so the total magnetic field seen by an alkali-metal atom in a molecule is approximately equal to B_1.

In this case, for hypothetical van der Waals molecules with a single vibration-rotation state, density matrix modeling predicts the nonlinear shift

$$\Delta_1^2 \nu = -\left(\frac{1}{2\pi T}\right) \sum_{\sigma=-2I}^{2I} \frac{W_\sigma (1+r_1\sigma)^3 \phi^3}{1+(1+r_1\sigma)^2\phi^2}, \tag{7}$$

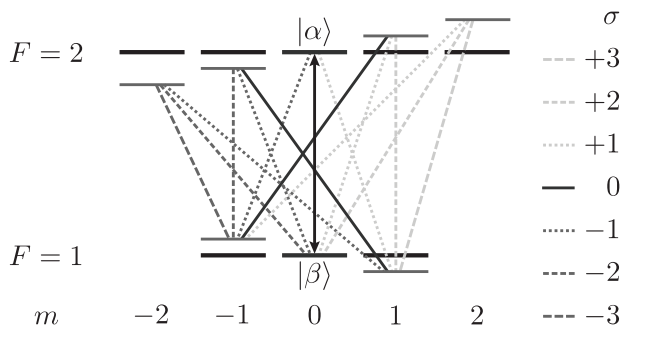


Fig. 4. Sketch of the different, effective hyperfine transitions indexed by σ in Eq. (7), which contribute to the 0-0 hyperfine coherence $\rho_{\alpha\beta}$ for ^{87}Rb. The transitions have different frequencies because of the first-order Zeeman shifts from the molecular field B_1.

where the dimensionless parameter

$$r_1 = \frac{2 g_s \mu_B B_1}{\langle \delta A \rangle [I]^2} = \frac{2\langle \gamma N \rangle}{\langle \delta A \rangle [I]^2} \tag{8}$$

characterizes the relative strengths of the spin-rotation and hyperfine-shift interactions, and the weight coefficients W_σ for the 0–0 hyperfine transition are

$$^{87}\text{Rb: } [W_3, \ldots, W_{-3}] = [9, 9, 23, 23, 23, 9, 9]/105 \tag{9}$$

$$^{133}\text{Cs: } [W_7, \ldots, W_{-7}] = [245, 245, 425, 425, 659, 659, 1231, 1231, 1231, 659, 659, 425, 425, 245, 245]/9009. \tag{10}$$

The weight coefficients W_σ depend on the choice of hyperfine transition, and as a result, so too does the nonlinear shift $\Delta_1^2 \nu$ of Eq. (7). The analysis that leads to Eq. (7) also predicts the molecular contribution s_{m} of Eq. (5) to the linear, limiting slope s at high pressures.

As sketched in Fig. 4 for ^{87}Rb, the summation index σ corresponds to the different, effective hyperfine transitions that contribute to the 0–0 hyperfine coherence $\rho_{\alpha\beta}$ after averaging over all orientations for B_1. The terms in the sum over σ in Eq. (7) are similar to $\Delta_0^2 \nu$, but have different effective phase shifts ϕ due to the first-order Zeeman shifts from B_1. The additional nonlinearity $\Delta^3 \nu = \Delta^2 \nu - \Delta_0^2 \nu$ observed in Fig. 3(c) is due to the interference of these terms as the molecular lifetime τ varies with the pressure p.

For $r_1 = 0$, the nonlinear shift Δ_1^2 of Eq. (7) simplifies to the previous model $\Delta_0^2 \nu$ of Eq. (3). In the limit of a large external field B_0, the analysis that leads to Eq. (7) also simplifies to Eq. (3).

VI. Pressure Gauge Nonlinearity

We measured the buffer gas pressure p with capacitance manometer gauges (MKS Instruments Baratron) to a precision of roughly ± 0.002 Torr and an accuracy of about $\pm 0.25\%$. We have found that these gauges typically have a significant quadratic nonlinearity. This pressure gauge nonlinearity can be mistaken for a nonlinear pressure shift, as Fig. 5 demonstrates for the case of ^{87}Rb in Ne.

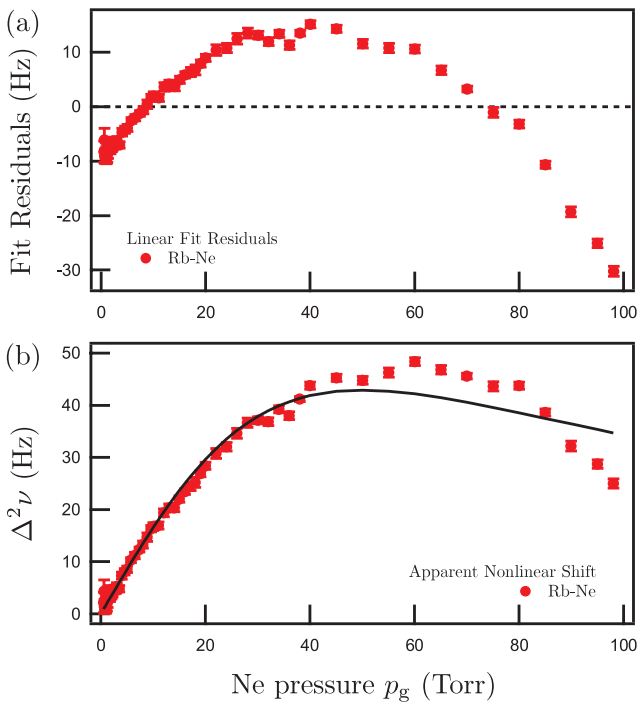


Fig. 5. Demonstration of a quadratic pressure gauge nonlinearity with measurements of the resonant frequency ν of ^{87}Rb in Ne at 40.0°C and $B_0 = 1$ G. The top panel (a) shows the residuals after fitting ν versus the gauge pressure p_g with a line, which have a quadratic dependence on the pressure p_g. The bottom panel (b) shows the apparent nonlinear shift $\Delta^2\nu$ after fitting ν versus p_g with the model $\Delta_0^2\nu$ of Eq. (3), which is due to the pressure gauge nonlinearity. For comparison, Fig. 6(b) shows the residuals after fitting ν versus the estimated pressure p from Eq. (11) with a line.

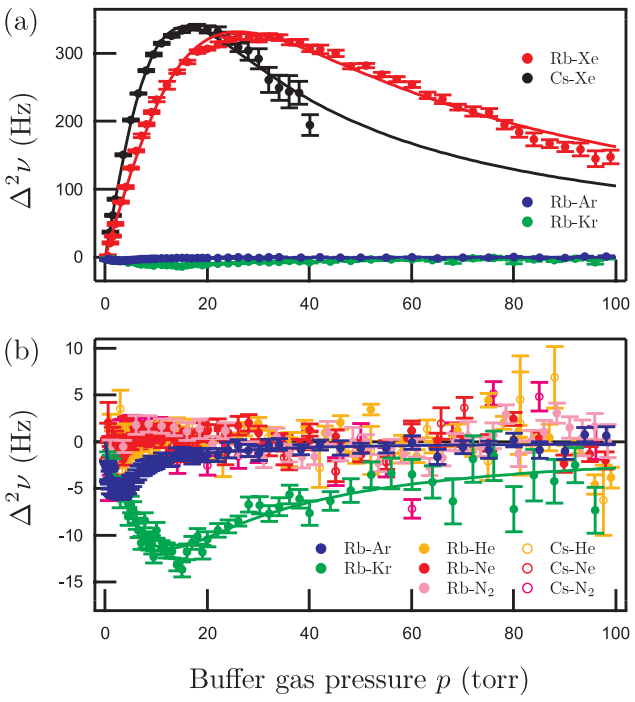


Fig. 6. Measured nonlinear shifts $\Delta^2\nu$ for ^{87}Rb at 40°C and $B_0 = 1$ G and for ^{133}Cs at 35°C and $B_0 = 0.2$ G. The top panel (a) shows the nonlinear shifts of Rb and Cs in Xe. The bottom panel (b) shows the nonlinear shifts of Rb in Ar and Kr and of Rb and Cs in He, Ne, and N_2. The curves shown are the model $\Delta_0^2\nu$ of Eq. (3). Fit parameters are listed in Table I.

We correct for this systematic error by estimating the actual pressure p from the measured pressure p_g as

$$p = r(\alpha)(p_g + \alpha p_g^2), \tag{11}$$

where the curvature α is determined for each pressure gauge through measurements of the shifts of the buffer gases He, Ne, and N_2. The coefficient

$$r(\alpha) = (1 + 75\alpha)/(1 + 150\alpha + 6000\alpha^2) \tag{12}$$

minimizes the least-squares linear bias of this correction, $\int (p - p_g)^2 dp_g$, over a 0-100 Torr pressure range. From our measurements, we determined the curvatures to be:

$$^{87}\text{Rb cell: } \alpha = (-3.40 \pm 0.22) \times 10^{-5} \text{ Torr}^{-1} \tag{13}$$

$$^{133}\text{Cs cell: } \alpha = (-1.71 \pm 0.46) \times 10^{-5} \text{ Torr}^{-1}. \tag{14}$$

Though the curvature for a pressure gauge may vary over time and with environmental conditions, these measured curvatures appear stable over the duration of our measurements.

After applying the correction of Eq. (11), we find that the pressure shifts of ^{87}Rb and ^{133}Cs in Ne are linear to within our experimental error, as shown in Fig. 6. However, the analysis for Ne is complicated by the observation of frequency shifts in some measurements at very low pressures, below about 1 Torr, which we think are systematic shifts of as yet undetermined origin. These low-pressure shifts were most noticeable with the Cs setup, where they were negative and could be as large as roughly -20 Hz at 0.7 Torr for Cs in He, Ne, or N_2. For the Rb setup, the low-pressure shifts were less reproducible and smaller, typically < 4 Hz at 0.5 Torr in Ar, Kr, Ne, and N_2, and were not observed in He. These shifts for Rb were usually negative, but could be of either sign for Rb in Ar. Detailed understanding of these low-pressure shifts requires further investigation.

VII. Nonlinear Pressure Shift Parameters

Table I lists the pressure-independent fit parameters for the nonlinear pressure shift data shown in Fig. 6. For Xe, the table includes results for both of the nonlinear pressure shift models, Eqs. (3) and (7). For Rb in Xe, the pressure-independent, spin-rotation parameter

$$r_1\phi p = \frac{\langle \gamma N \rangle \tau p}{[I]\hbar} \tag{15}$$

was set to an estimate calculated with the H_1^* and τp of Ref. [2]. For Cs in Xe, we were unable to estimate $r_1\phi p$ due to a lack of experimental data for τp, and so results are given for a fit to the data with $r_1\phi p$ as a free parameter. For the other gases, the table includes improved results over the previous work of Ref. [1], which are corrected for the pressure gauge nonlinearity.

As the data show, the effective ϕ due to molecules with Xe are negative, in contrast to the molecules with Ar and Kr.

TABLE I

FIT PARAMETERS FOR THE PRESSURE SHIFTS OF ^{87}RB AT 40°C AND $B_0 = 1$ G, AND OF ^{133}CS AT 35°C AND $B_0 = 0.2$ G.

Metal	Gas	$r_1\phi p$ rad Torr	Tp^2 sec Torr^2	ϕp rad Torr	s Hz Torr^{-1}
^{87}Rb	Xe	7.97	0.082 ± 0.010	-14.8 ± 1.5	-1184.0
		–	0.164 ± 0.013	-26.2 ± 1.2	-1183.7
	Ar	–	0.070 ± 0.012	2.21 ± 0.16	-53.71
	Kr	–	1.08 ± 0.31	12.5 ± 1.9	-558.1
	He				714.2
	Ne				387.3
	N_2				518.0
^{133}Cs	Xe	2.79	0.0185	-5.96	-2243.9
		–	0.059 ± 0.006	-15.8 ± 1.3	-2242.4
	He				1141.9
	Ne				579.9
	N_2				828.7

The linear shift coefficients s are in agreement with previous measurements [3], [6], [7]. The effective molecule formation rates $1/T$ are slower than those of Ref. [2], which could be a result of cancellation due to a variation of the phase shifts ϕ_v, and their sign, with different molecular vibration-rotation states v [1]. Table I also shows that the fit estimates for ϕp and Tp^2 are significantly affected by the choice of model, which highlights the importance of the spin-rotation interaction.

VIII. CONCLUSION

We report the first measurements of the nonlinear pressure shifts of ^{87}Rb and ^{133}Cs in Xe, and the first evidence that the spin-rotation interaction plays an important role in the hyperfine frequency shifts due to van der Waals molecules. We present a new model for nonlinear pressure shifts that includes the spin-rotation interaction in the limit of a weak external magnetic field, and which is a better fit to the Xe data. This model implies a dependence of the molecular shifts on the choice of the hyperfine transition and on the external magnetic field. No nonlinearity was observed for Rb and Cs in Ne, to within experimental error, suggesting that van der Waals molecules do not form in vapor cells with Ne buffer gas. We correct for a previously unknown pressure gauge nonlinearity with improved measurements of Rb in Ar and Kr and of Rb and Cs in He and N_2. Future work will explore the isotope dependence of the shifts due to van der Waals molecules with measurements of the nonlinear pressure shifts of ^{85}Rb in Ar, Kr, Xe, He, Ne, and N_2.

ACKNOWLEDGMENT

The authors are grateful to M. J. Souza for making cells, to F. Gong and S. W. Morgan for contributions to the original apparatus, and to J. H. Hendricks at NIST and D. Jacobs at MKS Instruments for helpful discussions about capacitance manometers. This work was supported by the Air Force Office of Scientific Research, the Department of Defense, and the Department of Energy.

REFERENCES

[1] F. Gong, Y.-Y. Jau, and W. Happer, "Nonlinear pressure shifts of alkali-metal atoms in inert gases," *Phys. Rev. Lett.*, vol. 100, no. 23, p. 233002, 2008. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.100.233002

[2] M. A. Bouchiat, J. Brossel, and L. C. Pottier, "Evidence for Rb-rare-gas molecules from the relaxation of polarized Rb atoms in a rare gas. Experimental results," *J. Chem. Phys.*, vol. 56, no. 7, pp. 3703–3714, 1972. [Online]. Available: http://link.aip.org/link/?JCP/56/3703/1

[3] M. Arditi and T. R. Carver, "Pressure, light, and temperature shifts in optical detection of 0-0 hyperfine resonance of alkali metals," *Phys. Rev.*, vol. 124, no. 3, pp. 800–809, 1961. [Online]. Available: http://dx.doi.org/10.1103/PhysRev.124.800

[4] R. A. Bernheim, "Spin relaxation in optical pumping," *J. Chem. Phys.*, vol. 36, no. 1, pp. 135–140, 1962. [Online]. Available: http://link.aip.org/link/?JCP/36/135/1

[5] Z. Wu, T. G. Walker, and W. Happer, "Spin-rotation interaction of noble-gas alkali-metal atom pairs," *Phys. Rev. Lett.*, vol. 54, no. 17, pp. 1921–1924, 1985. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.54.1921

[6] B. L. Bean and R. H. Lambert, "Temperature dependence of hyperfine density shifts. IV. ^{23}Na, ^{39}K, and ^{85}Rb in He, Ne, Ar, and N_2 at low temperatures," *Phys. Rev. A*, vol. 13, no. 1, pp. 492–494, 1976. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.13.492

[7] J. Vanier, R. Kunski, N. Cyr, J. Y. Savard, and M. Têtu, "On hyperfine frequency shifts caused by buffer gasses: Application to the optically pumped passive rubidium frequency standard," *J. Appl. Phys*, vol. 53, no. 8, pp. 5387–5391, 1982. [Online]. Available: http://link.aip.org/link/?JAP/53/5387/1

Update on a Comparison of Cesium Fountain Primary Frequency Standards

Thomas E. Parker
Time and Frequency Division
National Institute of Standards and Technology
Boulder, CO USA
tparker@boulder.nist.gov

***Abstract*—An update to a previous long term frequency comparison of cesium fountain primary frequency standards [1] has been made. This update covers the approximate three year interval from March 2008 to March 2011 and also includes two new standards. Also, simulated data have been used to estimate the biases and standard deviation of the Birge ratio as a function of the degrees of freedom.**

I. INTRODUCTION

In January 2010 the results of a long-term frequency comparison of cesium fountain primary frequency standards were reported in Metrologia covering the interval from August 2000 to June 2009 [1]. The new study presented here gives an update covering ten fountains from seven laboratories over the time frame from approximately March 2008 to March 2011. Included are two new fountains, NPL-CsF2 (at the National Physical Laboratory in the United Kingdom) and PTB-CsF2 (at the Physikalisch Technische Bundesanstalt in Germany), that were not present for the study in [1]. The methodology followed here is substantially the same as that in [1], but with one difference. Rather than using either NIST-F1 (at the National Institute of Standards and Technology in the USA) or SYRTE-FO2 (at the Laboratoire National de Métrologie et d'Essais, Systèmes de Référence Temps Espace in France) as references, only SYRTE-FO2 was used as the main reference. This change was made only because SYRTE-FO2 is almost always present for comparison. In the few instances when NIST-F1 was used as the reference, an adjustment based on the average frequency difference between NIST-F1 and SYRTE-FO2 was applied. Details of the comparison methodology are given in [1]. Also new in this paper are estimates from simulated data for biases and the standard deviation of the Birge ratio as a function of degrees of freedom. This information provides confidence levels for the Birge ratios.

II. FOUNTAIN COMPARISON METHOD

The comparisons were made using fountain data available in the BIPM (Bureau International des Poids et Measures) publication Circular T (www.bipm.org/jsp/en/TimeFtp.jsp?TypePub=publication). Pairs of data points that occurred close in time were used and in most cases there was some overlap. Both TAI (International Atomic Time) and AT1E (a post-processed maser ensemble at NIST) [1] were used as flywheel frequency references, and the average of the two independent results was calculated in order to gain a small reduction in the dead-time uncertainty. In all, 128 comparisons were made.

For a pair of runs all type A uncertainties, u_A, from both standards were combined in quadrature. Also the type B (systematic) uncertainties, u_B, from both standards were combined in quadrature. Over multiple runs the combined u_A averages down, while for multiple runs the weighted mean of the combined u_B was used. Consequently, u_B does not average down over time. u_A and u_B are added in quadrature to get the total uncertainty. It has been assumed that the u_B between standards are uncorrelated and that the u_B for the same standard is highly correlated over time. Neither assumption is strictly true, but the approach is a reasonable compromise.

Table 1 shows an example of a fountain comparison from a data pair. In the table MJD is the Modified Julian Date and y is the fractional frequency difference in units of 10^{-15}. u_{dead} is the uncertainty introduced by the noise in the flywheel due to misalignment of the run times shown in column one. $u_{dead} = 0$ if the start and stop times are identical. u_{TAI} is the transfer noise into TAI [2]. u_l includes dead time internal to the laboratory. u_A, u_l, u_{TAI}, u_{dead} are all type A uncertainties. u_C, u_{CA} and u_{CB} are the combined uncertainties of the comparison. All fountain uncertainties are 1 standard uncertainty (1 sigma).

The top left section in Table 1 shows the frequency of TAI with respect to NIST-F1 over the specified MJDs along with the associated uncertainties. The lower left section shows the same for SYRTE-FO2. The middle right section shows the comparison results, where u_{dead} comes from the lack of perfect overlap. AT1E was also used as a reference fly wheel in a manner similar to that in Table 1, and the two results were averaged to gain a small reduction in the dead-time

TABLE 1. Example of a fountain comparison data pair and the difference calculation for NIST-F1 and SYRTE-FO2. Fractional frequencies are in units of 10^{-15}.

NIST-F1

MJD	MJD	Duration	y(TAI-F1)	u_A	u_B	u_I	u_{TAI}
54314-54339	54326.5	25d	-3.4	0.3	0.3	0.1	0.4

u_{dead}	y(FO2-F1)	u_C	u_{CA}	u_{CB}	overlap
0.7	0.3	1.2	1.05	0.58	15d

SYRTE-FO2

MJD	MJD	Duration	y(TAI-FO2)	u_A	u_B	u_I	u_{TAI}
54309-54329	54319	20d	-3.7	0.3	0.5	0.1	0.5

uncertainty. The same procedure was used for all of the fountain comparisons.

III. BIRGE RATIO

The Birge ratio, R_B, is a consistency test and should be close to unity if stated uncertainties are correct. It is a test to see if the scatter in the data is consistent with the stated uncertainties. The definition of the Birge ratio is shown in equations (1) and (2) below, where χ_r^2 is the reduced chi-square. R_B can be calculated either for the same pair of standards over time or for the average frequency differences in a group of standards.

$$R_B = \sqrt{\chi_r^2} \tag{1}$$

$$R_B = \sqrt{\frac{\sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{u_i^2}}{(n-1)}} \tag{2}$$

The standard chi-square calculator (for example www.fourmilab.ch/rpkp/experiments/analysis/chiCalc.html) is stated to be not accurate when the degrees of freedom, DOF, are less than about ten. Therefore simulated data made up of 81,920 samples were used to confirm the standard deviation and bias of R_B as a function of DOF. The simulated data were generated using a common spread sheet random-number generator and the Box-Muller transformation to provide a Gaussian distribution. The results are shown in Figs. 1 and 2. The standard deviation (for a +/-34 % interval) as a function of DOF provides a confidence interval for the Birge ratio and is essentially the same for the simulated data as from the calculator as shown in Fig. 1. This is also true for the median as seen Fig. 2. The mean and median however are not the same because the distribution is asymmetric. Here the DOF is one less than the number of data points.

IV. FOUNTAIN COMPARISON RESULTS

Table 2 shows the results of the nine fountain comparisons for the interval of approximately March 2008 to March 2011.

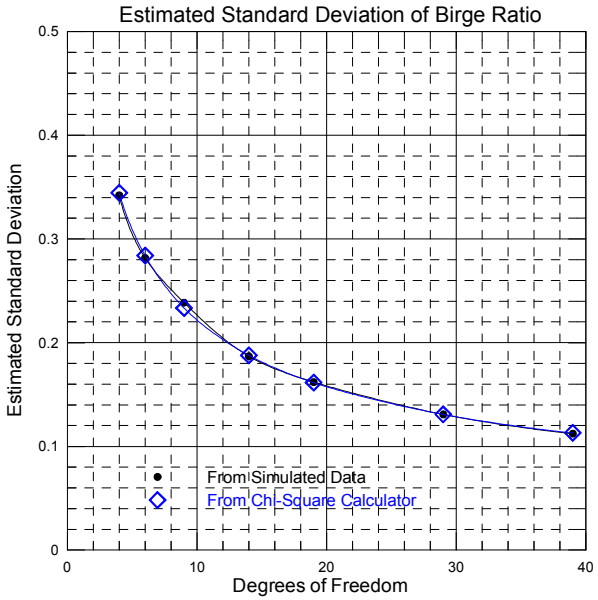


Figure 1. Estimated standard deviation of the Birge ratio from simulated data and from a chi-square calculator.

Each fountain is compared to SYRTE-FO2, and the average frequency difference, y_{avg}, is shown. The type A and type B uncertainties of the comparisons are U_{CA} and U_{CB}, and the combined total comparison uncertainty is U_C. The last column on the right gives the number of data pairs for each comparison. R_{BA} is the Birge ratio using only the type A uncertainties, and R_{BC} is the Birge ratio using the combined uncertainty. Here the Birge ratio is based on the scatter over time. In general one would expect R_{BC} to be less than 1 and R_{BA} to be larger than 1 [1]. There is cause for concern if R_{BA} is smaller than one in a statistically significantly fashion or if R_{BC} is larger than one in a statistically significantly fashion. This would indicate that the uncertainties are either overstated or understated respectively. Only SYRTE-FOM and IT-F1 (operated by the Istituto Nazionale di Ricerca Metrologica, (INRIM) in Italy) exhibit such behavior, with SYRTE-FOM having a small Birge ratio and IT-CSF1 having a large ratio. In both cases the deviation from 1 is more than two standard

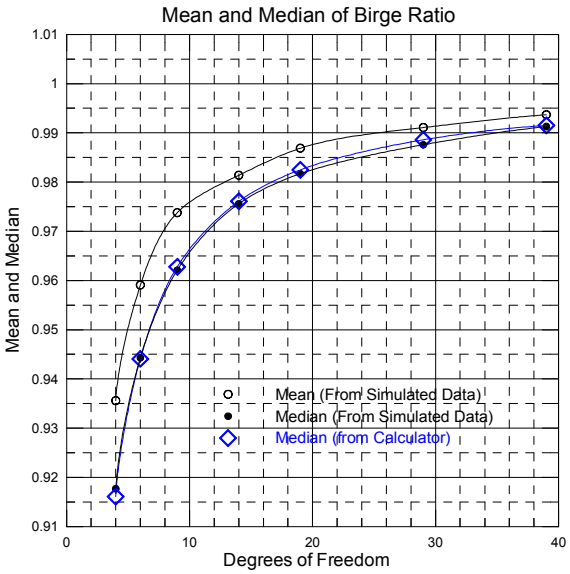


Figure 2. Biases in the Birge ratio at low degrees of freedom.

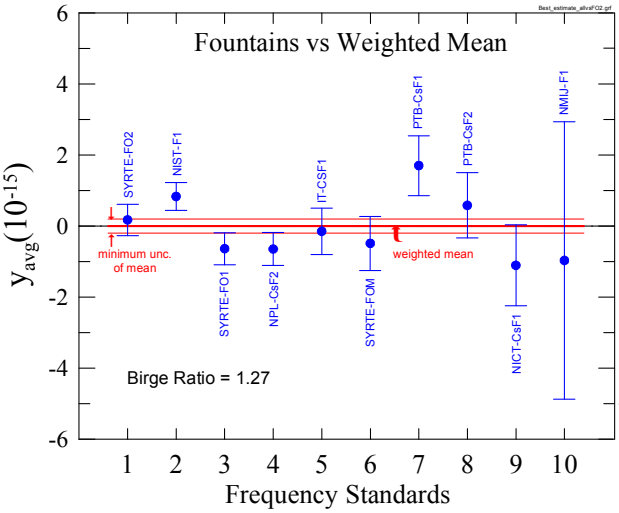


Figure 3. Average fountain frequencies vs the weighted mean.

TABLE 2. Fountain comparisons relative to SYRTE-FO2. All frequencies are in units of 10^{-15}. In this type of comparison an offset of 1.4 sigma or less is considered entirely normal.

Fountains	y_{avg}	U_C	U_{CA}	U_{CB}	R_{BA}	R_{BC}	# pairs
PTB-CsF1 vs SYRTE-FO2	+1.53	0.95	0.27	0.91	1.97	1.09	10
NIST-F1 vs SYRTE-FO2	+0.66	0.59	0.25	0.54	1.16	1.02	18
PTB-CsF2 vs SYRTE-FO2	+0.41	1.02	0.43	0.92	1.22	0.99	8
ITCS-F1 vs SYRTE-FO2	-0.32	0.79	0.31	0.73	1.89	1.56	15
SYRTE-FOM vs SYRTE-FO2	-0.66	0.88	0.24	0.85	0.56	0.43	14
SYRTE-FO1 vs SYRTE-FO2	-0.81	0.63	0.14	0.62	0.83	0.61	21
NPL-CsF2 vs SYRTE-FO2	-0.82	0.64	0.27	0.58	1.08	0.93	15
NMIJ-F1 vs SYRTE-FO2	-1.14	3.93	0.28	3.92	1.68	0.55	21
NICT-CsF1 vs SYRTE-FO2	-1.28	1.22	0.60	1.06	0.71	0.59	6

deviations. Significant errors in total uncertainties may come from sources other than the fountains themselves because substantial uncertainty is introduced by frequency transfer and dead time.

Three fountains in Table 2 show a positive average frequency offset relative to SYRTE-FO2, and six have a negative offset. In five cases the average frequency difference is larger than the total combined uncertainty, but no comparisons with SYRTE-FO2 exceed 2 sigma. In this type of comparison an offset of 1.4 sigma or less is considered entirely normal [1]. NMIJ-F1 is operated by the National Metrology Institute of Japan (NMIJ) and NICT-CsF1 is operated by the National Institute of Information and Communication Technology (NICT) in Japan.

A weighted mean of the offsets from SYRTE-FO2 can be calculated using the data in Table 2, and the average offsets from the mean for all ten fountains are shown in Fig. 3. The Birge ratio for the group of fountains shown in Fig. 3 is calculated using the combined uncertainties from each fountain. At 1.27 it is just over one standard deviation for nine DOF which is about 1.20 (see Figs. 1 and 2). The uncertainty of the weighted mean is about 2×10^{-16} if one assumes that all of the individual fountain uncertainties are uncorrelated. This is probably not true, but accurately quantifying the degree of correlation would be nearly impossible. Six of the ten standards in Fig. 3 are within 1 sigma of the mean and all are within 2 sigma. However, several standards are more than 2 sigma separated from each other. Specifically, NIST-F1 and PTB-CsF1 on the positive side of the mean versus SYRTE-FO1 and NPL-CsF2 on the negative side are separated by nearly 2.5 sigma.

V. CONCLUSIONS

Overall, the fountains are in fairly good agreement, although some deviate from the mean by nearly 2 sigma and from each other by nearly 2.5 sigma. The Birge ratio at 1.27 is just over one standard deviation and therefore does not indicate any serious problems. In the earlier study [1] the Birge ratio was 0.77, indicating less scatter relative to the stated uncertainties, which were generally larger than in the current study. In the current study, the scatter among the fountains is a little larger than one would expect from the stated uncertainties, but not enough to cause serious concerns. The Birge ratio indicates that there is a 90 % probability that, averaged over all the fountains, the uncertainties (including dead time and frequency transfer) have been understated by perhaps 20 %.

ACKNOWLEDGMENTS

Thanks to Steve Jefferts and Tom Heavner for helpful discussions.

REFERENCES

[1] Thomas E. Parker, "Long-term comparison of caesium fountain primary frequency standards", *Metrologia*, 47, pp 1-10, 2010.

[2] Gianna Panfilo and Thomas E. Parker, "A theoretical and experimental analysis of frequency tranfer uncertainty including frequency transfer into TAI", *Metrologia*, 47, pp 552-560, 2010.

Temperature Compensation via RF Power Adjusting in Chip Scale Atomic Clocks

Yeqing Li, Ke Deng, Qing Meng, Daiting Shi, Xuzong Chen, and Zhong Wang
School of Electronics Engineering & Computer Science (EECS), Peking University
Beijing, 100871, P. R. China
Email: yqli@pku.edu.cn

***Abstract*—In this paper, we propose a new method to reduce the temperature sensitivity of Chip Scale Atomic Clocks. We studied the relationship between RF power and the light shift as well as the relationship between the temperature and the collision shift, and the results are given. It is shown that the frequency shift caused by the change of the temperature can be compensated by the variation of light shift through the adjustment of RF power, in order to reduce the total output frequency shift.**

I. INTRODUCTION

Recently, chip scale atomic clocks (CSACs) [1, 2] have been recognized to be significant in the future portable electronic devices and other frequency standard fields [3-5]. Generally, CSACs are based on Coherent Population Trapping (CPT) [6-8] phenomenon which will occur in a Λ transition if the frequency difference of the two laser fields is close to the hyperfine splitting of the two ground states [9-11].

In CSACs, the atomic ground state hyperfine splitting frequency may be shifted as a result of collision between buffer gas atoms and alkali atoms. This frequency shift is sensitive to cell temperature. Because of the harsh application environment of CSACs, it is difficult to improve the temperature stability of the cell, which will limit the CSACs to achieve a good long-term stability. Many methods have been studied to reduce the temperature sensitivity of CSACs [12-14]. In our previous experiments, the rate of frequency shift with cell temperature is about 0.4Hz/K after optimizing the two buffer gases' ratio and pressure.

In order to reduce the dependence of the output frequency of CSACs on cell temperature, we propose that the frequency shift rate which is related to temperature can be compensated by adjusting the RF power level. The remarkable advantage of this approach is that we can relax the requirement on ambient temperature control. In the experiment, we measure the rates of frequency shift with temperature and RF power respectively. Then in the operation of CSACs, the cell temperature drift is detected and RF power level is adjusted exactly through the microprocessor in the control circuit to compensate the frequency shift coming from temperature drift.

Generally, the long-term instability of the temperature is around 1K. In this case, a frequency shift rate of 0.6Hz/K will lead to the clock frequency instability of 2×10^{-10}. After adjusting the RF power, the clock frequency instability is supposed to be improved to 10^{-11} region, which will broaden CSACs' application fields.

II. LIGHT SHIFT AND COLLISION SHIFT

Figure 1 shows the Λ type structure from simplifying the energy levels of the real atoms [15]. |1> and |2> are the two hyperfine splitting levels of ground state of alkali atoms, while |3> is the excited state. Γ* stands for the relaxation rate of the spontaneous emission of the upper level. γ_2 represents the rate of the coherence in the ground state tending to zero. Δ is the lasers detuning from optical resonance, while ω_1 and ω_2 are the laser angular frequencies.

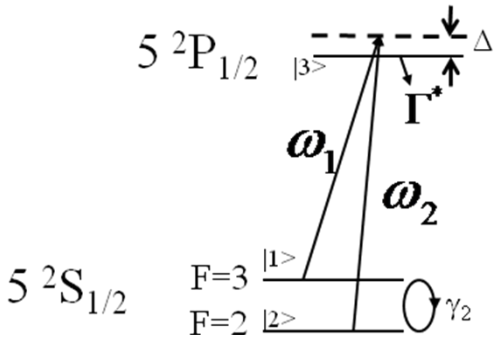


FIG. 1.Energy level of ^{85}Rb. Γ* is the decay rate from the excited state to the ground-state levels; Δ is the detuning; γ_2 is the ground-state relaxation rate of the coherence.

The total shift of the CPT resonance frequency can be expressed as

$$\nu_{total} = \nu_B + \nu_c + \nu_{ls} \quad . \tag{1}$$

Here ν_B is the magnetic field shift which is independent on the cell temperature. ν_c is the collision shift, and it is caused by the collisions between the buffer gases and the rubidium atoms and mainly influenced by the temperature as

$$\nu_c = P_0[\beta + \delta(T - T_0) + \gamma(T - T_0)^2], \quad (2)$$

where P_0 stands for the pressure of the buffer gases, β is the pressure shift coefficient, T is the atom cell temperature, T_0 is the reference temperature at which the coefficient δ is measured, and δ and γ are the first and second order of the temperature coefficient respectively. The relationship can be simplified in a small temperature range as

$$\nu_c = P_0[\beta + \delta(T - T_0)]. \quad (3)$$

ν_{ls} is the light shift, and both the RF modulation depth and the temperature will influence it, but generally we neglect the effect caused by the temperature. Every sidebands create light shift because the AC stark effect caused by the pump light. The AC stark effect is caused by a varying electric field, typically a laser, and leads to the shift of the atom energy levels. The contribution of each sideband to the total light shift can be measured as

$$\Delta\omega_n = \frac{1}{4}|\omega_R|^2 \frac{\Delta}{\Delta^2 + (\Gamma^*/2)^2}, \quad (4)$$

where ω_R represents the Rabi frequency, Δ is the lasers detuning from optical resonance and Γ* also stands for the decay rate from the excited state to the ground-state levels.

The total light shift can be calculated through a summation over all sidebands of the modulated laser which can be expressed as

$$\Delta\omega = \ldots\ldots + C_{-2}I_{-2} + C_{-1}I_{-1} + C_0I_0 + C_{+1}I_{+1} + C_{+2}I_{+2} + \ldots\ldots, \quad (5)$$

where Δω stands for the total light shift as same as the meaning of ν_{ls}, $C_{\pm n}$ means the coefficient of sideband ±n, and $I_{\pm n}$ is the intensity of each sideband. Taking the influence of all the sidebands into account, we can put the relationships in a simple way as

$$\Delta\omega = \alpha I, \quad (6)$$

With the changing of RF modulation depth, which will increase when RF power level goes up, the intensity will turn into another distribution, which means the coefficient α will change while the total light intensity is remaining the same.

III. Experiment Setup

In the experiment, the rubidium cell we used was 1cm in length and 1cm in diameter. The cell contained a mixture of natural rubidium, i.e. 72% ^{85}Rb and 28% ^{87}Rb [16-19], with buffer gases Ne and Ar. The temperature-controlled Vertical Cavity Surface Emitting Laser was modulated by the laser DC current and the RF generator, and these two parts were coupled by the Bias-Tee. The laser beam was collimated and was 3mm in diameter. After being circularly polarized and attenuated, it went through the rubidium cell which had magnetic shield and was temperature controlled.

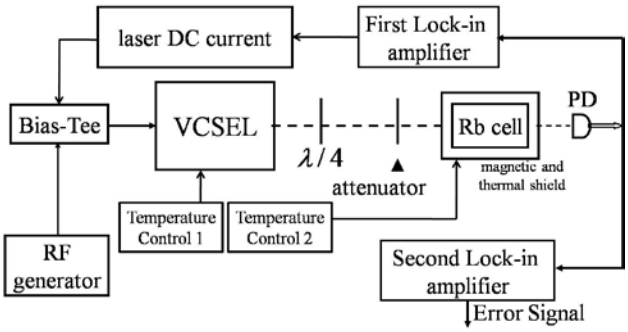


FIG. 2. Experiment Setup for measuring the shift of CPT frequency. RF generator and temperature control part can be used to adjust RF power level and temperature respectively.

The transmitted light was converted to electronic signal by a photo detector. Within the first lock-in loop, the laser's wavelength was locked on the ^{85}Rb D1 line absorption spectrum. The RF generator (model Agilent E8257D), which was phase locked into a commercial rubidium clock, was modulated at 190Hz. As a result of the difference between the CPT resonance frequency and the microwave frequency, an error signal was given. When the error signal was tuned to zero, the output RF generator frequency was the same as the CPT resonance frequency. We used full hyperfine frequency modulation in our experiment [20].

IV. Experiment Results

In FIG.3, it is shown that in the range -8~-3 of RF power, the CPT frequency shift is getting bigger with the increasing of RF power and the curve is approximately linear. However, the shift has a sharp reduction in the range of -3~-1 dBm. It's because that the increase of RF power in this range leads to an increase of intensity on the ±2 or other high-order sidebands which are not the pump light. As a result, the intensity of the pump light is reduced, the CPT phenomenon is weak, and the CPT frequency shift is small correspondingly. So the range of -8~-3 dBm is suitable for emerging CPT resonance since the pump light has enough intensity in that range.

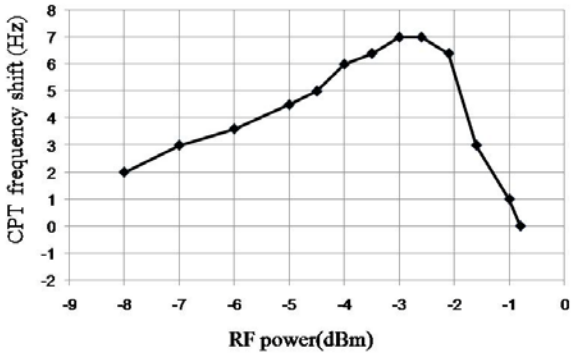


FIG. 3. CPT frequency shift with RF power level.

Figure 4 shows that the CPT frequency shift goes up when the temperature increases [21]. The amplitude of the CPT signal is given by

$$S \propto n \frac{\omega_{1R}^2 \omega_{2R}^2}{\left(\Gamma^*\right)^3} \frac{1}{2\left(\gamma_2 + \frac{\omega_{1R}^2 + \omega_{2R}^2}{2\Gamma^*}\right)}. \quad (7)$$

Here n is rubidium atom vapor density, ω_{1R} and ω_{2R} are rabi frequencies, γ_2 is the ground-state relaxation rate of the coherence and Γ^* is the relaxation rate. When the cell temperature is lower, the number n will become smaller. So the CPT signal is weak when the temperature is too low. However, as a result of the absorption of the light, the signal will become weak with a high cell temperature as well. So we have to control the temperature in a proper range, i.e. from 35℃ to 55℃ in FIG.4 with the size of cell considered.

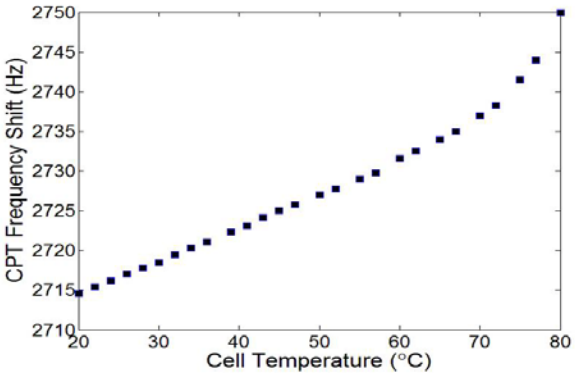


FIG. 4. CPT frequency shift with cell temperature.

As analyzed above, we can choose the range -8~-3 dBm of RF power and the range 35℃ to 55℃ of the temperature change to implement the compensation we proposed. For instance, we can set the initial work conditions at -5 dBm and 45℃, the middle values of the two curves. It is can be seen from FIG.3 that the CPT frequency shifted by around 3Hz with the RF power varying from -5 to -3. In FIG.4, the CPT frequency shift changed by about 3HZ with the cell temperature varying from 40 ℃ to 50 ℃. From the results and data, we can find that both of the curves had approximately linear areas and the variance of the frequency shift was comparable.

The results which show CPT frequency shift changing with the RF modulation depth and the cell temperature prove that the variance of the frequency shift is comparable and almost linear in some range. So considering the positive relationship between RF power level and modulation depth, adjusting RF power level can be used to compensate the CPT shift caused by the change of temperature.

In addition, in order to obtain the same modulation depth, the RF power may be adjusted differently in different lasers. But based on the observation in other related experiments, the trend of the CPT frequency shift changing with the RF power is the same in substance.

V. CONCLUSION

In this paper, we propose a new method to reduce the CSACs' temperature sensitivity and demonstrate it through theory analysis and experiment results. The total CPT frequency shift can be reduced by the compensation of the light shift varying with the RF power and the collision shift caused by the change of temperature.

Reducing the temperature sensitivity is crucial for the long term stability of CSACs. It may enhance the commercial feasibility of CSACs and also improve their reliability, especially when they are used in harsh application environmental conditions where it is difficult to improve the temperature stability of the cell. And this also can relax the requirement on the accuracy of the buffer gas pressure.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No.11074012.

REFERENCES

[1] R.Lutwak, "The Chip-Scale Atomic Clock – Recent Developments," the 23rd European Frequency and Time Forum and the 63rd consective meeting of the IEEE International Frequency Control Symposium (Besancon, France, 2009) pp.573-577.

[2] Y.-Y. Jau, A. B. Post, N. N. Kuzma, A.M. Braun, M.V. Romalis, and W. Happer, "Intense, narrow atomic-clock resonances," *Phys. Rev.* Lett., vol. 92. p. 110801, 2004.

[3] R. Lutwak, J. Deng, W. Riley, M. Varghese, J. LeBlanc, G. Tepolt, M. Mescher, D. K. Serkland, K.M. Geib, and G. M. Peake, 2005, "The Chip-Scale Atomic Clock – Low-Power Physics Package," in Proceedings of the 36th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 2004, Washington, D.C., USA (U.S. Naval Observatory, Washington, D.C.), pp. 339-354.

[4] R. Lutwak, D. Emmons, T. English, W. Riley, A. Duwel, M. Varghese, D. K. Serkland, and G. M.Peake, 2004, "The Chip-Scale Atomic Clock – Recent Development Progress," in Proceedings of the35th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 2-4 December 2003, San Diego, California, USA (U.S. Naval Observatory, Washington, D.C.), pp. 467-478.

[5] R. Lutwak, D. Emmons, W. Riley, and R.M. Garvey, "The Chip-Scale Atomic Clock – Coherent Population Trapping vs. Conventional Interrogation," in: Proceedings of the 34th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, Virginia, 2002, p. 539.

[6] N. Cyr, M. Tetu, and M. Breton, "All-optical microwave frequency standard: a proposal," *IEEE Trans. Instrum. Meas.* Vol. 42 p. 640, 1993.

[7] J. Vanier, "Atomic clocks based on coherent population trapping: a review," *App. Phys. B*. vol. 81, p.421, 2005.

[8] R. Wynands and A. Nagel, "Precision specrtoscopy with cohernt darkstates", Appl. Phys. B, Laser Optics, vol.68, no. 1, pp. 1-25, Jan.1999.

[9] J. Vanier, M.W. Levine, D. Janssen, M.J. Delaney, "On the use of intensity optical pumping and coherent population trapping techniquesin the implementation of atomic frequency standards", IEEETransaction on Instrumentation and Mesurement, 52, pp. 822-831,2003.

[10] J. Vanier, A. Godone, and F. Levi, in proceedings of the IEEE International Frequency control Symposium, pp. 96-99, 1999.

[11] M. Merimaa, T. Lindvall, I. Tittonen, and E. Ikonen, "All-optical atomic clock based on coherent population trapping in ^{85}Rb," J. opt. Soc. Am. B., vol. 20, p. 273, 2003.

[12] D.Miletic, P. Dziuban, R. Boudot , M. Hasegawa, "Quadratic dependence on temperature of Cs 0-0 hyperfine resonance frequency in

single Ne buffer gas microfabricated vapour cell," Electronics Letters, Vol. 46, p. 1069 – 1071, 2010.

[13] Ke Deng, Xuzong Chen, and Zhong Wang, "Minimization of the Temperature Coefficient of Resonance Frequency Shift in the Coherent Population Trapping Clock," Optics Letters 36,1740 , 2011.

[14] R. Lutwak, A. Rashed, M. Varghese, G. Tepolt, J. LeBlanc, M. Mescher, D. K. Serkland, K. M. Geib, G. M. Peake, and S. Römisch, "The Chip-Scale Atomic Clock –Prototype Evaluation," 39th Annual Precise Time and Time Interval (PTTI) Meeting, pp.269-290, 2008.

[15] J. Vanier, A. Godone and F. Levi, "Coherent population trapping in cesium: Dark lines and coherent microwave emission," *Phys. Rev. A.* vol. 58, p. 2345, 1998.

[16] M. Zhu, L. S. Cutler, "Theoretical and experimental study of light shift in a CPT-based Rb vapor cell frequency standard," in: 32nd Annual Precise Time and Time Interval (PTTI) Meeting, Reston, VA, 2000, p. 311.

[17] S. Knappe, R. Wynands, J. Kitching, H. G. Robinson and L. Hollberg, "Characterization of coherent population-trapping resonances as atomic frequency references,"*J. Opt. Soc. Am. B*., vol. 18, p. 1545, 2001.

[18] LIU Lu, GUO Tao, DENG Ke, LIU Xin-Yuan, CHEN Xu-Zong, WANG Zhong, "Frequency Stability of Atomic Clocks Based on Coherent Population Trapping Resonance in ^{85}Rb", CHIN.PHYS.LETT. 24, 1883 (2007).

[19] S. Knappe, V. Gerginov, P. D. D. Schwindt, V. Shah, H. G. Robinson, L. Hollberg, and J.Kitching, "Atomic vapor cells for chip-scale atomic clocks with improved long-term frequency stability," *Opt. Lett.* Vol. 30, p. 2351, 2005.

[20] Ke Deng, Tao Guo, Juan Su, Dengzhu Guo, Xinyuan Liu, Lu Liu, Xuzong Chen, and Zhong Wang, "Full hyperfine frequency modulation in the implementation of coherent population trapping atomic clocks ," Physics Letters A 373, 1130, 2009.

[21] K. Deng, T. Guo, D. W. He, X. Y. Liu, D. Z. Guo, X. Z. Chen, and Z. Wang, "Effect of Buffer gas ratios on the Relationship between cell temperature and frequency shifts of the coherent population trapping resonance," Applied Physics Letters 92, 211104, 2008.

Laser-pumped double-resonance clock using a micro-fabricated cell

M. Pellaton, C. Affolderbach, G. Mileti
Laboratoire Temps-Fréquence (LTF)
Université de Neuchâtel
Neuchâtel, Switzerland
gaetano.mileti@unine.ch

Y. Pétremand, N.F. de Rooij
Institute of Microengineering (IMT)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland

***Abstract*—In view of a novel miniature atomic clock with low power consumption, we investigate on the potential for a laser-pumped double-resonance atomic clock based on a micro-fabricated Rb vapor cell. We obtain a clock short-term stability of $2\times10^{-11}\,\tau^{-1/2}$ and the stability stays below 1×10^{-11} up to $\tau = 10^4$s, which demonstrates the feasibility of the approach.**

I. INTRODUCTION

Developments towards miniaturized or even chip-scale atomic clocks (CSAC) [1] up to today were mostly based on the Coherent Population Trapping (CPT) scheme [2]. This all-optical technology has the advantages of being compact with low power consumption. The alternative Double-Resonance (DR) scheme [3] has been widely abandoned for arguments of power consumption and size of the microwave cavity required. However, implementation of laser optical pumping can replace the bulky and power-consuming discharge lamp [4, 5], and micro-fabrication has opened new ways in microwave structure design [6]. These facts make it interesting to investigate further on the DR clock scheme, which avoids several issues of the CPT approach such as higher signal background or the difficulty to control precisely the spectrum of the modulated laser over time. It has also been demonstrated that DR short term stability is roughly 5 times better than CPT's in comparable experimental conditions [7].

II. CELL PRODUCTION

The fabrication process for the micro-fabricated cell used in this study is similar to the anodic bonding technology reported previously [8] and is described in [9]. First, a silicon wafer is "drilled" by deep reactive ion etching (DRIE), and then is anodically bonded to a glass wafer. This wafer stack is diced to obtain preforms, that are then handled individually. Rb filling is done with a commercially available Rb dispenser from SAES getters, Italy. Finally cell sealing is achieved by anodic bonding of a second glass plate, performed in a buffer-gas atmosphere with appropriate gas mixture and pressure.

This work was supported by the Swiss National Science Foundation (Sinergia grant CRSI20-122693/1) and the European Space Agency (ESA).

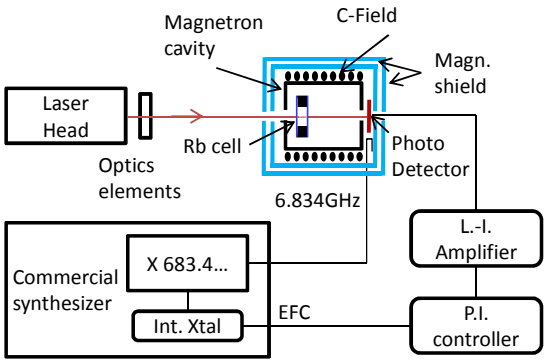


Figure 1. Experimental setup.

III. EXPERIMENTAL CLOCK SETUP

The setup is presented in Fig. 1. Pump light on the Rb D1 line is provided by a DFB laser stabilized to saturated-absorption lines obtained from a small in-house made glass-blown evacuated Rb reference cell. The light beam sent to the clock cell has 1mm diameter and is linearly polarized. The physics package is composed of the micro-fabricated cell placed in a compact magnetron-type microwave cavity [10] to orient and enhance the microwave field applied to the cell. The cavity is surrounded by a C-field coil that produces a static magnetic field parallel to the light beam to isolate the clock transition. Mu-metal shields are used to isolate the system from external magnetic fields. The 6.834 GHz microwave is produced from the 2.278 GHz output of a commercial synthesizer via a frequency tripler. This microwave frequency is modulated at $\approx$ 187 Hz to generate the error signal using a lock-in (L.-I.) amplifier. The clock loop is closed by sending the error signal to a Proportional/Integrator (P.I.) controller that acts on the EFC (Electronic Frequency Control) of the synthesizer.

IV. RESULTS

A typical clock signal is presented in Fig. 2. A contrast of nearly 2 % is achieved with a linewidth of 6.3 kHz (FWHM).

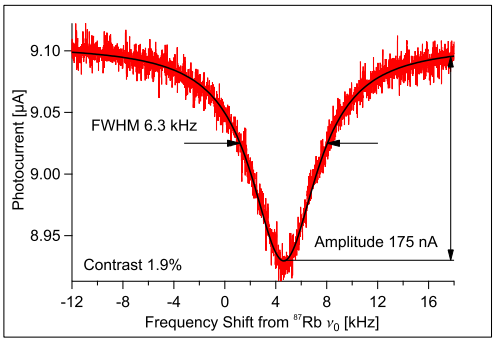


Figure 2. Typical double-resonance clock signal.

The corresponding error signal has a discriminator slope of D = 30 pA/Hz for a cell temperature of 373 K and a pump light power of 40 µW. Linewidth extrapolation to zero light intensity gives a value of 2.2 kHz. Using the formula [11]

$$\sigma_y(\tau) = \frac{N}{\sqrt{2}Df_0}\tau^{-\frac{1}{2}} \qquad (1)$$

where N is the total detection noise and f_0 the frequency of the microwave clock transition, a signal-to-noise limited short-term stability of $\sigma_y(\tau) = 1.8\text{x}10^{-11}\,\tau^{-1/2}$ and a shot-noise limit of $6\text{x}10^{-12}\,\tau^{-1/2}$ are calculated.

At medium and long-term timescales, the main degradations of clock stability result from fluctuations in temperature and light intensity. Fig. 3 shows the temperature shift of the clock frequency, consistent with predictions based on the buffer gases content of the cell [12]. A temperature-shift coefficient is measured as $8.2\text{x}10^{-10}$ /K (at the operating temperature of 373 K), which limits the clock stability to $4\text{x}10^{-12}$ at longer time scales. Straight-forward optimization of the buffer-gas mixture [13] will shift the point of minimized temperature sensitivity (currently around 76°C) towards the operating temperature and will result in an improvement of this stability limit by up to two orders of magnitude.

Intensity light-shift results for different laser frequencies (i.e. laser lock points) are shown on Fig. 4. With the laser locked to the $|F_g=2\rangle\longrightarrow|F_e=1\rangle$ sub-Doppler transition obtained from the reference cell (our experimental clock conditions) the intensity light-shift coefficient is $1.9\text{x}10^{-9}$ /µW, resulting in a clock instability contribution of $8\text{x}10^{-12}$ at 1000 s. This value can be reduced by a factor of 4 by locking the laser on

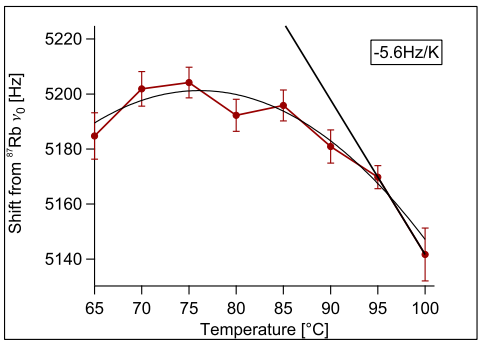


Figure 3. Measured temperature shift (data points) and parabolic fit to the data. The straight line is a linear fit at 100°C.

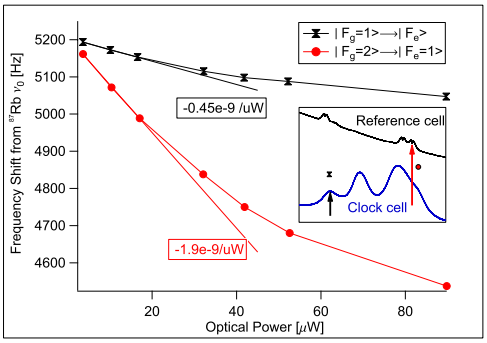


Figure 4. Light-shift measured for different laser frequencies. The inset shows the laser lock points used.

the buffer-gas broadened $|F_g=1\rangle\longrightarrow|F_e\rangle$ Doppler transition obtained on the micro-fabricated clock cell, but at the price of a reduced signal-to-noise ratio that degrades the short-term clock stability.

The measured clock frequency stability is $2\text{x}10^{-11}\,\tau^{-1/2}$ for τ < 100s, in excellent agreement with the signal-to-noise limit (see Fig. 5). The clock stability stays below $1\text{x}10^{-11}$ up to 10^4s averaging time, in agreement with the estimated limitations from the temperature and intensity light-shift effects. The main limiting factor here is the intensity light-shift.

V. CONCLUSIONS

We presented an experimental DR clock based on a micro-fabricated Rb vapor cell that already meets the typical stability requirements for a miniature atomic clock. The intensity light-shift is identified as the main factor limiting the long-term stability. Optimizing the cell's inner atmosphere (buffer gas and Rb isotopic ratio) will allow us to further improve on the clock stability. Combining these encouraging results with a new micro-fabricated microwave cavity will open the way for a novel extra small DR clock with competitive frequency stability.

ACKNOWLEDGMENT

We thank F. Gruet, P. Scherler, M. Durrenberger (all at UniNE-LTF) and the CSEM clean room staff for their contributions.

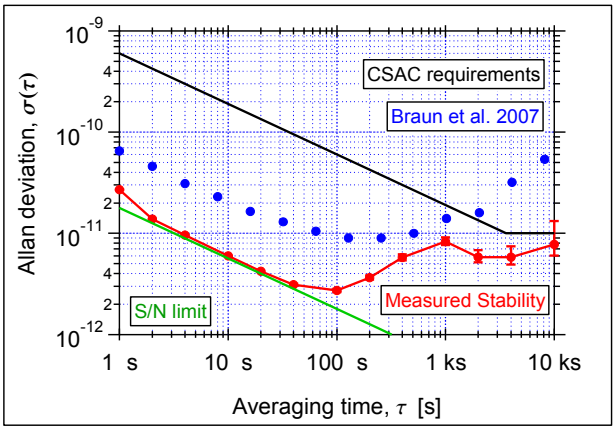


Figure 5. Measured Allan deviation of the clock, compared to other relevant clock data reported.

REFERENCES

[1] S. Knappe, "MEMS Atomic Clocks," Comprehensive Microsystems, vol. 3, pp. 571–612, September 2007.

[2] J. Vanier, "Atomic clocks based on coherent population trapping: a review," Appl. Phys. B, vol. 81, pp. 421–442, July 2005.

[3] J. Vanier and C. Mandache, "The passive optically pumped Rb frequency standard: the laser approach," Appl. Phys. B, vol. 87, p. 565–593, May 2007.

[4] J. Camparo, "The rubidium atomic clock and basic research," Physics Today, vol. 60, no. 11, pp. 33–39, November 2007.

[5] L. Guo-Bin, Z. Feng, and G. Si-Hong, "Study of a low power dissipation, miniature laser-pumped rubidium frequency standard," Chinese Physics B, vol. 18, no. 9, pp. 3839–3843, September 2009.

[6] A. M. Braun, et al., "RF-interrogated end-state chip-scale atomic clock," in 39th Annual Precise Time and Time Interval (PTTI) Meeting, 2007.

[7] R. Lutwak, D. Emmons, W. Riley, and R. M. Garvey, "The chip-scale atomic clock – coherent population trapping vs. conventional interrogation," in 34th Annual Precise Time and Time Interval (PTTI) Meeting, 2002.

[8] L.-A. Liew, S. Knappe, J. Moreland, H. Robinson, L. Hollberg, and J. Kitching, "Microfabricated alkali atom vapor cells," Applied Physics Letters, vol. 84, no. 14, pp. 2694–2696, April 2004.

[9] J. Di Francesco et al.,"Evaluation of the frequency stability of a VCSEL locked to a microfabricated Rubidium vapour cell," Proc. SPIE, vol. 7720, p. 77201T, April 2010.

[10] C. Affolderbach, F. Droz, and G. Mileti, "Experimental demonstration of a compact and high-performance laser-pumped rubidium gas cell atomic frequency standard," IEEE Transactions on Instrumentation and Measurement, vol. 55, no. 2, pp. 429 – 435, April 2006.

[11] G. Mileti, "Study of the S/N performance of passive atomic clocks using a laser pumped vapor," in 9th European Frequency and Time Forum, D. Hauden, Ed., 1995, pp. 271–276.

[12] J. Vanier, R. Kunski, N. Cyr, J. Y. Savard, and M.Têtu, "On hyperfine frequency shifts caused by buffer gases: Application to the optically pumped passive rubidium frequency standard," J. Appl. Phys., vol. 53, no. 8, pp. 5387–5391, August 1982.

[13] T. Bandi, C. Affolderbach, C. Calosso, and G. Mileti, "High-Performance Laser-Pumped Rubidium Frequency Standard for Satellite Navigation," Electron. Lett., in press.

Temperature dependence of a Cs vapor cell clock: pressure shift, signal amplitude, light shift.

Olga Kozlova, Jean-Marie Danet, Stéphane Guérandel and Emeric de Clercq
LNE-SYRTE, Observatoire de Paris, CNRS-UMR 8630, UPMC,
61 avenue de l'Observatoire, 75014
Paris, France

***Abstract*—We present the investigation on the temperature dependence of the pressure (collisional) shift of the Cs clock transition in presence of buffer gases (Ne, N_2 and Ar). The temperature dependence of the signal amplitude and light shift are also investigated. From the obtained coefficients of temperature dependence it is possible to foresee the buffer gas mixture which cancels the temperature sensitivity at a given temperature. These results as well as the results of the investigation on the signal amplitude and the light shift can be useful for improving the frequency stability of the Cs vapor-cell clocks.**

I. INTRODUCTION

Compact and stable frequency standards are needed in many applications such as telecommunications or navigation. Alkali vapor-cell atomic clocks are currently used for this purpose. In such clocks a buffer gas is added to the alkali vapor in order to reduce the relaxation by collisions on the cell walls and to reduce the Doppler broadening by Dicke effect. However the collisions between alkali atoms and the buffer gas can shift the frequency of the clock transition by several kHz per torr, and this shift is temperature dependent. This effect can limit the long term frequency stability of cell clocks and indeed has limited the stability of our clock prototype which uses a pulsed coherent population trapping (CPT) scheme. That is why we have undertaken a metrological study of these shifts for some buffer gases of interest, with the aim of reducing or cancelling the temperature dependence.

We have measured the pressure and temperature coefficients of the clock frequency shift induced by three buffer gases: Ne, N_2, Ar, see Section II. This allows to foresee the composition of a gas mixture canceling the frequency temperature sensitivity at the wished working temperature. It is experimentally verified on a N_2-Ar mixture. The working temperature can be the one for which the signal amplitude is maximal. We show in section III that this temperature also depends on the laser intensity. In section IV we show that the light shift is also temperature dependent.

This work is partly supported by Direction Générale de l'Armement (DGA) under contract n° 2009 34 0052. O. K. was supported by DGA.

II. PRESSURE SHIFT AND TEMPERATURE DEPENDENCE CANCELATION

A. Pressure and temperature coefficients

The frequency shift, induced by alkali-buffer gas collisions, can be written [1]:

$$\Delta\upsilon = P_0\left(\beta + \delta(T - T_0) + \gamma(T - T_0)^2\right) \quad (1)$$

where P_0 is the buffer gas pressure in the cell at the reference temperature T_0, β is the pressure coefficient ($Hz \cdot Torr^{-1}$), δ is the linear temperature coefficient ($Hz \cdot Torr^{-1} \cdot K^{-1}$), γ is the quadratic temperature coefficient ($Hz \cdot Torr^{-1} \cdot K^{-2}$) and T is the buffer gas temperature.

Following our previous works [2, 3], we present here the results of the pressure and temperature coefficients for three buffer gases (Ne, N_2 and Ar). In order to avoid any bias we use 4 cells filled with Ne, 5 with N_2 and 5 with Ar made by three different suppliers. The cells are filled with different pressures ranging between 15 and 180 torr. The actual pressure in the cell after sealing is unknown, so the accuracy on the frequency shift rates measurements is generally limited by the uncertainty on the gas pressure. We have measured the pressure of the buffer gas in each cell by measuring the frequency shift of the optical transitions of the Cs D_1 line [3]. The pressure is then calculated using the published optical shift rates [4, 5] corrected for the temperature (for more details see [6]). Once the pressure of the buffer gas is determined, the temperature dependence of the clock frequency is measured for each cell with our CPT pulsed Cs clock described in [3].

TABLE I. EXPEREMENTAL PRESSURE AND TEMPERATURE COEFFICIENTS FOR CS WITH TRHEE BUFFER GASES

Buffer gas	**Pressure coefficient β, ($Hz \cdot Torr^{-1}$)**	**Linear temperature coefficient δ, ($Hz \cdot Torr^{-1} \cdot K^{-1}$)**	**Quadratic temperature coefficient γ, ($mHz \cdot Torr^{-1} \cdot K^{-2}$)**
Ne	686 ± 14	0.266 ± 0.006	-1.68 ± 0.05
N_2	922.5 ± 4.8	0.824 ± 0.006	-2.51 ± 0.03
Ar	-194.4 ± 1.6	-1.138 ± 0.010	0.0 ± 0.3

In Table I, measured pressure and temperature coefficients for Cs clock transition are shown. The accuracy of the present work is still limited by the uncertainty on the buffer gas pressure. We have precisely determined the quadratic coefficients for N_2 and Ne. The coefficient of Ar is much smaller and we have estimated an upper limit.

It is important to notice that gases with a quadratic coefficient exhibit an inversion temperature, where the temperature dependence vanishes, $T_{inv} = T_0 - \delta/2\gamma$. Ne buffer gas shows a strong quadratic temperature coefficient. The inversion temperature computed from the coefficients of the fitted polynomials for each cell, in order to avoid pressure additional uncertainties, is T_{inv} = (79 ± 3) °C, which confirms previous evaluations [7, 8]. It makes Ne very attractive for use in the Cs chip scale clocks which needs a high working temperature. The calculated inversion temperature with N_2 is (164 ± 13) °C, while no inversion temperature can be given for Ar.

B. Buffer gas mixtures

One can also cancel the temperature dependence at a chosen temperature by a mixture of gases with temperature coefficients of opposite signs [9, 10, 11, 12]. From the obtained coefficients for three buffer gases the mixture composition canceling temperature dependence at a given temperature can be calculated:

$$r(P_2/P_1) = -\frac{\delta_1 + 2\gamma_1(T_{inv} - T_0)}{\delta_2 + 2\gamma_2(T_{inv} - T_0)} \qquad (2)$$

where $r\ (P_2/P_1)$ is the pressure ratio of two buffer gases, $\delta_{1,2}$ and $\gamma_{1,2}$ are the temperature coefficients for buffer gases, T_0 is the reference temperature (0°C), T_{inv} is the inversion temperature.

We have calculated the equation (2) for three possible buffer gas compositions (Ar-N_2, Ar-Ne and N_2-Ne).

- $r(P_{Ar}/P_{N2}) = 0.724(12) - 0.00441(9) \cdot (T_{inv} - T_0)$
- $r(P_{Ar}/P_{Ne}) = 0.234(9) - 0.00295(11) \cdot (T_{inv} - T_0)$
- $r(P_{N2}/P_{Ne}) = -\dfrac{0.266(6) - 0.00336(10) \cdot (T_{inv} - T_0)}{0.824(6) - 0.00502(6) \cdot (T_{inv} - T_0)}$

Note that in case of mixture containing Ar buffer gas (for which the γ coefficient is zero) the equation (2) could be simplified to $r = -[\delta_1/\delta_2 + (\delta_1/2\ \gamma_2)\cdot(T_{inv} - T_0)]$. The Fig. 1, 2 and 3 illustrate the calculated dependences.

In order to verify the calculated values several cells with different Ar-N_2 mixture were carefully filled in the laboratory and studied. For each buffer gas mixture cell the inversion temperature was measured. Fig. 4 shows the temperature dependence of the clock frequency in the Ar-N_2 mixture cells with different proportion of two gases.

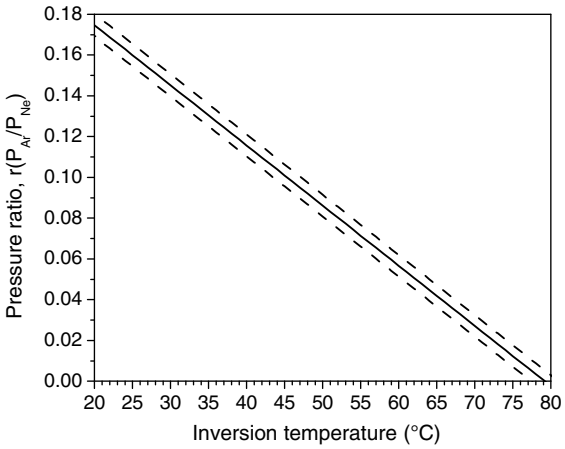


Figure 2. Ar-Ne pressure ratio versus inversion temperature. Black solid line - calculated values from the obtained temperature coefficients, black dashed lines – 1σ uncertainty for the calculated values.

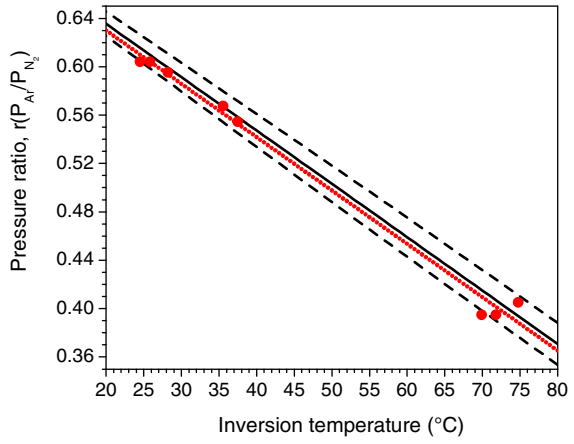


Figure 1. Ar-N_2 pressure ratio versus inversion temperature. Black solid line - calculated values from the temperature coefficients, black dashed lines – 1σ uncertainty for the calculated values. Red circles – experimentally measured inversion temperatures T_{inv} for 8 cells. Red dotted line – fitted line.

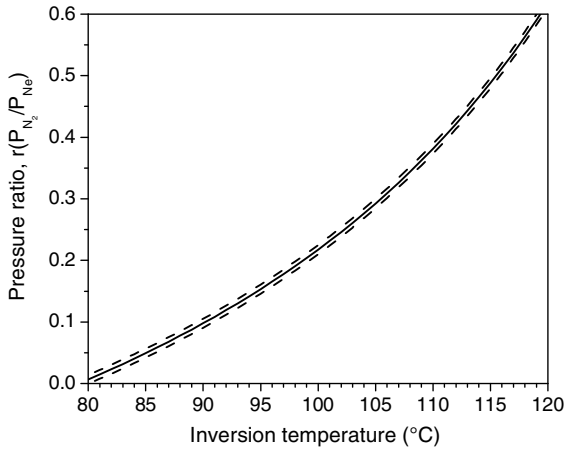


Figure 3. N_2-Ne pressure ratio versus inversion temperature. Black solid line - calculated values from the obtained temperature coefficients, black dashed lines – 1σ uncertainty for the calculated values.

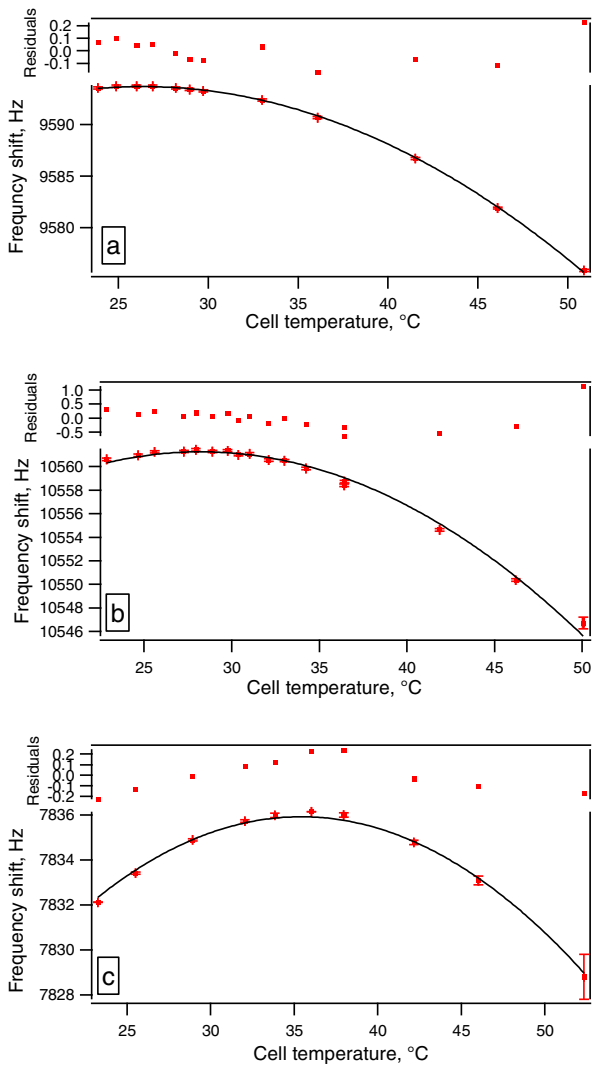


Figure 4. Temperature dependence of the clock frequency in Ar-N_2 Cs cells with different compositions: a) r = 0.608; b) r = 0.595; c) r = 0.568. Experimental points are fitted with N_2 pressure and the ratio r as free parameters, the pressure coefficients and temperature coefficients for both gases as fixed.

The comparison between calculated and experimentally measured values is presented on Fig 1.

The expression given by the fit of the experimental points is $r = 0.718(9) - 0.00441(18)\cdot(T_{inv} - T_0)$. The good agreement between calculated and experimental values confirms the validity of the measured microwave shift coefficients.

The coefficients values of Table I and the equations for the mixture gas composition can be a useful guideline for manufacturers of Cs vapor cell frequency standards.

III. SIGNAL AMPLITUDE

In order to determine the optimal working temperature the signal amplitude was measured as a function of the cell temperature, see Fig. 5. At low temperature the signal increases with the temperature, due to the increase of the Cs density. At higher temperatures the signal decreases because the medium becomes optically thick and also because of the increasing relaxation rate. We have notice that the maximum signal amplitude of the CPT resonance with cell temperature depends on the laser intensity (Fig. 5). Such dependence was observed in both CW and Ramsey interrogation modes.

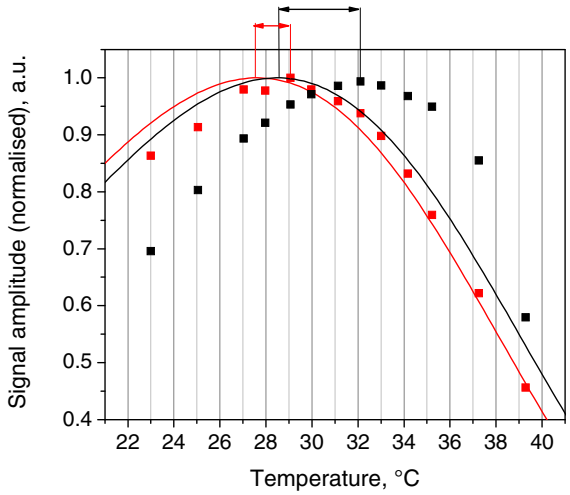


Figure 5. Signal amplitude versus cell temperature for two different laser intensities. Red and black points – experimental data for 60μW/cm^2 and 300μW/cm^2, respectively. Red and black curves – theoretical model for for 60μW/cm^2 and 300μW/cm^2, respectively. Red and black arrows show the discrepancy between theory and experiment. Cell with Ar-N_2 mixture, pressure ratio r (P_{Ar}/P_{N2}) =0.595. CW mode of interrogation.

A theoretical model was developed in order to explain this dependence. It is based on the calculation of two absorption coefficients: one on CPT resonance and the other off CPT resonance, both at optical resonance. Contrary to the model developed in [13], there is no empirical parameters (such as Cs density or Rabi frequencies) adjusted in our model. The developed model and the experimental data are in qualitative agreement: the temperature for which the signal is maximal increases with the laser intensity. We find better agreement for lower intensity (as in [13]). For the lower laser intensity (60μW/cm^2) the experimental value of the temperature for the maximal amplitude is 29.0°C, while the theoretical value is 27.7°C. For the higher laser intensity (300μW/cm^2) the experimental value is 32.0°C and the theoretical value is 28.5°C. Note that our model is sensitive to the value Cs density, which is known with maybe ~10% uncertainty [14, 15, 16], and this uncertainty limits our model. A fitted value of the Cs density is probably necessary. However this does not explain the unexpected large variation of the temperature of the maximum with the laser intensity. Further investigations are needed.

IV. LIGTH SHIFT

The light shift is one of the major effects, with the buffer gas pressure shift, which can limit the long-term stability of alkali vapour cell clocks [1]. It must then be well controlled. We have measured the light shift in cells with different buffer

gases, as a function of the laser intensity and the cell temperature. We have noticed that the light shift is sensitive to the cell temperature: the slope of the light shift versus laser intensity changes value and even changes sign (Fig. 6). So for a given temperature of the cell it is possible to have a zero slope and to cancel the light shift. Here we present the results for N_2 buffer gas, but the similar behavior was observed in cells with Ar and with Ar-N_2 mixtures.

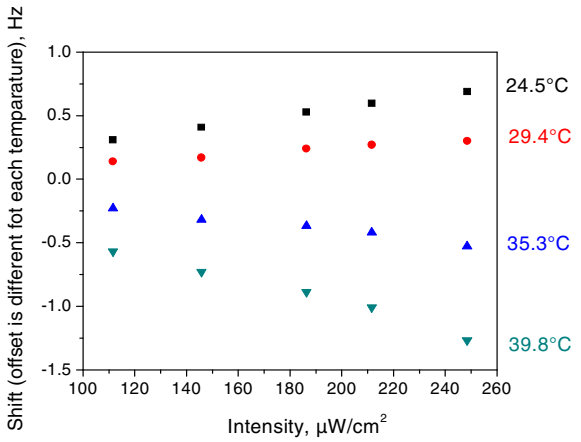


Figure 6. Light shift for different temperatures. For better visibility a different offset is applied for each temperature. Each color corresponds to the temperature (from 24.5°C to 39.8°C). Ramsey mode (T_r=4ms) interrogation. Cell 30 torr N_2.

Fig. 7 shows the slope of the light shift versus cell temperature. The slope should be null for a temperature of about 31°C in this case.

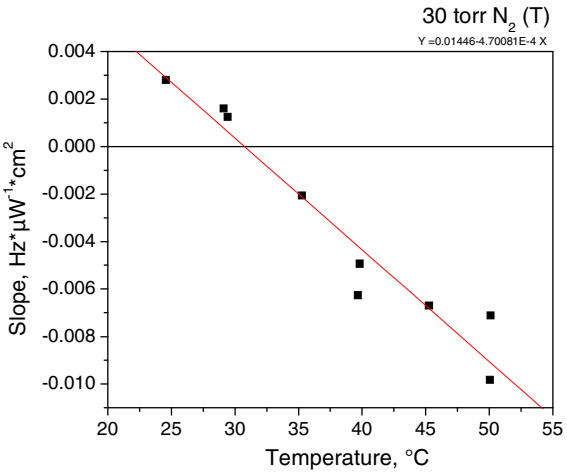


Figure 7. Slope of the light shift for different temperatures. Ramsey mode (Tr=4ms) interrogation. Cell 30 torr N_2.

This sensitivity can be explained by the temperature dependence of the detuning of the optical transitions. Indeed, the light shift is expected to have a dispersion line-shape with the detuning of the optical resonance [17], and the optical shift induced by the buffer gas is temperature dependent. But in these experiments the laser frequency was locked on the saturated absorption spectrum in another Cs cell, it was then fixed, independent of the temperature of the buffer gas-Cs cell of the clock.

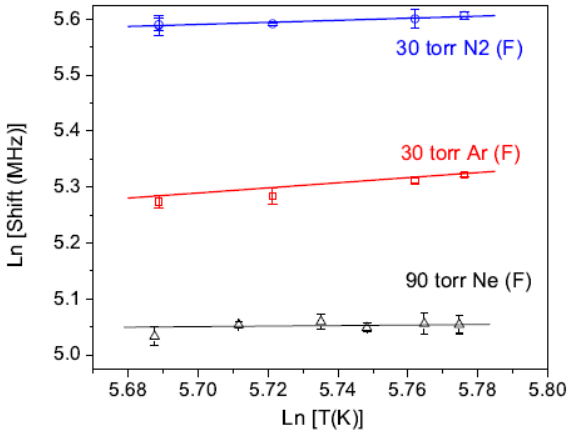


Figure 8. Temperature dependence of the shift of the D_1 line in presence of Ne, N_2 and Ar buffer gas, logarithmic units.

The temperature dependence of the optical shift induced by the buffer gas can be written [18, 19]:

$$\frac{S_m}{S_r} = \left(\frac{T_m}{T_r}\right)^n, \qquad (3)$$

where S_m and S_r are the shifts measured at the temperature T_m and T_r respectively, the value of n is determined by the Cs-buffer gas interactions and the energy dependence of the cross section. We have measured the temperature dependence of the optical shift for three buffer gases, Ne, N_2 and Ar, see Fig. 8. The obtained values for n are: 0 ± 0.2 (Ne), 0.18 ± 0.03 (N_2), 0.56 ± 0.07 (Ar) [6].

When n is known, we can plot again Fig. 7, but as a function of the laser detuning instead of temperature. Fig. 9 shows the slope of the light shift versus optical detuning. The value of the optical detuning is calculated for each temperature according to (3).

Note that the zero slope of the light shift does not correspond to the zero optical detuning (the zero slope corresponds to the optical detuning 4 – 5 MHz). In fact the absorption spectrum in the cell with buffer gas is broadened by Cs-buffer gas collisions. In our case (30 torr N_2) this broadening leads to a Lorentzian width in the Voigt profile of about 460 MHz. For the temperatures used in the experiment the Doppler width is about 360 MHz. As the light shift depends also on the width of the optical transition, another temperature dependence is induced by the width variation. Moreover the half of the hyperfine splitting of the excited state is 584 MHz, so the overlapping of the profiles of two excited states is possible. This could explain why the light shift cancels for a detuning different of zero.

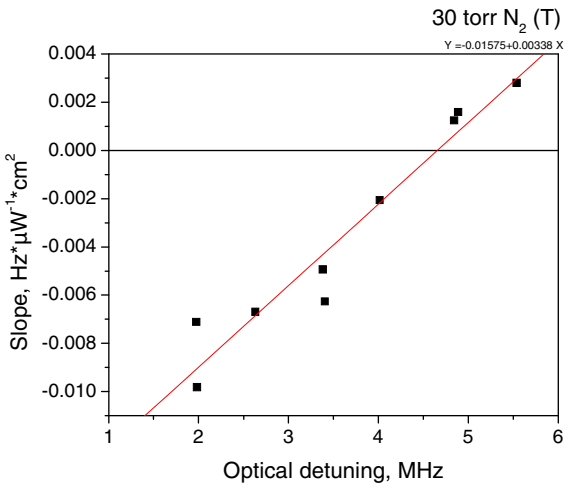


Figure 9. Slope of the light shift for the optical detuning. The value of the optical detuning is calculated for each temperature from the temperature dependence of the optical shift. Ramsey mode (Tr=4ms) interrogation. Cell 30 torr N_2.

Besides of the temperature dependence of the optical transitions, it is probably necessary to take also into account the increasing absorption of the incident laser intensity and the effect of optical thickness. Further studies are needed for a full theoretical explanation. However it is worth to note that the light shift can be canceled by tuning the laser frequency, or also by tuning the cell temperature.

V. CONCLUSION

We have reported here results of the measurement of the pressure and temperature coefficients for Cs ground state transition in presence of Ne, N_2 and Ar buffer gases measured with CPT pulsed clock. From the obtained coefficients the composition of mixtures of buffer gases (Ar-N_2, Ar-Ne and N_2-Ne) is calculated as a function of the wished temperature canceling the temperature sensitivity of the Cs clock frequency. This prediction was experimentally checked and confirmed in the case of an Ar-N_2 mixture. The obtained results will allow one to choose appropriate gas or a mixture of buffer gases to cancel the temperature dependence of the Cs vapor cell clock at a working temperature.

The working temperature can be chosen to maximize the signal amplitude. We have observed that the cell temperature of maximum signal amplitude of the CPT resonance depends on laser intensity. A preliminary theoretical model qualitatively describes this behavior, but with a much smaller temperature variation.

The temperature dependence of the light shift was observed. This dependence could be explained by the temperature dependence of the collisional shift of the optical transitions and then the detuning of the laser. This temperature dependence was measured for three buffer gases Ne, N_2 and Ar.

The results of the investigation on the signal amplitude and light shift would be useful to optimize the working parameters of the CPT pulsed clock. Further investigations are in progress.

ACKNOWLEDGMENT

We are pleased to acknowledge P. Bonnay and A. Gérard for very careful manufacturing of Cs cells. We are grateful to Michel Lours and the SYRTE electronic team for their help.

REFERENCES

[1] J. Vanier and C. Audoin "The quantum physics of atomic frequency standards", IOP Publishing, Bristol and Philadelphia, Ch. 3, 1989.

[2] O. Kozlova, R. Boudot, S. Guérandel, E. De Clercq, "Cs buffer gas frequency shift: method and measurements", in Proc. Of the joint meeting EFTF-IFCS 2009, Besançon, France, 2009.

[3] O. Kozlova, R. Boudot, S. Guérandel, E. De Clercq, "Measurements of Cs-buffer gas collisional frequency shift using pulsed CPT interrogation", in Proc. EFTF 2010, Noordwijk, The Netherlands.

[4] G. A.Pitz, D. E. Wertepny, and G. P. Perram, "Pressure broadening and shift of the cesium D_1 transition by the noble gases and N_2, H_2, HD, D_2, CH_4, C_2H_6, CF_4 and ^{3}HE", Phys.Rev. *A*, vol.80, 062718 (2009)

[5] A.Andalkar and R.B.Warrington, "Higt-resolution measurement of the pressure broadening and shift of the Cs D_1 and D_2 lines by N_2 and He buffer gases", Phys.Rev A, vol.65, 032708 (2002)

[6] O. Kozlova, S. Guérandel, E. De Clercq, "Temperature and pressure shift of the Cs clock transition in presence of buffer gases Ne, N_2 Ar", Phys.Rev A, unpublished.

[7] O. Kozlova, R. Boudot, S. Guérandel, and E. de Clercq, "Temperature dependence cancellation of the cs clock frequency in the presence of ne buffer gas," (2011), IEEE Trans. Instrum. Meas., in press.

[8] D. Miletic, P. Dziuban, R. Boudot, M. Hasegawa, R. K. Cutani, G. Mileti, V. Giordano, and C. Gorecki, "Quadratic dependence on temperature of Cs 0–0 hyperfine resonance frequency in single Ne buffer gas microfabricated vapour cell", Electron. Lett. 46, 1069 (2010).

[9] J. Vanier, R. Kunski, N. Cyr, J. Y. Savard, and M. Têtu, "On hyperfine frequency shifts caused by buffer gases: Application to the optically pumped passive rubidium frequency standard", J. Appl. Phys., 53, 5387, 1982.

[10] R. Lutwak D. Emmons, W. Riley, and R. M. Garvey, The chip-scale atomic clock – coherent population trapping vs. conventional interrogation, Proc. 34th Annual Precise Time and Time Interval (PTTI) Meeting, 3-5 December 2002, Reston, Virginia, USA (U.S. Naval Observatory, Washington, D.C.), pp. 539-550.

[11] S. Knappe et al., "Atomic vapor cells for chip-scale atomic clocks with improved long-term frequency stability", Opt. Lett., 30, 2351, 2005.

[12] R. Boudot et al;, "First-order cancellation of the Cs clock frequency temperature dependence in Ne-Ar buffer gas mixture", Optics Express, 19, 3106, 2011.

[13] S. Knappe, J. Kitching, L. Hollberg, R. Wynands "Temperature dependence of coherent population trapping resonances", Appl. Phys. B 74, 217-222 (2002)

[14] J. B. Taylor and I. Langmuir, "Vapor pressure of Caesium by the Positive Ion Method", Phys. Rev. 51, 753, 1937.

[15] C. B. Alcock, V. P. Itkin, and M. K. Horrigan, "Vapor pressure of the Metallic Elements", Canadian Metallurgical Quarterly, 23, 209, 1984.

[16] Y.-Y. Jau, N. N. Kuzma, and W. Happer, Phys. Rev. A, 69, 061401, 2004.

[17] J. Vanier, A. Godone, and F. Levi, "Coherent population trapping in cesium:Dark lines and coherent microwave emission", Phys. Rev. A, 58, 2345, 1998.

[18] J. F. Kielkopf, "Predicted alkali collision broadening by nobel gases based on semiempirical potentials", J. Phys. B 9, L547 (1976).

[19] M. V. Romalis, E. Miron, and G. D. Cates, "Pressure broadening of Rb D_1 and D_2 lines by ^{3}He, ^{4}He and Xe: Line cores and near wings", Phys. Rev. A 56, 4569 (1997).

Status of a compact cold-atom CPT frequency standard

F.-X. Esnault, E. A. Donley, and J. Kitching
Atomic Devices and Instrumentation Group,
NIST Time and Frequency Division,
325 Broadway, Boulder CO 80305, USA
Email: fesanult@nist.gov

E. N. Ivanov
School of Physics
Univ. of Western Australia
35 Stirling Hwy, Crawley WA 6009, Australia

***Abstract*—We describe the progress towards the realization of a cold-atom frequency standard based on coherent population trapping (CPT). We explain our particular CPT configuration and give details on the experimental setup.**

I. Introduction

The performance of atomic clocks based on coherent population trapping (CPT) has improved considerably over the last few years in terms of stability and reliability [1], [2], [3]. Frequency stabilities in the low $10^{-10}\ \tau^{-1/2}$ range are now realized within package volumes of a few tens of cubic centimeters [3]. Nevertheless, a significant frequency drift ($\geq 10^{-12}$/day) is observed in these devices. The use of wall coatings and buffer gases, which are required for high-contrast narrow-linewidth CPT resonances, degrades the long-term stability through cell aging and large temperature-dependent pressure shifts. Light shifts also contribute to the drift of CPT-based clocks.

To circumvent these issues, we began a study of a microwave CPT frequency standard based on laser-cooled atoms (CA-CPT). Cold atoms can provide good long-term performance, while the all-optical CPT interrogation enables good stability through narrow atomic resonances [1] and allows miniaturization of the physics package by eliminating the microwave cavity. This project aims at reaching a frequency stability of better than $10^{-11}\ \tau^{-1/2}$ and an accuracy of around 10^{-13} within a miniaturized physics package. Our work focuses both on the experimental realization and the metrological evaluation of a CA-CPT clock and on efficient ways of producing a cold atom sample in a millimetric volume (see [4]). Consideration of the expected systematics, including light shifts, second-order Zeeman effects and Doppler effects are given in [5].

II. CPT interrogation

Many schemes and configurations with differing atom types, optical excitation lines, light polarization, etc., have been proposed to improve the metrological properties of a clock based on a CPT resonance [1]. Improvements can be made, for example, by use of a polarization configuration that does not involve so-called "trap states" (extreme m_F sublevels) [6], [7], [8], and by use of a larger intensity in the CPT beams to increase the pumping rate into the "dark state".

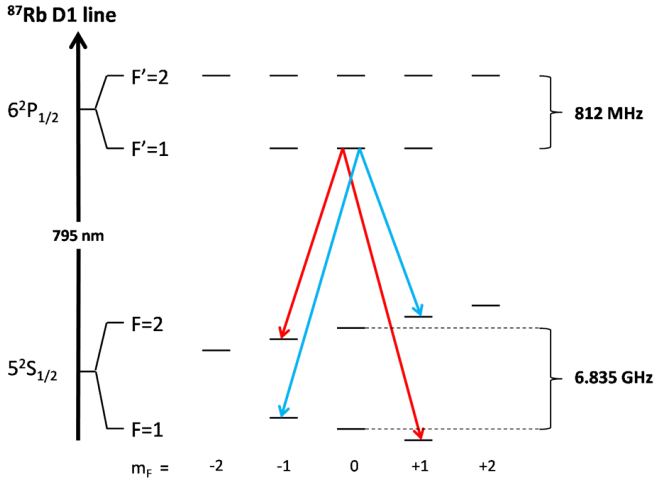


Fig. 1. Energy diagram of the ^{87}Rb D1 line at 795 nm and transitions involved in the $lin||lin$ configuration.

However, this latter option has two major drawbacks. A large intensity broadens the resonance, so that the high contrast is quickly counter-balanced by a broader linewidth $\Delta\nu$; secondly, it increases the magnitude of the second-order Stark shift (light shift).

The sequence and configuration that we propose, a pulsed Ramsey-like sequence on a cold ^{87}Rb sample combined with a $lin||lin$ configuration [8], can achieve high contrast without being limited by power broadening and large light shifts. The advantages of this approach are described in the following.

A. *lin parallel lin configuration*

The $lin||lin$ configuration was first proposed and studied by Taichenachev *et al.* in [8] and then investigated further in [9], [10], [11], [12], [13]. The CPT resonance is the superposition of two Λ systems (see Fig. 1) connecting the ground state sublevels $|F=1, m_F=\pm 1\rangle$ and $|F=2, m_F=\mp 1\rangle$ through the excited level $|F'=1, m_F=0\rangle$. Because this configuration does not involve any "trap state" we expect an increased contrast. Indeed, contrasts of up to 40 % have been reported in [8] and of 25 % in [11]. From an experimental perspective,

this scheme is well suited for a frequency modulated diode laser, as both beams have the same polarization.

B. Pulsed interrogation

Ramsey-like pulsed interrogation with CPT interactions was first performed in the late eighties with a thermal sodium beam and spatially separated interactions [14], [15], [16]. More recently, several groups have used the technique in alkali vapor cells [7], [17] and on cold atoms [18] with two optical pulses separated in time by T_R. The first pulse pumps the atoms into the "dark state" (duration τ_p); then the atoms evolve freely for a period of T_R, after which the second pulse (duration τ_m) measures the accumulated phase. An absorption measurement gives rise to an interference pattern typical of Ramsey interrogation. In this scheme, because the atoms evolve most of the time in the dark, the linewidth $\Delta\nu$ no longer depends on the atomic saturation but is Fourier limited and scales as $1/(2T_R)$. For a miniaturized cold atom clock, the interrogation period is limited by the atomic free fall under gravity to about 10 ms duration ($500\ \mu m$ displacement), leading to a 50 Hz linewidth.

III. Experimental Setup

The experimental apparatus has been fully assembled. This section presents some aspects of our designs and ideas concerning the vacuum chamber and the optical bench.

A. Vacuum chamber

We use a vacuum chamber produced by *Cold Quanta*[1] for atom-chip Bose-Einstein condensation experiments. The chamber is shown in Fig. 2 and described in detail in [19], [20]. The main chamber, where a 3D optical molasses is realized, is separated from a 2D-MOT zone containing a Rb dispenser by two pinholes of 1 mm diameter, ensuring differential pumping. The vacuum level in the main chamber is currently below 10^{-7} Pa, thanks to a 2 l/s ion pump and a getter pump.

The whole chamber is mounted horizontally to realize CPT interrogation parallel or perpendicular to gravity. A set of three pairs of Helmholtz coils is used to cancel the residual magnetic field and to define a proper quantization axis along the CPT beams.

B. Laser cooling setup

We use a single diode laser to feed both the 2D-MOT and the 3D-optical molasses regions. We have about 16 mW of power after a double-pass AOM and fiber coupling. The light is tuned about 2 Γ below the $|F=2\rangle \rightarrow |F'=3\rangle$ transition of the ^{87}Rb D2-line, where Γ is the resonance linewidth. The 2D-MOT is generated by use of two elliptical retroreflected beams of area 12 $\times$ 25 mm^2 with 5 mW of power in each beam. The quadrupole field is produced by two permanent magnets (12 G/cm). A 1 mW push beam (1 cm diameter) provides a 2D$^+$-MOT geometry. For the optical molasses we use retro-reflected 3 mm beams (diameter @ $1/e^2$) with about 1.5 mW in each pair. With this setup, we are able to capture a few $\times 10^6$ atoms per second, which is currently limited by laser power. The fluorescence is collected with a fresnel lens (NA = 1) with an efficiency of about 4 % on a high-gain photodiode.

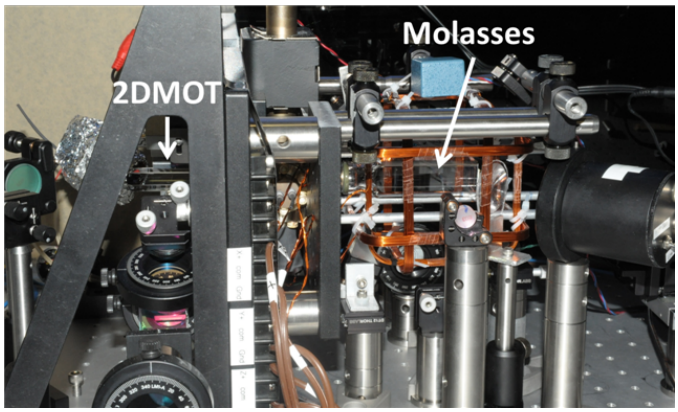


Fig. 2. Picture of the current apparatus.

C. Optical PLL setup

In many cases, the bichromatic laser field needed for excitation of the CPT resonances is generated by modulating the injection current of a diode laser. This produces a comb of optical frequencies of varying amplitudes. The unwanted residual sidebands degrade the clock performance by contributing to detection noise and the AC Stark shift, but they do not contribute to the CPT signal.

To avoid spurious sidebands in the CPT light we use two independent phase-locked diode lasers. We phase-lock two commercial DFB/DBR lasers at 795 nm with measured Lorentzian linewidths of 1 MHz and 0.6 MHz respectively. The optical beat note is realized around the ^{87}Rb hyperfine splitting at 6.8 GHz. The design of the optical phase lock loop is shown in Fig. 3. The amplified optical beat note is mixed with a 6.65 GHz signal coming from a microwave synthesizer (referenced to a 10 MHz H-Maser signal), filtered, amplified, and compared with a 150 MHz signal coming from a direct digital synthesizer inside a digital phase frequency detector (DPFD AD9901). This DPFD provides an error signal that is filtered and fed back on the slave laser's current. By use of high-speed electronics and through careful measurement of the FM response of both lasers to determine the best master/slave combination, we are able to achieve phase-locked operation with a bandwidth of about 8 MHz. We measured the fractional power in the coherent carrier to be about 70 %, which corresponds to a phase error variance of $0.35\ \mathrm{rad}^2$ over a 20 MHz integration bandwith. We measured the phase noise of the optically generated signal by comparing it to an external RF synthesizer, as shown in Fig. 3. The results are shown in

[1]Products or companies named here are cited only in the interest of complete scientific description, and neither constitute nor imply endorsement by NIST or by the US government. Other products may be found to serve equally as well.

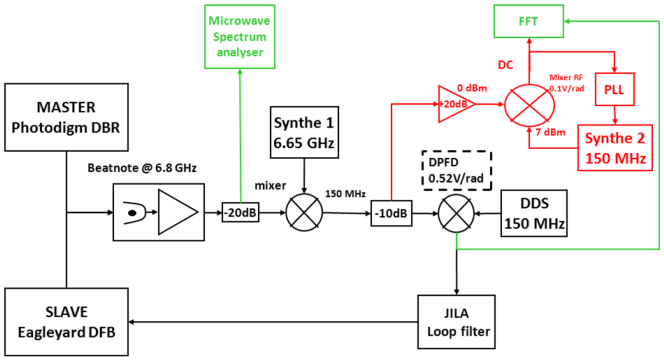


Fig. 3. Optical phase-lock-loop design

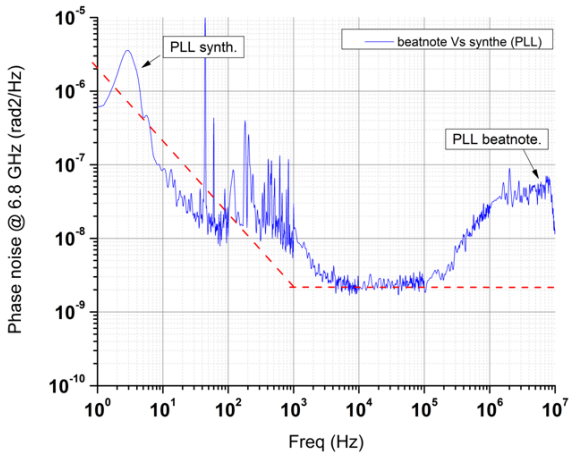


Fig. 4. Measured phase noise of the optically carried 6.8 GHz signal. The small bump around a few hertz is an artifact due to the very narrow band PLL used in the phase noise measurement.

Fig. 4. Out to several kilohertz, the phase noise is determined by joint contribution of the DDS and Synth2, and then we observe a noise floor corresponding to the PLL residual noise. Above 100 kHz the increase in the noise spectrum is due to the loss of PLL gain. With such phase noise, we estimate that the Dick effect will degrade the clock stability at a level of $5 \times 10^{-12}\tau^{-1/2}$. Because the sensitivity function of a pulsed CPT interrogation has yet to be derived, we used the appropriate function for a standard microwave Ramsey interrogation. The parameters used are a cycle duration of 100 ms, $T_R = 10$ ms and $\tau_m = \tau_p = 50$ μs.

ACKNOWLEDGMENT

The authors thank Terry Brown and James Fung-A-Fat from JILA's Electronic Shop for developing the PLL circuit. F-X Esnault is grateful to the Delegation Generale pour l'Armement and NIST for providing the post-doc fellowship. This paper is a contribution of the National Institute of Standards and Technology, an agency of the U.S. government, and is not subject to copyright.

REFERENCES

[1] J. Vanier, "Atomic clocks based on coherent population trapping: a review," *Appl. Phys. B*, vol. 81, no. 4, pp. 421–442, 2005.

[2] S. Knappe, "Emerging topics : MEMS atomic clocks," *Comp. Microsys.*, vol. 3, pp. 571–612, 2007.

[3] R. Lutwak, "The Chip-Scale atomic clock - recent developments," in *Proc. 2009 Joint Meeting IEEE Int. Freqency Control Symp. and EFTF Conf.*, Besancon, France, 2009, pp. 573–577.

[4] Tara Cubel-Liebisch, Eric Blanshan, Elizabeth Donley and John Kitching, "MOT loading enhancement with stimulated forces" *in Proceeding of the 2011 IFCS-EFTF joint meeting, San Francisco, CA, USA*, 2011.

[5] F.-X. Esnault, E. Donley, and J. Kitching, "Towards a compact cold atom frequency standard based on coherent population trapping," in *in Proc. of 2010 IEEE International Frequency Control Symposium, Newport Beach, CA, USA* (2010), pp. 465-469.

[6] E. Arimondo, "Coherent population trapping in laser spectroscopy," in *Progress in Optics. Elsevier Science Publication B V, Amsterdam, The Netherlands*, vol. 35, 1996.

[7] T. Zanon, S. Guerandel, E. de Clercq, D. Holleville, N. Dimarcq, and A. Clairon, "High contrast ramsey fringes with Coherent-Population-Trapping pulses in a double lambda atomic system," *Phys. Rev. Lett.*, vol. 94, no. 19, 2005.

[8] A. V. Taichenachev, V. I. Yudin, V. L. Velichansky, and S. A. Zibrov, "On the unique possibility of significantly increasing the contrast of dark resonances on the D1 line of 87 Rb," *JETP Lett.*, vol. 82, no. 7, pp. 398—403, 2005.

[9] E. Breschi, G. Mileti, G. Kazakov, B. Matisov, R. Lammegger, and L. Windholz, "Study of the laser linewidth influence on "lin || lin" coherent population trapping," in *Proc. 2007 Joint Meeting IEEE Int. Freqency Control Symp. and EFTF Conf.*, 2007, pp. 617–622.

[10] E. Breschi, G. Kazakov, R. Lammegger, G. Mileti, B. Matisov, and L. Windholz, "Quantitative study of the destructive quantum-interference effect on coherent population trapping," *Phys. Rev. A*, vol. 79, no. 6, p. 063837, 2009.

[11] S. A. Zibrov, I. Novikova, D. F. Phillips, R. L. Walsworth, A. S. Zibrov, V. L. Velichansky, A. V. Taichenachev, and V. I. Yudin, "Coherent population trapping resonances with linearly polarized light for all-optical miniature atomic clocks," *0910.4703*, Oct. 2009.

[12] E. E. Mikhailov, T. Horrom, N. Belcher, and I. Novikova, "Performance of a prototype atomic clock based on lin||lin coherent population trapping resonances in Rb atomic vapor," *0910.5881*, Oct. 2009.

[13] K. Watabe, T. Ikegami, A. Takamizawa, S. Yanagimachi, S. Ohshima and S. Knappe, "High-contrast dark resonances with linearly polarized light on the D1 line of alkali atoms with large nuclear spin," *App. Opt.* vol. 48, pp. 1098-1103, 2009.

[14] P. R. Hemmer, S. Ezekiel, and J. C. C. Leiby, "Stabilization of a microwave oscillator using a resonance raman transition in a sodium beam," *Opt. Lett.*, vol. 8, no. 8, pp. 440–442, 1983.

[15] P. R. Hemmer, G. P. Ontai, and S. Ezekiel, "Precision studies of stimulated-resonance raman interactions in an atomic beam," *J. Optical Soc. Am. B*, vol. 3, no. 2, pp. 219–230, 1986.

[16] P. R. Hemmer, M. S. Shahriar, V. D. Natoli, and S. Ezekiel, "Ac stark shifts in a two-zone raman interaction," *J. Optical Soc. Am. B*, vol. 6, no. 8, pp. 1519–1528, 1989.

[17] G. Pati, K. Salit, R. Tripathi, and M. Shahriar, "Demonstration of Raman–Ramsey fringes using time delayed optical pulses in rubidium vapor," *Opt. Commun.*, vol. 281, no. 18, pp. 4676–4680, 2008.

[18] X. Chen, G. Q. Yang, J. Wang, and M. S. Zhan, "Coherent Population Trapping-Ramsey Interference in Cold Atoms," Chinese Physics Letters 27, 113201 (2010).

[19] D. M. Farkas, K. M. Hudek, E. A. Salim, S. R. Segal, M. B. Squires, and D. Z. Anderson, "A compact, transportable, Microchip-Based system for high repetition rate production of Bose-Einstein condensates," *0912.0553*, Dec. 2009.

[20] D. M. Farkas, A. Zozulya, and D. Z. Anderson, "A compact Microchip-Based atomic clock based on ultracold trapped Rb atoms," *0912.4231*, Dec. 2009.

Recent Atomic Fountain Clock Comparisons at LNE-SYRTE

J. Guéna, M. Abgrall, D. Rovera, P. Rosenbusch, B. Chupin, , M. Lours, G. Santarelli, P. Laurent, S. Bize and A. Clairon
LNE-SYRTE, Observatoire de Paris, UMR CNRS 8630
Paris, France
jocelyne.guena@obspm.fr

M.Tobar
University of Western Australia
Crawley, Australia

***Abstract*—We report on the current status of the LNE-SYRTE fountain clock ensemble composed of three fully operational fountain, one of which is a dual Cs/Rb clock. We focus on the comparisons between them and their impact as primary frequency standards contributing to the accuracy of TAI..**

I. MOTIVATIONS

The LNE-SYRTE clock ensemble comprises three atomic fountain clocks FO1, FOM and FO2 [1]. FO2 is a dual fountain using both rubidium and cesium [2]. FOM is a transportable cesium fountain. The clock ensemble also includes an ultra stable frequency reference based on a cryogenic sapphire oscillator (CSO) and disseminated to all three fountains [3].

The main application of this fountain clock ensemble is to contribute to the realization of the S.I. second by calibrating of national and international timescales. All three fountains are primary frequency standards regularly contributing to TAI.

Developments during the last years focused on improving long term stability and accuracy, and ability to operate unattended for extended periods of time. This is an important step towards the completion of the ground segment for the PHARAO/ACES space mission scheduled for 2014. The second important goal of the fountain clocks is to perform fundamental tests, such as the test of the stability of the fundamental constants [4]. Here the Cs/Rb dual alkali fountain clock provides an original tool. A third application is the utilization of the fountains as frequency references for absolute frequency measurements of optical clocks. At SYRTE this concerns the two strontium lattice clocks now fully operational, and in near future the Hg clock under development. In addition, other optical clocks in foreign laboratories have been measured against the transportable fountain FOM. This is an important step in preparation for a redefinition of the SI second.

Here we give an overview of the current status of the clock ensemble, and present long term fountain comparisons performed over the last three years.

II. FOUNTAIN PERFORMANCES AND COMPARISONS

The four fountain clocks operate near the quantum noise limit using the CSO as an interrogation oscillator [3]. The CSO ultrastable reference at 12 GHz is down converted to 9 GHz and 1 GHz. The FO2 fountain being located near the CSO uses directly the ultrastable signals to generate the Cs/Rb clock frequencies. For FO1 and FOM, as well as the optical clock laboratory, located in distant buildings within SYRTE, the 1 GHz and 9 GHz signals are disseminated through actively compensated fiber links, which are simplified versions of the link described in [5]. Since the CSO drifts above 1000 s, it is slowly phase-lock to one of the four SYRTE H-masers, the one that is linked to BIPM through GPS and TWTFT links. Thereby, each fountain measures the maser frequency.

Each fountain uses optical molasses beams for atom loading and cooling starting from a 2D-MOT cold atomic beam presource, except the FOM which loads from a vapour source. Each fountain, except the FOM, is operated in differential mode HD/LD (high/low density) with atom numbers $N_{HD} = 2N_{LD}$ using adiabatic passage [6] to allow for real-time evaluation of the collision shift with about 1 % accuracy. This limits the fountain stability to $\sigma(2\nu_{BD} - \nu_{HD})$. Typical instability at high density is 3.5×10^{-14} at 1s for FO1 and FO2, and 7×10^{-14} for FOM using the CSO interrogation signal.

Each fountain clock synthesis is operated in a pulsed mode using a dedicated interferometric switch [7] with Rabi pulses durations adjusted to eliminate potential microwave leaks during the Ramsey interrogation. We regularly check the absence of phase transients synchronous with the fountain cycle, or induced by the switching process, using a heterodyne phase transient phase analyzer [7].

Each fountain produces measurements, corrected for all systematics, at 1 measurement per cycle. The individual fountain measurements are averaged over each $1/100^{th}$ of day (864 s) to serve as a synchronous basis to make the frequency

This research is funded in part by CNRS and by LNE.

differences between them, taken two by two. This direct comparison eliminates the local oscillator noise.

An example of frequency instability between FOM and FO2 over 50 effective days of comparison is presented in Fig. 1. A stability of about 2×10^{-16} is obtained at $\tau = 4.5$ days.

The typical accuracy of the fountains is 3.8×10^{-16} for FO1 and FO2-Cs, 4.5×10^{-16} for FO2-Rb, and 7×10^{-16} for FOM, wherein the leading uncertainty of 3×10^{-16} comes from first Doppler effects associated with distributed cavity phase shifts. Recent quantitative evaluation of this effect with FO2 reduces this error to 8.4×10^{-16} for FO2-Cs [8].

Fountain comparisons are now performed regularly. The frequency differences of the three Cs fountain clock pairs over years 2008-2010 is presented on a monthly basis in Fig.2. The data corresponds to a total number of days of comparison of about 800 days. The error bars are total uncertainties for the clock pair which are dominated by the systematic uncertainty U_b. The weighted average differences D and related U_b are given in Table 1. We see that the FOM-FO1 and FOM-FO2 differences are well within the U_b, while the FO1-FO2 is slightly larger than one U_b. Although uncertainties are mainly of systematic origin, we test the statistical significance of these mean differences. To this aim, we evaluate the χ^2 of the weighted differences with respect either to the mean D or to zero. The goodness-of-fit Q and Q_o respectively, as defined in numerical recipes [9], are given in Table I. The Q's very close to 1 indicate that fluctuations about the means D's are much smaller than the total uncertainties. Q_o which tests deviations (and/or fluctuations) from a null difference shows that FOM-FO2 and FOM-FO1 agree within total uncertainties, while the low Q_o for FO1-FO2 would indicate a slightly underestimated systematic effect, of a few 10^{-16}. Similar analysis for a comparison of worldwide fountain primary frequency standards exploiting their calibrations of TAI is reported in [10].

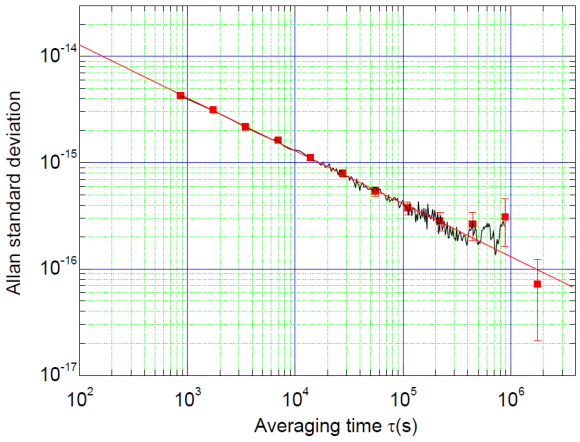


Figure 1. Example of Allan standard deviation for the fractional frequency difference between FOM and FO2Cs over 50 days. The continuous black line is the Allan deviation for all averaging times τ.

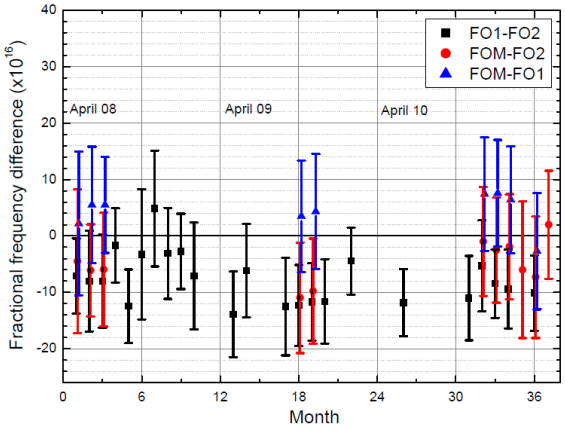


Figure 2. Fractional frequency differences between the three SYRTE Cs fountain pairs over the last three years. Error bars are total uncertainties dominated by the systematic uncertainties.

TABLE I. χ^2 TEST OF THE CS FOUNTAIN DIFFERENCES PLOTTED IN FIG.2

	D (x 10^{-16})	U_b (x10^{-16})	D/U_b [a]	Q	Q_0
FOM-FO2	-4.85	9.1	0.53	0.999	0.935
FOM-FO1	4.65	9.2	0.50	0.999	0.95
FO1-FO2	-8.1	6.2	1.31	0.999	0.037

a. D/Ub represents the difference in unit of standard deviation.

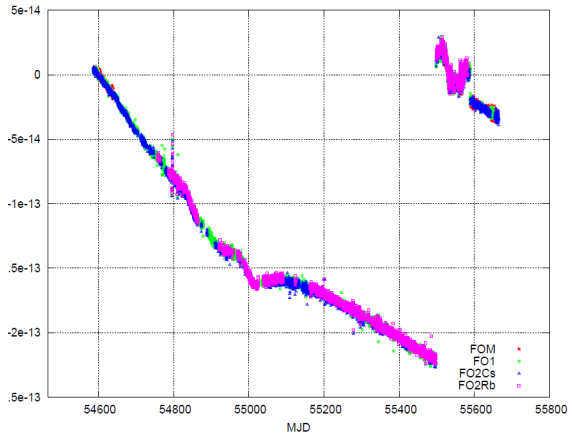


Figure 3. SYRTE fountain measurements of one H-maser at SYRTE (H 140-0890 from Kvarz, then 140-0816 from Symmetricom and again 140-0890). Changes in reference maser occurred due to failure.

III. SYRTE CS FOUNTAIN CONTRIBUTIONS TO TAI

Since 2007 the four fountain clocks, including FO2-Rb, are measuring the frequency of one of the SYRTE H-masers almost continuously, as shown in Fig. 3. These data were exploited to participate to the accuracy of TAI. Since 2007, about 75 formal reports have been transmitted from SYRTE to BIPM over a total of 161 reports from all standards worldwide, *i.e.* the contribution is by almost 50%. Each report

corresponds to the measurement of the H-maser by one fountain over 20-30 days, reaching a statistical uncertainty of about 1.2×10^{-16} and uncertainty of 1.5×10^{-16} due to dead times.

IV. SUMMARY AND SOME OUTLOOKS

In summary, we have presented long term comparisons between the SYRTE fountain clocks down to the $1x10^{-16}$ level. The fountains differences are consistent with the total stated uncertainties for each clock pair, except for the FO1-FO2 pair for which a systematic effect of a few 10^{-16} may be slightly underestimated. This is under current investigation. Recent improvements in the fountain operation and for long term comparison are provided by monitoring softwares capable of assessing the status of the overall system and automatically preprocessing clock data. These recent developments also result in an enhanced capability to contribute to International Atomic Time (TAI). Operational compensated fiber links within SYRTE allows for reliable comparisons between the fountain clocks and the optical lattice Sr and Hg clocks.

REFERENCES

[1] S. Bize et al., "Advances in atomic fountains", C.R. Physique, vol. 5, p.829, 2004. S. Bize et al., "Cold atom clocks and applications", J. Phys. B: At. Mol. Opt. Phys., vol. 38, S449, 2005.

[2] J. Guéna et al., "Demonatration of a dual alkaliRb/Cs fountain clock", IEEE Trans. UFFC, vol. 57, 2010, pp. 647–653.

[3] D. Chambon, et Al., "Design and realization of a flywheel oscillator for advanced time and frequency metrology," Rev. Sci. Instr., vol. 76, Sept. 2005, art. no. 094704,

[4] H. Marion et al., "Search for Variations of Fundamental Constants using Atomic Fountain Clocks", Phys. Rev. Lett., vol. 90, 2003, 150801.

[5] F. Narbonneau et al., "High resolution frequency standard dissemination via optical fiber metropolitan network", Rev. Sci. .Instr., vol. 77, 2006, 064701.

[6] F. Pereira Dos Santos, H. Marion, S. Bize, Y. Sortais, and A. Clairon, "Controlling the Cold Collision Shift in High Precision Atomic Interferometry", Phys. Rev. Lett., vol. 89 (23), 2002, 233004.

[7] G. Santarelli et al., "Switching atomic fountain clock microwave interrogation signal and high resolution phase measurements", IEEE Trans. Ultrason. Ferroelect. Freq. Contr., vol. 56 (7), 2009, 1319–1326.

[8] J. Guéna, R. Li, K. Gibble, S. Bize and A. Clairon, "Evaluation of Doppler shifts to improve the accuracy of primary atomic fountain clocks", Phys. Rev. Lett , vol. 106, 2011, 130801.

[9] Numerical Recipes in C, Second Edition (1992), p. 663. Q is the probability that a value of chi-square as *poor* as the value χ^2 found should occur by chance. According to this reference, "If Q is larger than, say, 0.1, then the goodness-of-fit is believable".

[10] T. J. Parker, "Long term comparison of caeium fountain primary frequency standards" , Metrologia, vol. 47, 2010, pp. 1–10.

Disciplined Rubidium Oscillator For Harsh Environments

T. McClelland, I. Shtaerman, E. Zarjetski, R. Baransky, M. Khurgin
Frequency Electronics, Inc.
Uniondale, NY/USA
tomm@freqelec.com

***Abstract*— A disciplined rubidium (Rb) oscillator has been developed for use in harsh environments. This oscillator operates from -40°C to +82°C, in the presence of random vibration of up to 19 g RMS; while continuously maintaining lock of both the Rb control loop and the disciplining loop. Performance under vibration is enhanced by incorporation of a quartz oscillator with active MEMS accelerometer based vibration compensation. The device warms up in less than 15 minutes, and can be disciplined to a 1 pps input, such as the signal from a GPS receiver. One of the unique features of this device is the capability to align the internally generated RF and 1 pps signals to the input 1 pps signal with a resolution of < 1 nsec. This is accomplished using a novel PLL approach for locking the internal RF signals to the disciplining 1 pps input. A number of design refinements to the Rb physics package have also been implemented which significantly improve the performance during "holdover" conditions when no disciplining signal is present. A major design goal of this device is to reduce the accumulated time error during a 24 hour holdover period to < 245 nanoseconds.**

The design of this device is described, and performance data is presented to demonstrate actual performance of initial production units compared to the design objectives.

I. INTRODUCTION

A Rb oscillator designed for use in harsh environments has been developed. The device is also capable of being disciplined by an external one pulse per second (1 PPS) signal, such as is typically available from a GPS receiver. This Rb oscillator utilizes a conventional, lamp excited vapor cell physics package, but with some key updates to the electronics, which make possible significant improvements to the performance generally obtained from this class of oscillator [1], [2], [3].

Photographs of this device are shown in Figure 1. , and Figure 2. , and an outline drawing, showing dimensions is provided in Figure 3. The unit is 4 x 6 x 1.5 inches, and weighs 29 ounces (820 grams).

Figure 1. Photograph of disciplined rubidium oscillator

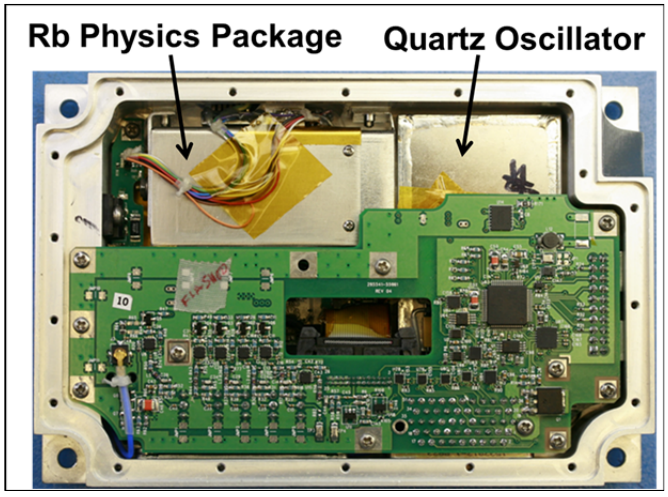


Figure 2. View of Rb oscillator with cover removed

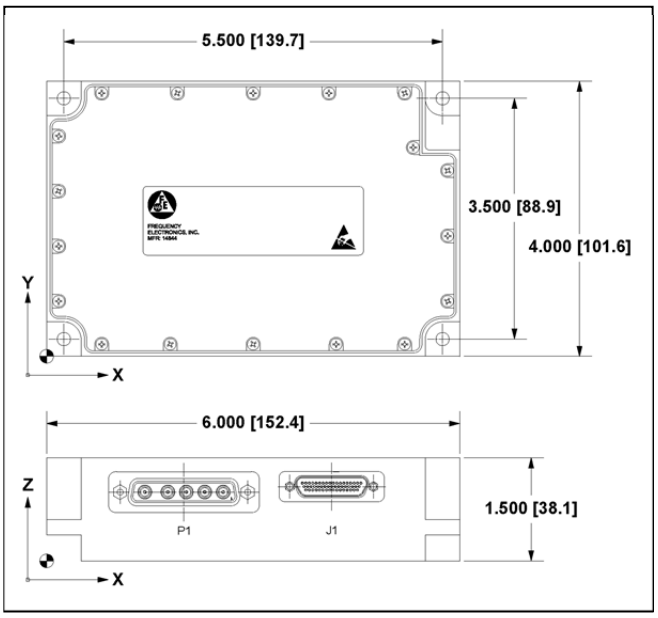


Figure 3. Outline drawing of Rb oscillator. Dimensions are in inches [millimeters].

II. DESIGN

A functional block diagram is provided in Figure 4. The device accepts a 1 pulse per second (PPS) input as a disciplining input, and provides both 10 MHz and 1 PPS outputs.

The 10 MHz outputs are derived from a quartz oscillator which incorporates active acceleration compensation in order to improve the performance during vibration. This active compensation is achieved by using the signals from MEMS accelerometers attached to the quartz resonator to correct for the acceleration induced frequency shifts in the resonator. A block diagram showing this is provided in Figure 5.

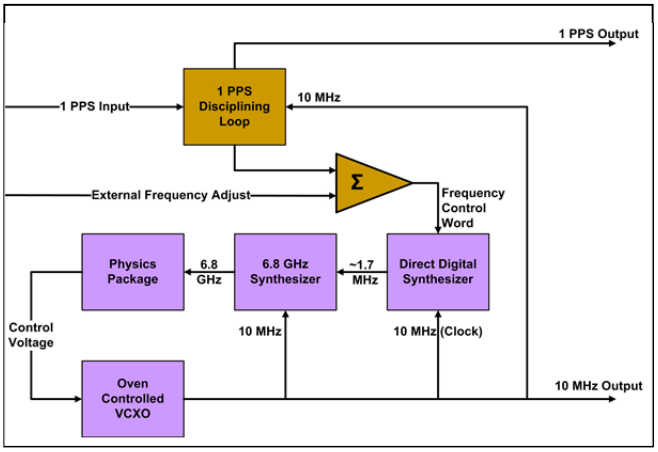


Figure 4. Block diagram of Rb oscillator

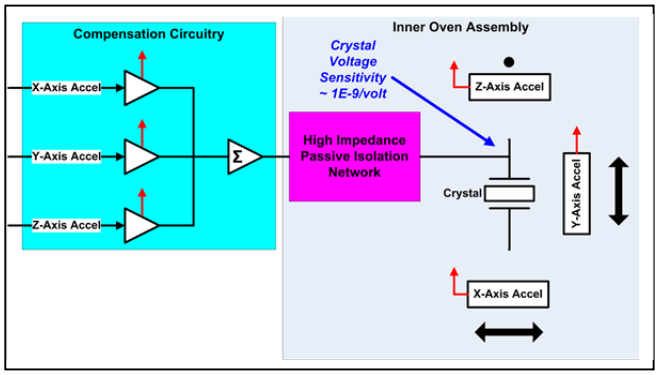


Figure 5. Block diagram of acceleration compensation circuitry

The rubidium control loop utilizes a direct synthesis approach to generate a signal at the Rb hyperfine resonance frequency of ~6.8 GHz. This signal is synthesized external to the Rb physics package, eliminating the need to incorporate any microwave circuitry inside the physics package. Previous designs of this type incorporate a step recovery diode comb generator inside the physics package, which generally involves the placement of components with ferromagnetic (kovar) packages near the Rb resonance cell. This magnetic material, when present, causes frequency variations which limit the achievable frequency stability.

The microwave excitation of the physics package modulates the transmitted light optical signal, which is used to steer the VCXO.

The 1 PPS disciplining loop adjusts the Rb frequency based on the 1 PPS input. This phase locked loop (PLL) utilizes a proprietary time interval counter which employs a direct digital synthesizer (DDS) clocked by the 10 MHz internal reference. The PLL measures the phase of the DDS output at the time of arrival of the input 1 PPS signal. The phase of the DDS output can be measured/adjusted with 12 bit accuracy, which allows the phase of the *output* 1 PPS signal to be aligned with the *input* 1 PPS signal with an accuracy of < 0.8 nanoseconds.

This PLL adjusts the output frequency of the Rb oscillator by adjusting the frequency of another DDS which is part of the microwave synthesis chain. The frequency can be adjusted in this way with a resolution of $<1 \times 10^{-14}$.

With a 1 PPS input, the disciplining loop adjusts the Rb oscillator frequency by steering the DDS frequency. With no 1 PPS input, manual frequency adjustment of the Rb oscillator is similarly accomplished; by external command which controls the frequency of the DDS.

III. PERFORMANCE

Considerable efforts (e.g. chip scale atomic clock programs) have recently been directed to driving down the size, weight, and power of Rb standards, while accepting performance compromises[4], [5]. This effort has a different goal. Rather than trying to achieve the smallest possible size, weight, and power, the goal is to achieve the best possible performance; while not significantly increasing the size, weight, and power compared to typical commercially

available Rb oscillators. Certainly, progress has been made toward this end.

A. Short Term Frequency Stability

The short term frequency stability of this Rb oscillator over a 5 day period of time is shown in Figure 6. , and Figure 7. This data is for an un-disciplined unit, and thus represents a measure of the capability of the free running Rb system. The fact that the Allan deviation exhibits $1/\sqrt{\tau}$ behavior out to 10,000 seconds is an improvement over what has generally been achieved in the past with this class of Rb oscillator [1], [2], [3].

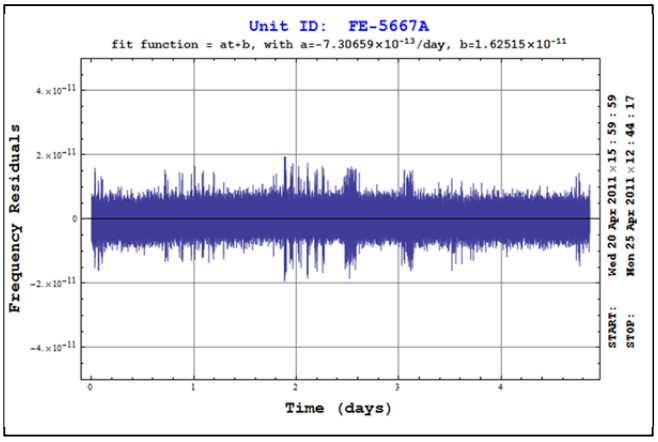


Figure 6. Frequency stability of ***un-disciplined*** Rb oscillator over a 5 day time period, at constant temperature (25°C). A linear drift of -7.3 x 10^{-13}/day has been removed from the data.

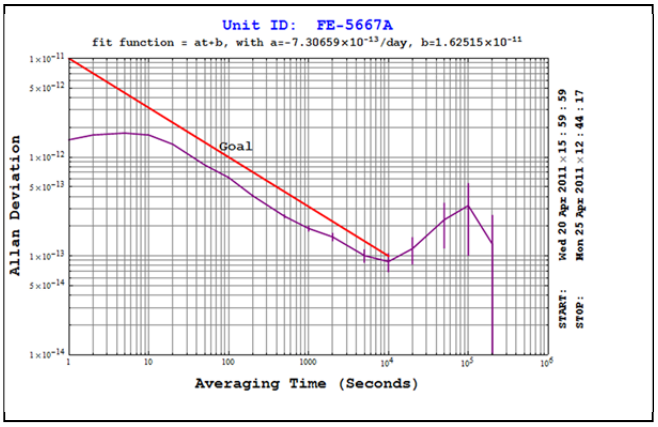


Figure 7. Allan deviation of ***un-disciplined*** Rb oscillator, from frequency data plotted in Figure 6. Performance at $\tau = 1$ second reflects the quartz oscillator noise, while the $\tau = 10$ to 10,000 second performance reflects the Rb loop white frequency noise.

B. Frequency vs. Temperature

Frequency vs. temperature for this device, again while un-disciplined, is provided in Figure 8. From -25°C to +75°C the frequency variation is <3 x 10^{-11}, performance which is an order of magnitude better than what has typically been achieved in the past [1], [2], [3].

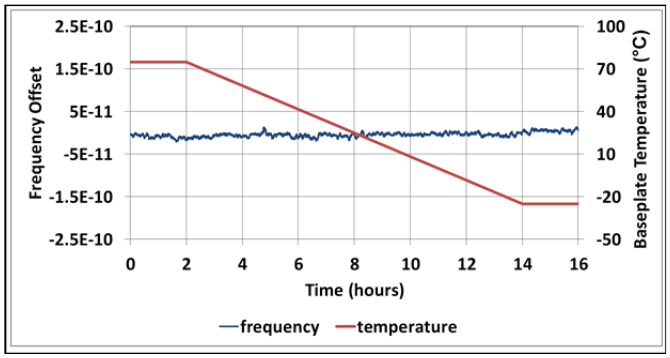


Figure 8. Frequency vs temperature of Rb oscillator from +75°C to -25°C. Test performed with ***no*** disciplining input

C. Holdover

Disciplined Rb oscillators are often used as primary frequency and time references in communication networks. One of the major concerns in this type of application is the behavior of the oscillator during brief periods of time ("holdover") during which the disciplining source is not available; for instance due to local jamming of GPS signals, urban canyon loss of GPS signals, or antenna failure. Depending on the application, holdover times for Rb oscillators are specified anywhere from 15 minutes to 100 days, but most often for a period of 24 hours.

Accumulated time error for this Rb oscillator during a 24 hour holdover period is shown in Figure 9. The test results shown in the figure demonstrate performance under simulated worst case temperature conditions given a maximum diurnal temperature variation of 17°C. In this case, the time error accumulated in 24 hours is under 25 nanoseconds. The unit is designed to be able to maintain <245 nanoseconds time error during any 24 hour holdover period. This performance is more than an order of magnitude better than what has generally been achieved in the past with compact Rb oscillators.

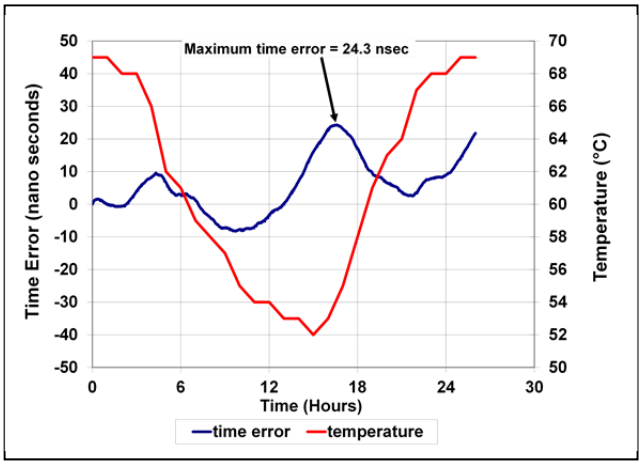


Figure 9. "Holdover" test to determine accumulation of time error while Rb oscillator is ***un-disciplined***. Time error accumulated over 24 hours is shown. Unit was subjected to a diurnal temperature variation of ~17°C, simulating a worst case condition, during this test.

D. Vibration

This Rb oscillator is designed not only to survive, but to function while undergoing severe vibration. As described above, the quartz oscillator incorporates active vibration compensation in order to improve performance under vibration conditions. Performance under extreme conditions (vibration = 19 g_{RMS}) is shown in Figure 10. The figure shows performance under vibration both with and without the active compensation circuitry enabled. For the results shown in the figure, the unit was subjected to vibration shown in TABLE I. From the figure, it is clear that active compensation improves the performance by more than 20 dB for vibration between 10 and 100 Hz.

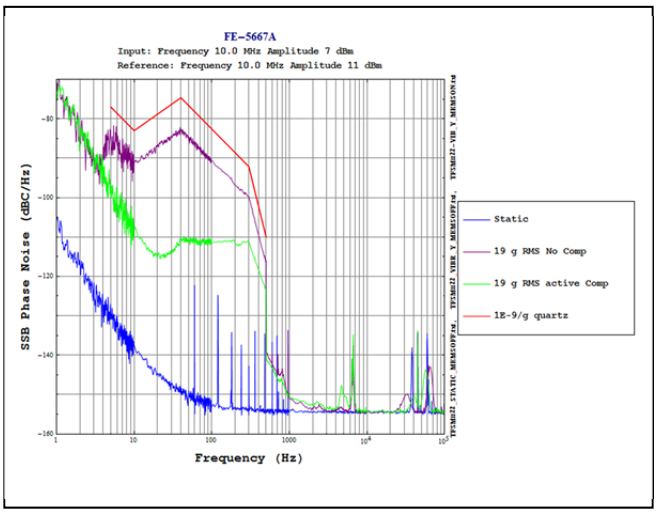


Figure 10. Phase noise of Rb oscillator. Both static phase noise and phase noise under vibration along x-axis are shown. The vibration profile used to collect this data is shown in TABLE I. These tests were performed on an ***un-disciplined*** Rb oscillator

TABLE I. RANDOM VIBRATION PROFILE

Frequency (Hz)	PSD (g^2/Hz)
5	0.01
10	0.01
40	1.1
300	1.1
500	0.05
total	**19 g RMS**

E. Warm-up

Fast warm-up is critical for many applications in which relatively short periods of operation may occur. This is the situation for units deployed in wheeled vehicles and aircraft. Performance immediately after turn-on is a challenge for Rb oscillators due to the fact that the major physics package elements must be ovenized. Disciplining to an external 1 PPS signal compounds the difficulty, especially when the disciplining source is a typical GPS signal with 20 to 120 nanoseconds of jitter.

Warm-up of this Rb oscillator is shown in Figure 11. During this test, the unit is continuously disciplined by a 1 PPS input signal. From the figure it is seen that 15 minutes after turn-on, the phase of the Rb oscillator generated signals are aligned with the phase of the input 1 PPS to within 10 nanoseconds. The frequency is accurate to within a few parts in 10^{-11} at the 15 minute point.

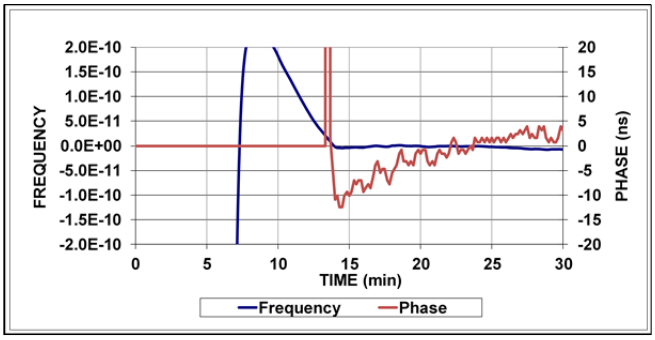


Figure 11. Warm-up of Rb oscillator at -40°C with 1 PPS input. Test performed ***with*** disciplining input signal.

F. Magnetic Field Sensitivity

Performance under conditions of varying magnetic field is often overlooked in Rb oscillators of this type. However, this is actually quite important in many of the dynamic environment applications such as moving vehicles, and aircraft. This Rb oscillator has exceptionally good performance in changing magnetic field conditions, as demonstrated in Figure 12. This figure shows that for a magnetic field variation of ±4 gauss, the frequency variation is less than 1 x 10^{-11}. This is at least 5 times better than other similar units.

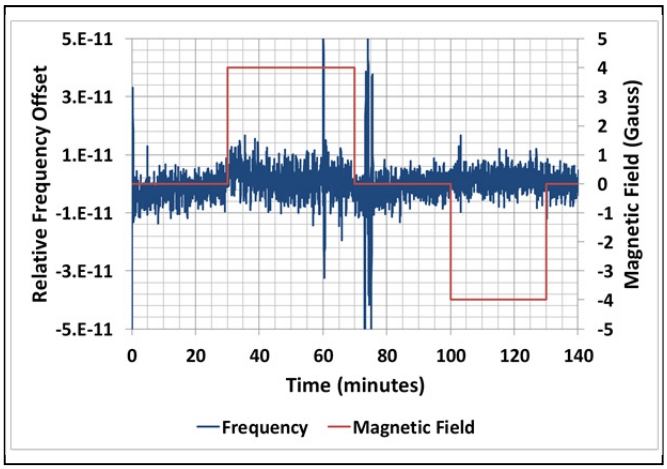


Figure 12. Magnetic field sensitivity of Rb oscillator, ***un-disciplined***. Frequency change is shown on the left axis, corresponding magnetic field on the right axis.

IV. CONCLUSION

A disciplined rubidium oscillator for use in harsh environments has been developed. This oscillator achieves much better performance than comparable devices currently available. The emphasis in designing this unit has been to

improve performance; not, as is the case for chip scale atomic clock efforts, to reduce size, weight and power even if performance compromises are necessary. Short term frequency stability, warm-up, holdover time error capability, performance over temperature, vibration operation, and magnetic field sensitivity five to ten times better than other Rb oscillators of this type have been demonstrated.

REFERENCES

[1] J. Ho, I. Pascaru, C. Stone, and T. McClelland, "New Rb frequency standard designs for telecommunications applications," Proceedings, IEEE International Frequency Control Symposium, 1998, p. 80.

[2] T. McClelland, J. Ho, I. Pascaru, and C. Stone, , "Rb frequency standard with expanded operating temperature range (-55 to +95°C)," Proceedings, IEEE International Frequency Control Symposium, 1999, p. 137.

[3] W. Weidemann, "Subminiature Rb oscillator model FRS," 40th Annual Frequency Control Symposium, 1986, p. 470.

[4] R. Lutwak, "The chip scale atomic clock – recent developments," Proceedings, IEEE International Frequency Control Symposium, 2009, p. 573.

[5] M. Zhu, and J. DeNatale, "Application of reduced light shift optical pumping method to chip scale atomic clock," Proceedings, IEEE International Frequency Control Symposium, 2009, p. 1183.

Stability of Relaxation Processes in a Quantum System as a Main Factor for Achieving the High Long-Term Frequency Stability of a Standard

Evgeny Pestov

[1] The Fed. State Scient./Prod. Ent. "Geologorazvedka", 19/1 Knipovich str., St. Petersburg 192019
[2] The Russian Institute of Radionavigation and Time, RIRT, 2 Rastrelli sq., St.-Petersburg 191124, Russia
E-mail: infra-balt@peterlink.ru

Abstract **– The processes of synchronous Zeeman's relaxation are investigated with simultaneous observing population difference, M_z (τ_1) – and coherence, M_x (τ_2), - components within a spin of system of ^{87}Rb atoms in cells with antirelaxation coating, when changing cell *volume*, operating *temperature* and pumping *light intensity* with D_1 – line.**
The existence of total relaxation 'degeneracy' is experimentally determined which is characterized through ($\tau_1 \approx \tau_2$) – condition being necessary [1] for achieving a high long-term frequency stability of standards (in particular, quantum magnetometers).
The procedure for investigating relaxation processes developed by the author could be useful for the realization of atomic frequency standards with a high long-term frequency stability.

I. INTRODUCTION

Of all atomic frequency standard's parameters the major one is the long-term frequency stability. All efforts of frequency standards' researchers and creators are ultimately directed to achieve a high long-term frequency stability.
The central unit providing a long-term frequency stability is the atomic discriminator. Whatever the quantum system is used, two main factors affect obtaining the maximum value of atomic discriminator's long-term frequency stability:
1. Technique for exciting a resonance '0-0' – transition which implies the realization of the regime of atomic system's forced oscillations by means of, in particular, the external alternating magnetic field, $H_1 \cdot \cos\omega t$.
2. Dynamics of relaxation processes within a quantum system.

II. Impact of the Regime of Quantum System's forced Oscillations on Long-term Stability.

1.1. It is found that the resonance exciting field, $H_1 \cdot \cos\omega t$, shifts a resonance frequency of the quantum system, $\nu_{0\text{-}0}$, by a small value, $\delta\omega^{(2)}_{b\text{-}s} \sim \gamma H_1^2 / 4H_0$, which is a constant correction of the second order for the field amplitude H_1, and is known as the Bloch-Siegert shift. At frequencies of hyperfine '0-0' – transitions, this correction is small ($\sim 10^{-15}$), not accounted for in practice and does not affected a value of long-term stability.
1.2. One should consider as a success of the recent decade such a point as the detection of a drift varying-in-time correction of the first order of $\delta\omega^{(1)}_{ch\text{-}p} = (\gamma \cdot H_1 \cdot z/a) \cdot \cos\vartheta$ – kind made E.Pestov and A.Chirkov [1]. This correction caused by a forced oscillations' regime itself, is always presents in frequency standards. It cannot be removed in practice and is a source of the long-term drift of resonance frequency's center, $\nu_{0\text{-}0}$, with a long period – from several units of day to tens or more. The variable correction, $\delta\omega^{(1)}_{ch\text{-}p}(t)$, affects significantly the standard's long-term frequency stability.

III. Impact of Relaxation Processes' Dynamics within a Quantum System on Long-Term Stability of a Standard. Results of the Investigation

For providing a stability of various quantum systems' states, it is necessary in practice to solve a number of difficult technical tasks. Let us consider a problem related to the need in supporting the stability of relaxation processes within absorption cells of rubidium frequency standards (RFS) and point out the ways to improve a long-term stability of frequency standards of this type.

As is shown in the same work [1], achieving the ultimate long-term stability of any quantum system when exciting a resonance by the harmonic field, $H_1 \cdot \cos\omega t$, is only possible when realizing within this system a regime of stationary oscillations (RSO) in which a drift variable correction of the resonance frequency, $\delta\omega^{(1)}_{ch\text{-}p}(t)$, tends to a negligible value. For providing RSO existence and steadiness (under conditions of frequency detuning, $\Delta\omega = |\omega - \omega_0| = 0$), resonance signal's amplitude, a_s , population difference, z_s , and phase difference, ϑ_s , between phases of resonance field and atom should be supported with a super high accuracy during the operation of a frequency standard. It is difficult to solve this task in practice, in particular, with respect to ϑ_s [1]. In addition, RSO steadiness is significantly affected by relaxation processes disturbing a frequency of the clock atomic transition.

From [1] it follows that, for improving RSO steadiness (and, consequently, improving a high standard's long-term stability), it is necessary to realize within a quantum system the equal intensity of relaxation processes as regards population component (M_z) and coherence component (M_x) of operating states, i.e. to provide a condition of equality for efficient relaxation times, $\tau_1 \approx \tau_2$, respectively. The distinction in intensities of this processes (at $\tau_2 < \tau_1$) which takes place in practice is an additional source of lowering the long-term stability of a standard (AFS).

As will be shown later, for realizing the case $\tau_1 \approx \tau_2$, we recommend to use gas cells with anti-relaxation wall (ARW) coating. For example, Cs cells with ARW-coating were for the first time used as early as 1971 in the USSR within Cs magnetometer at "Cosmos-355" geophysical satellite [2]. The Cs quantum sensor operated under conditions of open space and demonstrated super results.

The manufacture of cells and wide-scale investigations of relaxation processes in ^{133}Cs, ^{87}Rb, ^{39,41}K cells with ARW -coating were conducted in IZMIRAN (USSR). [The author of this work participated in these investigations].

In recent years the interest is growing as regards the use of cells with ARW-coating in frequency standards due to more efficient optical pumping and lower intensity of relaxation processes compared to those in buffer gases. It manifests itself as increase in SNR value and decrease in resonance line width, $\Delta\nu$. For example, when using Rb - cylindrical cells of our manufacture (coating is tetracontane) with dimensions of Ø22x22, the width of '0-0'-resonance line decreased by 4-5 times compared to the lines within the similar cells with buffer gases. The absorption cells with ARW-coating have a larger value of the figure of merit, $Q \sim [S(\tau_1)/N]/\Delta f(\tau_2)$, and this opens a prospect for creating the competitive ^{87}Rb frequency standards.

On the strength of this data, such a problem becomes actual as conducting investigations of relaxation processes' dynamics in the ^{87}Rb cells with ARW-coating for detecting a condition of relaxation times equality, $\tau_1 \approx \tau_2$, when changing cell *volume*, operating *temperature*, optical pumping *intensity*.
In this work, the results are presented as regards experimental investigations of Zeeman's relaxation times of population difference, τ_1, and coherence, τ_2, [3], in $F_g = 2$ state of ^{87}Rb atomic spin-system within cells with ARW-coating as functions of pumping lamp's light intensity, $I_1(D_1)$, and temperature, t^0C.

τ_1 (I_1, t^0)- and τ_2 (I_1, t^0) - dependences within cylindrical ^{87}Rb cells of various dimensions are obtained using a technique developed by the author for measuring τ_1 and τ_2 *synchronous* relaxation in spin systems [4] which allows to simultaneously observe both relaxation processes – population difference component, M_z, and coherence component, M_x. Analysis of current works on relaxation investigation in spin systems shows that these investigations are only restricted by a single relaxation component – population difference, M_z [5,6].

Fig. 1 shows dependences of relaxation times τ_1, τ_2 on a change in pumping light's intensity with D_1-line, $k \cdot I_1$, for ^{87}Rb cells of various dimensions:
plots 1 (τ_1), 2 (τ_2) – sphere, Ø38 (inner 36.5); cylinders, 3 (τ_1), 4 (τ_2) – Ø22x22 (inner Ø19x19) mm;
5 (τ_1), 6 (τ_2) – Ø15x15 (inner Ø12x12) mm. $I_1 = 1 \sim 25$ $\mu W/cm^2$; here k = 0.1; 0.15; 0.32; 0.5; 0.7 – transmission factors of neutral light filters. (1-6) curves characterize the relaxation processes' dynamics of M_z and M_x components depending on a change in cells' volume at the fixed cell temperature, $T \approx 23^0$C.

Fig. 2 demonstrates similar dependences of relaxation times, $\tau_1(k \cdot I_1)$, $\tau_2(k \cdot I_1)$, for cylindrical ^{87}Rb cell of Ø15x15 – dimensions. (1-4) dependences characterize the relaxation processes' dynamics of M_z and M_x components depending on a change in the cell temperature.
The high-melting paraffin $C_{40}H_{82}$ of high-degree purity was selected as coating.

Each event of atom's collision with a wall generates on this wall a fluctuating local field along all coordinates: $\overline{H_x^2}$, $\overline{H_y^2}$, $\overline{H_z^2}$ - are the average energies of the components of the variable magnetic field *induced* in the time of the atom's interaction with the cell wall-coating.
Microscopic relaxation theory (Redfield, Kubo, Tomita) gives expressions for relaxation times, T_1 and T_2 , [7]

$$1/T_1 = \gamma^2 [\, \overline{H_x^2} + \overline{H_y^2} \,] \cdot \tau_c / (1 + \omega_0^2 \cdot \tau_c^2) \qquad (1)$$

$$1/T_2 = \gamma^2 [\, \overline{H_z^2} \cdot \tau_c + \overline{H_y^2} \cdot \tau_c / (1 + \omega_0^2 \cdot \tau_c^2 \,)] \qquad (2)$$

These local fields, $\overline{H_x^2}$,.., are the cause of decrease of the population difference and resonance broadening.
It follows from the expression (1) that T_1 affects only the population, Δn, because of the transverse fluctuating fields, $< H_x^2 >$ and $< H_y^2 >$, transfer the atom from one sublevel (m_F) to another one ($m_F \pm 1$) at the resonance frequency, ω_0 , and lower a signal value. The expression for T_2 (2) is of more complicated kind and characterizes a loss in phase coherence which is accompanied with a resonance broadening.
Let us consider some *specific features* of relaxation as regards populations (T_1) and phase (T_2) on coating.

1. When increasing in a cell temperature, the correlation time, τ_c , i.e. time of interaction with a wall, decreases. This means that T_1 and T_2 increase.

2. Rates of T_1 and T_2 growth due to a change in temperature are different. It follows from formulas (1) and (2) that the T_1 time will increase *faster* than T_2 with a decrease in the correlation time.

3. The addendum, $\overline{H_z^2} \cdot \tau_c$, for the phase relaxation (T_2) reflects the dipole-dipole interaction and causes a resonance broadening. It lowers also the T_2 increase velocity on coating when the temperature increases.

4. For wall relaxation times, T_1 and T_2, the inequality $T_1 \neq T_2$ is always satisfied, and $T_1 > T_2$.

Let us write the expressions for times τ_1, τ_2 taking into account spin-exchange and optical pumping.

$$1/\tau_1 = 1/T_1 + 1/\tau_{sp-ex}^{(1)} + 1/\tau_p \qquad (3)$$

$$1/\tau_2 = [1/T_2'(\overline{H_z^2}) + 1/2\,T_1 + 1/T_{S\text{-}N}] + 1/\tau_{sp-ex}^{(2)} + 1/\tau_p \qquad (4)$$

The addendums in the first brackets are responsible for the phase relaxation *on coating*, where $1/T_{S\text{-}N}$ is spin-spin relaxation due to the interaction of a working atom's spin with the coating's nuclear spin.
Let us rewrite the formula (3), (4) in a compact style.

$$1/\tau_1 = 1/T_1 + 1/\tau_{sp-ex}^{(1)} + 1/\tau_p \qquad (5)$$

$$1/\tau_2 = 1/T_2^* + 1/\tau_{sp-ex}^{(2)} + 1/\tau_p \qquad (6)$$

These formulas describes experimental dependences which will be presented below.
In the course of experiment, a pumping intensity, $k \cdot I_1$, is changed. When decreasing $k \cdot I_1 \rightarrow 0$, we obtain dark elaxation times, τ_1^* and τ_2^*, acting within a cell, which reflect wall and spin-exchange relaxations (7,8):

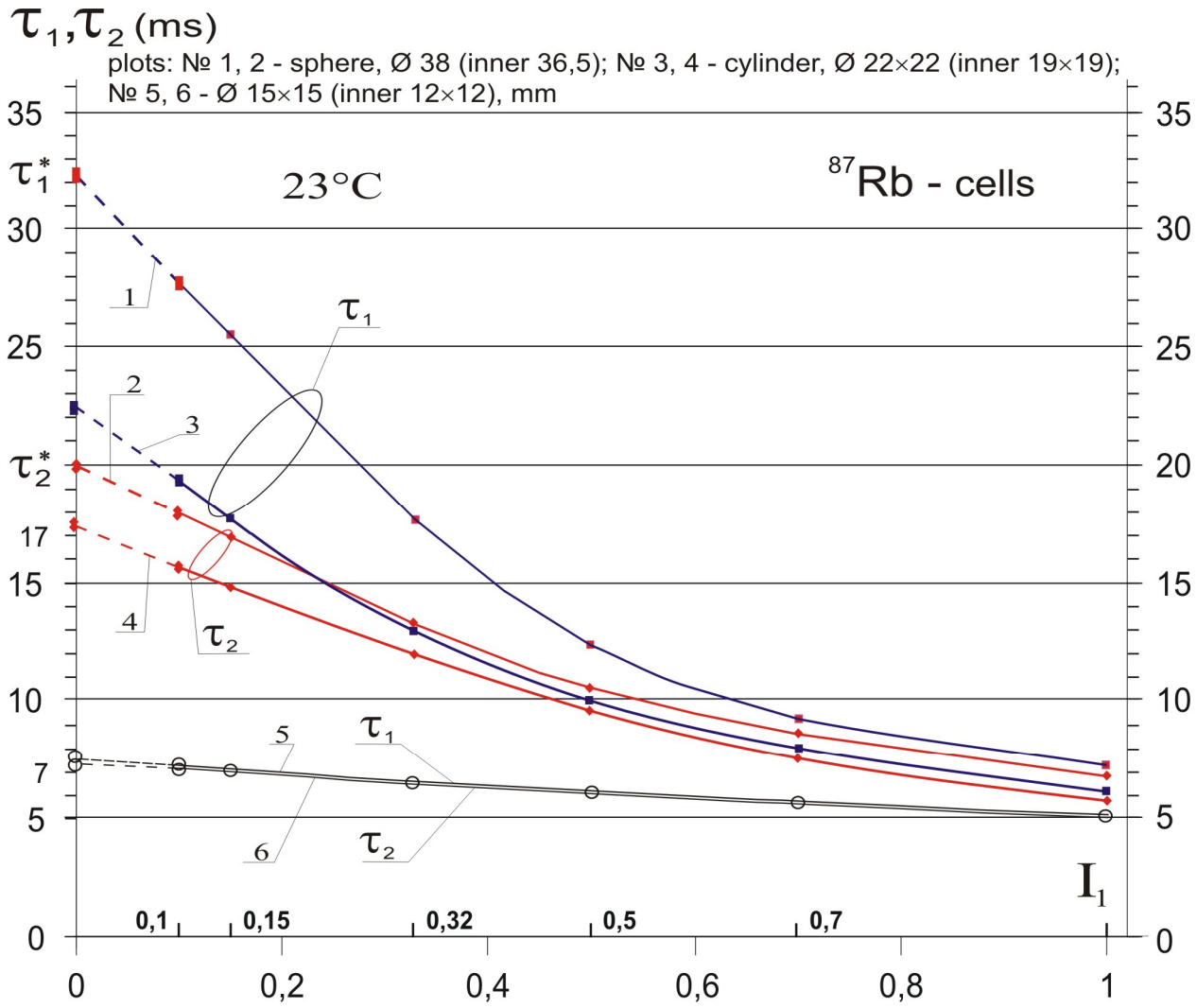


Fig. 1. Dependences of relaxation times τ_1, τ_2 on a change in pumping light's intensity with D_1 –line for ^{87}Rb cells of various dimensions

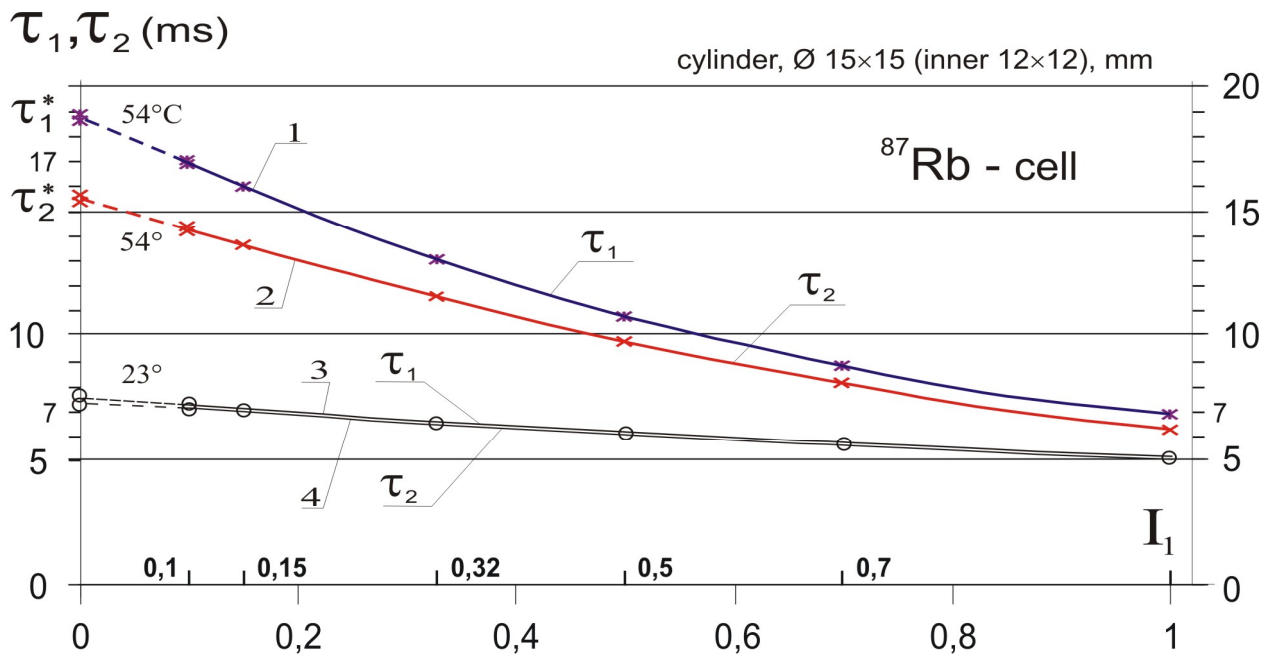


Fig. 2. Dependences of relaxation times τ_1, τ_2 on a change in pumping light's intensity with D_1 –line for cylindrical ^{87}Rb cell of Ø 15×15 (mm) in the various temperature

$$1/\tau_1^* = 1/T_1 + 1/\tau_{sp-ex}^{(1)}, \quad (7)$$

$$1/\tau_2^* = 1/T_2^* + 1/\tau_{sp-ex}^{(2)}, \quad (8)$$

where T_1, T_2^* are the relaxation time on the wall coating; τ_1^* and τ_2^* values are on the ordinate axis, Fig.1, 2.

Let us consider specific features of results obtained in this work.

1. Change in a cell volume (t^0 = const = 23^{0}C), Fig.1.

• When increasing a cell volume (cylinder, Ø15x15 → Ø22x22 → sphere Ø38, mm), a difference, $\Delta = \tau_1^* - \tau_2^*$, is increased between dark relaxation times, τ_1^* and τ_2^*, respectively:
$\Delta_{Ø15x15} \leq$ 0.5 ms; $\Delta_{Ø22x22} \sim$ 5 ms; $\Delta_{Ø38} \sim$ 12 ms. In so doing, τ_1^* time increases significantly faster than τ_2^*. These circumstances show that, then increasing a length of atomic free path, a frequency of atomic collisions with the wall and between themselves decreases. As a result, such a parameter decreases as the spectral power density of fluctuating local fields on the wall, $<H_x^2>$ and $<H_y^2>$, revealing in T_1 and τ_1^* increase.

• τ_2^* value is always less than τ_1^*, i.e. $\tau_2^* < \tau_1^*$. The reason is the presence of $1/T_2^*$ ($<H_z^2>$, $1/T_{S-N}$) addends in the formula for τ_2^* (4) which decrease τ_2^* and broaden the resonance.

• The interesting phenomenon is observed for a cylindrical cell with a lesser volume, Ø15x15 (curves no. 5, 6). For this cell, some *total* 'degeneracy' for relaxation times, ($\tau_1 \approx \tau_2$), takes place which is being characterized not only by the dark 'degeneracy', $\tau_1^* \approx \tau_2^*$, at $I_1 = 0$, but also by the 'degeneracy' within all range of pumping light's intensity change, $k \cdot I_1$. This state of the relaxation 'degeneracy' within a quantum system, according to [1], just provides a condition for achieving a high long-term frequency stability of a standard.

• The pumping light, while acting on a quantum system as the light relaxation, helps to generate a state being near to the relaxation 'degeneracy' ($\tau_1^* \approx \tau_2^*$). For example, for Ø22x22 cell (curves no.3, 4), this state occurs at $k \cdot I_1 \geq 0.5 \cdot I_1$ where $I_1 = 1$ corresponds to the emission power density ~ 25 μW/cm^2.

2. Change in a cell temperature (volume V = const), Fig.2

• Fig.2 gives an example of experimental dependences $\tau_1, \tau_2 = f(k \cdot I_1)$, for a cell of Ø15x15 (inner Ø12x12). It follows from this data that, for this cell, the dark relaxation 'degeneracy', ($\tau_1^* \approx \tau_2^*$), disappears at the temperature $t^0 \sim 54^0$C.

It is interesting in practice, however, to find τ_1, τ_2 (I_1) – dependences from the temperature change in the case of the total 'degeneracy' (similar with 5,6 – dependences, Fig.1) for determining a critical temperature, t^0_c, beyond which the dark 'degeneracy' is removed. It follows from these dependences, $\tau_1^*(t^0)$, $\tau_2^*(t^0)$ obtained and given in Fig.3, that a condition for the *total* relaxation 'degeneracy' for this cell is fulfilled up to the critical temperature, $t^0_c \sim 36^0$.

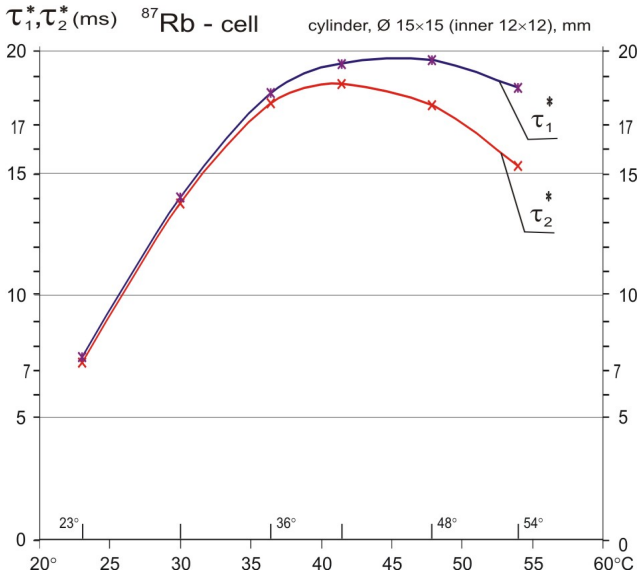


Fig.3. The example of the total 'degeneracy' τ_1^* and τ_2^*

Frequency Shifts of '0-0' Transition in RFS Caused by Action of Coating

This wall shift is negative and is of several hundreds Hertz. Preliminary investigations performed by us have shown that frequency shifts could be compensated through adding the buffer gas of small pressure to the cell.

IV. CONCLUSIONS

1. For a cylindrical cell with ARW-coating, the dimensions are found, ∅15×15 (∅12×12), at with a total relaxation 'degeneracy' is observed at temperature values up to critical, $t^0_c \sim 36^0$ C. This means, for example, that, when changing the parameters within a wide temperature range ($20^0 \div 36^0$C) and pumping light intensity, $k \cdot I_1$, within (0÷25) μW/cm^2, ($\tau_1 \approx \tau_2$) - condition is fulfilled which is necessary for achieving a high long-term frequency stability of standards (in particular, RFS and magnetometers).

2. The dynamics of relaxation processes within a spin system with hyperfine transitions is similar to processes of Zeeman's relaxation and differs only in values of relaxation time, τ_1 and τ_2. Therefore we believe that the procedure for relaxation investigation and results of this work would be useful, when performing similar investigations on hyperfine transitions in the course of developing atomic frequency standards.

3. The total relaxation 'degeneracy', ($\tau_1^* \approx \tau_2^*$), characterizes the isotropy of relaxation processes in the space.

4. The performed investigations enable one to determine conditions for achieving a high frequency stability

(short-term, medium-term, and long-term) and finding operating parameters for realization such stability values in practice.

REFERENCES

[1] E.N. Pestov and A.G. Chirkov,
"New Results of the Investigations of the NATURE of the Long-Term Instability in Quantum Frequency Standards and Magnetometers. NON-Removable Resonance Drift", *Proc*. 16th EFTF' 02, St. Petersburg, Russia, p. E063-066 (2002)

[2] Sh. Sh. Dolginov and A.N. Kozlov, "Magnetometers for Space Measurements", Rev. de Phys. Appl. **5**, № 1, 178-182 (1970)

[3] E.N. Pestov, Reports of INTAS-CNES Project № 06-1000024-9321 (2007-2008)

[4] E.N. Pestov, "Instantaneous and in-Phase Simultaneous Excitation of Coherent Signals from Several Quantum Systems", JETP, **102**, № 6, 1781-1787 (1992)

[5] M.V. Balabas, Zh. Tekh. Fiz., **80,** № 9, 91-96 (2010)

[6] D. Budker, L. Holberg, D.F. Kimball, et al., Phys. Rev. A **71**, 012903-09 (2005)

[7] C.P. Slichter. Principles of Magnetic Resonance, Berlin, Heidelberg, New York (1980)

Viscosity Sensor Based on *c*-axis Tilted AlN Thin Film Bulk Acoustic Wave Resonator

Lifeng Qin[1], Yizhong Wang[1], Jing-Feng Li[2], and Qing-Ming Wang[1*]
[1]Department of Mechanical Engineering and Materials Science, University of Pittsburgh, Pittsburgh, PA 15261
[2] Department of Materials Science and Engineering, Tsinghua University, Beijing, China
*qmwang@engr.pitt.edu

Abstract— **It has been shown that pure thickness shear mode can be excited in thin film bulk acoustic wave resonators (FBARs) using AlN thin films with special *c*-axis tilt angles (46.1° and 90°). The pure shear mode FBARs may be used as sensors in liquid medium with improved sensitivity than the commonly used thickness shear mode quart crystal resonator. In this paper thickness shear mode FBAR has been investigated for viscosity sensor application. The equation for predicting electric impedance of shear mode FBARs with a liquid layer was derived from the basic piezoelectric constitutive equations. The viscosity sensitivity of AlN FBARS was achieved by calculation of resonant frequency shift due to the liquid loading layer. In the calculation, it is assumed that the FBAR is formed by AlN film with 2μm thickness, 300μm by 300μm electrode area and with mechanical quality factor of 400. Three different liquids (water, acetone and olive oil) were used as the loading layer to evaluate the sensitivity of FBAR viscosity sensors. It was found that the viscosity sensitivities of AlN FBAR with 46.1^0 and 90^0*c*-axis tilted angle are very close, and do not change much with different liquid materials. The simulation results can be used for design and application of AlN FBARs.**

I. INTRODUCTION

In the past years, FBARs have attracted much attention for application in RF and microwave frequency control and signal processing Fields [1-3]. In addition to applications in electronic circuits as frequency control and signal processing devices, FBARs were also investigated for sensor applications due to its high sensitivity [4-7]. AlN and ZnO thin films are widely used for FBAR fabrication, and they are usually with normal-plane *c*-axis orientation, which makes FBARs operated in thickness longitudinal mode. The damping of thickness longitudinal resonator in liquid is severe, inducing poor sensor response. While the damping of shear mode resonator in liquid is much less significant, hence, for liquid application FBAR operated in shear mode is promising.

It has been shown that thickness shear mode resonance can be excited besides longitudinal mode when *c*-axis of AlN thin film is tilted off the thickness normal direction. For example, *c*-axis tilted AlN thin films were achieved by Bjurström and Wingqvist et al using reactive magnetron sputter deposition system; they examined the variation of electromechanical coupling and the quality factor of the AlN resonators with *c*-axis tilt angle, and investigated the performance of resonator for sensor application in liquid [8-10]. Martin et al. presented dual mode resonators based on *c*-axis tilted AlN films with high resonant frequency, around 8GHz for longitudinal mode and 4 GHz for shear mode [11, 12]. Chuang et al. demonstrated solidly mounted dual mode FBARs based on 0^0, 15^0 and 30^0 *c*-axis tilted AlN films, and the effect of mass loading on the resonators were experimentally conducted.[13].

Our previous study [14] also shows that the tilted polar *c*-axis orientation in AlN and ZnO thin films will induce normal plane and in-plane polarization components, which leads to the coexistence of the longitudinal mode and shear mode in the resonators. Moreover, it was found that pure thickness shear mode exists when *c*-axis is tilted at a special angle. In the case of AlN FBAR, when *c*-axis is tilted at angle 46.1° or 90°, only thickness shear mode exist, which, as mentioned before, can be a very promising FBAR structure for sensor applications in liquid medium. At this point, theoretical analysis of FBAR sensors based on polar *c*-axis tilted AlN thin film is needed to guide device design and fabrication. To our knowledge, there is no theoretical study on *c*-axis tilted AlN FBAR for liquid viscosity sensing application from the literature. In this study, we focus on shear mode FBARs based on AlN thin film with 46.1° and 90° *c*-axis tilt angles. The electric impedance of pure thickness shear mode AlN FBAR with a liquid loading was derived by the basic piezoelectric equations. Then the resonant frequency shift for different liquid loading was examined. Finally the viscosity sensitivity of shear mode AlN FBAR was calculated and discussed.

II. THOERY

A. Dual mode AlN FBARs.

Before we discuss AlN FBAR for viscosity sensor application, we briefly present the theoretical analysis of dual mode FBARs using *c*-axis tilted AlN, which is the basis for analytical study of shear mode AlN FBAR with liquid loading. Following a similar procedure [15], the impedance equation for *c*-axis tilted AlN FBARs can be derived [14].

Fig. 1(a) shows FBAR based on AlN thin film with normal *c*-axis, and (x_1, x_2, x_3) is the original material coordinate system; Fig. 1(b) shows the schematic of FBAR based on *c*-axis tilted AlN thin film, where the top and bottom electrode

are ignored for model simplification. A rectangular Cartesian coordinate system ($x_1^{'}$, $x_2^{'}$, $x_3^{'}$) is chosen with top electrode on $x_3^{'} = h$ and bottom electrode on $x_3^{'} = 0$. The *c*-axis of AlN film is tilted at an angle θ to the $x_3^{'}$ direction. Coordinate system ($x_1^{'}$, $x_2^{'}$, $x_3^{'}$) can be treated as the result of rotation of (x_1, x_2, x_3) about x_2 with an angle θ, as shown in Fig. 1(c).

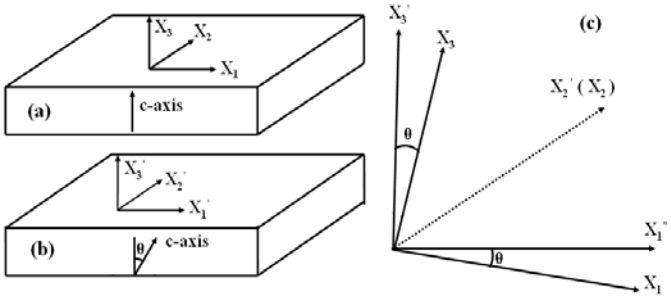


Fig.1 (a) FBAR based on AlN thin film with normal *c*-axis; (b) FBAR based on *c*-axis tilted AlN; (c) the relation of (x_1, x_2, x_3) and ($x_1^{'}$, $x_2^{'}$, $x_3^{'}$).

In ($x_1^{'}$, $x_2^{'}$, $x_3^{'}$), we have the following acoustic field equations and piezoelectric constitution equations[16]:

$$\frac{\partial T_1}{\partial x_1^{'}} + \frac{\partial T_6}{\partial x_2^{'}} + \frac{\partial T_5}{\partial x_3^{'}} = \rho \frac{\partial^2 u_1}{\partial t^2}$$
$$\frac{\partial T_6}{\partial x_1^{'}} + \frac{\partial T_2}{\partial x_2^{'}} + \frac{\partial T_4}{\partial x_3^{'}} = \rho \frac{\partial^2 u_2}{\partial t^2} \quad (1)$$
$$\frac{\partial T_5}{\partial x_1^{'}} + \frac{\partial T_4}{\partial x_2^{'}} + \frac{\partial T_3}{\partial x_3^{'}} = \rho \frac{\partial^2 u_3}{\partial t^2}$$

$$S_1 = \frac{\partial u_1}{\partial x_1^{'}}, S_2 = \frac{\partial u_2}{\partial x_2^{'}}, S_3 = \frac{\partial u_3}{\partial x_3^{'}}$$
$$S_4 = \frac{\partial u_3}{\partial x_2^{'}} + \frac{\partial u_2}{\partial x_3^{'}} \quad (2)$$
$$S_5 = \frac{\partial u_3}{\partial x_1^{'}} + \frac{\partial u_1}{\partial x_3^{'}}$$
$$S_6 = \frac{\partial u_2}{\partial x_1^{'}} + \frac{\partial u_1}{\partial x_2^{'}}$$

$$T_p = c_{pq}^{E'} S_q - e_{kp}^{'} E_k \quad (3)$$

$$D_i = e_{iq}^{'} S_q + \varepsilon_{ik}^{S'} E_k \quad (4)$$

where T_p, S_q, D_i and E_k are the components of stress, strain, electric displacement and electric field intensity, $c_{pq}^{E'}$ are the elastic stiffness constants under constant electric field intensity, $e_{kp}^{'}$ are the piezoelectric stress constants, $\varepsilon_{ik}^{S'}$ are the permittivity constants under constant strain, ρ is the density of AlN and u_i is the displacement in the direction of $x_i^{'}$ ($i, k = 1, 2, 3$ and $p, q = 1, 2, 3, 4, 5, 6$). For one dimension assumption, we have:

$$\frac{\partial T_i}{\partial x_1^{'}} = \frac{\partial T_i}{\partial x_2^{'}} = 0, E_1 = E_2 = 0, E_3 \neq 0, \frac{\partial D_3}{\partial x_3^{'}} = 0 \quad (5)$$

Under a sinusoidal excitation, the voltage V and current I across FBAR can be determined by

$$V = \int_0^h E_3 dx_3^{'}, \; I = j\omega A D_3 \quad (6)$$

where A is electrode area, ω is angular frequency. For traction forces at the boundary ($x_3^{'} = 0, h$) we have:

$$T_5(0) = T_3(0) = T_5(h) = T_3(h) = 0 \quad (7)$$

Also we found:

$$c_{34}^{E'} = c_{45}^{E'} = e_{34}^{'} = 0 \quad (8)$$

From equation (1)-(8), the impedance of dual mode FBAR can be solved:

$$Z = \frac{1}{j\omega C_0}(1 - (k^L)^2 \frac{\tan(\gamma^L / 2)}{\gamma^L / 2} - (k^S)^2 \frac{\tan(\gamma^S / 2)}{\gamma^S / 2}) \quad (9)$$

Here C_0 is the static capacitance; k_L^2 and k_S^2 are electromechanical coupling coefficients; γ^L and γ^S is phase delay for longitudinal (L) and shear mode (S), respectively.

Based on the impedance expression of FBAR (9), the resonator could be operated in a single mode when a tilt angle makes $k_L = 0$ (pure shear mode) or $k_S = 0$ (pure longitudinal mode). Our calculation results show when *c*-axis tilt angle is 46.1^0 or 90^0, AlN FBARs can be operated in pure thickness shear mode ($k_L = 0$).

B. Shear mode AlN FBAR with liquid loading

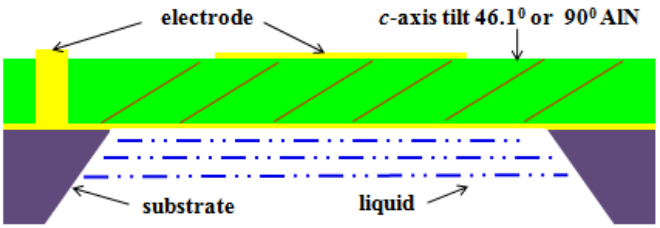


Fig. 2 Schematic of viscosity sensor based on *c*-axis tilted AlN FBAR.

As mentioned before, the existence of pure shear mode makes AlN FBARs possible for sensor application in liquid. Also, we know that for acoustic sensor, the higher resonant frequency usually means higher sensor sensitivity. Considering FBAR whose operation frequency can be up to tens of GHz, it is expected that sensor sensitivity will be

greatly improved. Fig. 2 shows the schematic of viscosity sensor based on AlN FBAR with *c*-axis tilt angle 46.1° or 90°.

For analyzing sensor performance of shear FBAR for viscosity measurement, we need derive the impedance expression of FBAR with liquid loading. Assume the interface of liquid and resonator is at $x_3' = 0$, now the boundary condition (7) changes to

$$\begin{aligned} &T_5(h) = T_3(h) = T_3(0)=0 \\ &T_5(0) = -j\omega Z_{Liquid} u_1(0) \\ &Z_{Liquid} = \sqrt{j\rho_{Liquid}\omega\eta_{Liquid}} \end{aligned} \tag{9}$$

where ρ_{liquid} and η_{liquid} are the density and viscosity of liquid. From (1)-(6), (8), and (9), the impedance of shear AlN FBARs with liquid loading can be solved,

$$Z = \frac{1}{j\omega C_0}\left(1-\left(\frac{(k^S)^2}{\gamma^S}\frac{2\tan(\gamma^S/2)-j\dfrac{Z_{Liquid}}{Z_S}}{1-\dfrac{Z_{Liquid}}{Z_S}\cot(\gamma^S)}\right)\right) \tag{10}$$

where Z_S and γ_S are the characteristic impedance and phase delay of AlN; C_0 is the static capacitance; Z_{liquid} is the acoustic loading of liquid.

III. RESULTS AND DISCUSSION

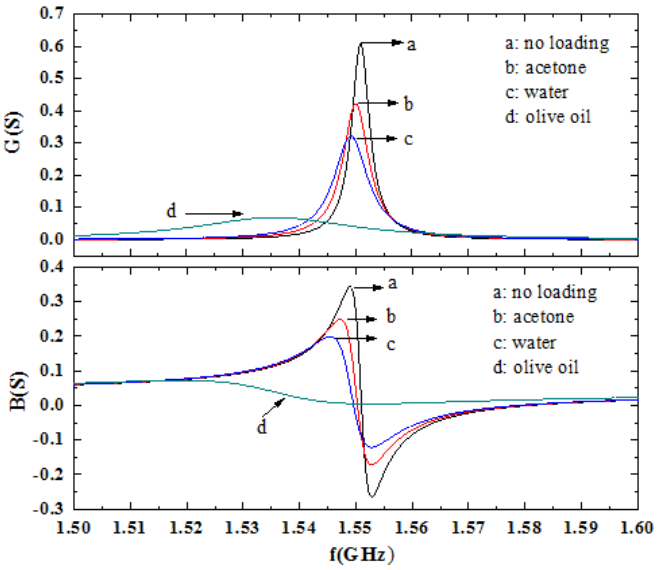
A. *Simulation of impedance spectra of shear mode AlN FBARs with liquids loading*

TABLE I. Parameters for simulation

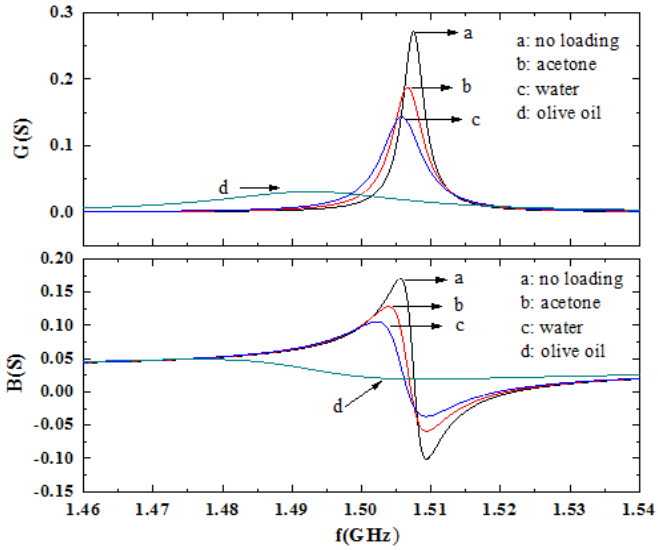
Piezoelectric layer	Mechanical quality factor	Thickness (μm)	Electrode area (μm×μm)
AlN	400	2	300×300
Liquid loading	**Temperature (^{0}C)**	**Density (kg/m^3)**	**Viscosity (mPa•S)**
Acetone	20	792.5	0.31
Water	20	1000	1
Olive oil	20	920	84

Using the electric impedance equation for FBAR with tilted polar c-axis orientation, the resonator spectrum with liquid loading can be simulated. The parameters used in the simulation are listed in Table I. Fig. 3 shows the calculated impedance spectra of FBARs with different liquid loadings. In the case of no liquid loading, it can be seen that the resonant frequencies of these two resonators (for tilt angle 46.1° or 90°) are quite similar. The resonant frequency is about 1.55GHz for AlN FABR with *c*-axis tilt angle 46.1°, and about 1.51GHz for AlN FABR with *c*-axis tilt angle 90°, respectively). The resonant frequency f_S for an acoustic wave resonator can be described by $f_S = \frac{\upsilon}{\lambda} = \frac{\upsilon}{2h}$, where υ is the acoustic velocity determined by material properties; the acoustic wave length (λ) is 2 times of the thickness (h) of resonator. Due to the crystal orientation dependence of material properties including elastic constants, piezoelectric constants and dielectric constants, the acoustic velocity changes with *c*-axis orientation. Hence the resonant frequencies of these two pure shear mode AlN FBARs are different, although they have same thickness.

In the case of liquid loading, for the two pure thickness shear mode resonators we can clearly see that the resonant frequency decreases due to liquid loading, and the resonant peak becomes small and broadened, indicating that the energy loss caused by viscous loading. The loading effect of olive oil on the FBARs is much more significant than that of water and acetone. This is attributed to the higher acoustic loading impedance of olive oil.



(a)



(b)

Fig. 3 Simulation of impedance spectrum of shear mode AlN (a) *c*-axis tilt angle 46.1^0; (b) *c*-axis tilt angle 90^0

B. Sensitivity of shear mode AlN FBAR viscosity sensor

The sensitivity (S) of shear mode AlN FBAR viscosity sensor can be defined as,

$$S = \frac{1}{f_0}\frac{df}{d\sqrt{\rho_{Liquid}\eta_{Liquid}}} = \frac{1}{f_0}\lim_{\Delta\eta\to 0}\frac{\Delta f}{\sqrt{\rho_{Liquid}(\eta_{Liquid}+\Delta\eta_{Liquid})}-\sqrt{\rho_{Liquid}\eta_{Liquid}}} \quad (11)$$

Here we assume that the liquid density does not change. f_0 is initial resonant frequency; Δf is resonant frequency shift due to viscosity change $\Delta\eta_{Liquid}$.

Table II lists the viscosity sensitivities of shear mode FBARs based on AlN thin films with *c*-axis tilt angle 46.1° and 90°, respectively, where the viscosity sensitivity was calculated through calculating resonant frequency shift for 1% viscosity change, and the resonant frequency is corresponding to the maximum conductance (G). It can be seen that the viscosity sensitivity of these two shear mode resonators are quite similar with a value of 1.12e-3 or 1.13e-3 $(kgm^{-3}Pa•S)^{-0.5}$, respectively, and they are almost same for different liquid loading material.

For comparison, we also calculated the viscosity sensitivity of a 6 MHz thickness shear mode quartz crystal resonator, which has a value around 1.43e-4 $(kgm^{-3}Pa•S)^{-0.5}$, also listed in Table II. It is found that the sensitivity of shear mode AlN FBAR is about 8 times as 6 MHz quartz crystal resonator, which seems a little bit low considering the huge difference between their resonant frequencies. One reason is that the definition of viscosity sensitivity given in Eq.(11), where the relative frequency shift is used instead of absolute frequency shift, which lowers the effect of resonant frequency on the value of viscosity sensitivity during calculation.

TABLE II. Viscosity sensitivities of thickness shear mode AlN FBARs

liquid	Viscosity sensitivity($(kgm^{-3}Pa•S)^{-0.5}$)		
	2um AlN FBAR (*c*-axis tilt 46.1°)	2um AlN FBAR (*c*-axis tilt 90°)	6MHz Quartz resonator
Acetone	1.12e-3	1.12e-3	1.428e-4
Water	1.12e-3	1.13e-3	1.428e-4
Olive oil	1.13e-3	1.13e-3	1.430e-4

IV. CONCLUSION

In summary, viscosity sensitivity of thickness shear mode AlN FBARs with *c*-axis tilt angle 46.1° and 90° has been theoretically studied. The electric impedance for thickness shear mode AlN FBARs with a liquid layer has been derived from the basic piezoelectric constitutive equations. Viscosity sensitivities for three liquid loading (water, acetone and olive oil) were achieved by the calculation of relative resonant frequency shift for 1% viscosity change. The viscosity sensitivities of shear mode 2um AlN FBARs with *c*-axis tilt angle 46.1° and 90° are similar, with a value around 1.12e-3 or 1.13e-3 $(kgm^{-3}Pa•S)^{-0.5}$, and they do not change much with these three different liquid loading. Viscosity sensitivity of 2μm shear mode AlN is found about 8 times as 6MHz TSM quartz crystal resonator. The simulation results can be used for FBARs design and sensor application.

REFERENCES

[1] K. M. Lakin, G. R. Kline, R. S. Ketcham, and S. G. Burns, "Thin film resonator based low insertion loss filters," IEEE Ultrasonic Symposium, pp. 371–376, 1986.

[2] K. M. Lakin, "Thin film resonators and filters," IEEE Ultrasonic Symposium, Vol. 2, pp. 895–906, 1999.

[3] R. Weigel, D. P. Morgan, J. M. Owens, A. Ballato, K. M. Lakin, K.-y. Hashimoto, and C. C. W. Ruppel, "Microwave Acoustic Materials, Devices, and Applications," IEEE Transactions on Microwave Theory and Techniques, vol. 50, pp. 738-749, Mar. 2002.

[4] R. Gabl, H. -D. Feucht, H. Zeininger, G. Eckstein, M. Schreiter, R. Primig, D. Pitzer, and W. Wersing, "First results on label-free detection of DNA and protein molecules using a novel integrated sensor technology based on gravimetric detection principles," Biosensors and Bioelectronics, vol. 19, pp. 615-620, Jan. 2004.

[5] H. Zhang, M. S Marma, E. S. Kim, C. E McKenna, and M. E Thompson, "A film bulk acoustic resonator in liquid environments," J. Micromech. Microeng., vol. 15, no. 10, pp. 1911–1916, 2005

[6] M. Benetti, D. Cannatà, and F. Di Pietrantonio, "Microbalance chemical sensor based on thin-film bulk acoustic wave resonators," Appl. Phys. Lett. 87, 173504 (2005).

[7] Y. Chen, P. I. Reyes, Z. Duan, G. Saraf, R. Wittstruck, Y. Lu, O. Taratula and E. Galoppini, "Multifunctional ZnO-Based Thin-Film Bulk Acoustic Resonator for Biosensors," Journal of ELECTRONIC MATERIALS, vol . 38, pp. 1605-1611, 2009.

[8] G. Wingqvist, J. Bjurström, L. Liljeholm, I. Katardjiev, and A. L. SpetZ, "Shear mode AlN thin film electroacoustic resonator for biosensor applications," IEEE Sensor Proceedings, pp. 492-495, 2005

[9] J. Bjurström, G. Wingqvist, and I. Katardjiev, "Synthesis of Textured Thin Piezoelectric AlN Films With a Nonzero C-Axis Mean Tilt for the Fabrication of Shear Mode Resonators," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 53, pp. 2095-2100, Nov. 2006.

[10] G. Wingqvist, J. Bjurström, L. Liljeholm, V. Yantchev, and I. Katardjiev, "Shear mode AlN thin film electro-acoustic resonant sensor operation in viscous media," Sensors and Actuator B, vol. 123, pp. 466-473, 2007.

[11] F. Martin, M.-E. Jan, S. Rey-Mermet, B. Belgacem, D. Su, M. Cantoni, and P. Muralt, "Shear Mode Coupling and Tilted Grain Growth of AlN Thin Films in BAW Resonators," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 53, pp. 1339-1343, Jul. 2006.

[12] F. Martin, M.-E. Jan, B. Belgacem, M.-A. Dubois, P. Muralt, "Shear mode coupling and properties dispersion in 8GHz range AlN thin film bulk acoustic wave (BAW) resonator," Thin Solid Films, vol. 514, pp. 341-343, 2006.

[13] C. Chung, Y. Chen, C. Cheng, and K. Kao, "Synthesis and Bulk Acoustic Wave Properties on the Dual Mode Frequency Shift of Solidly Mounted Resonators," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 55, pp. 857-864, Apr. 2008.

[14] L. Qin, Q. Chen, H. Cheng and Q.-M. Wang, "Analytical study of dual-mode thin film bulk acoustic resonators (FBARs) based on ZnO and AlN films with tilted c-axis orientation," Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions , vol.57, no.8, pp.1840-1853, August 2010

[15] N. F. Foster, G. A. Coquin, G. A. Rozgonyi, F.A.Vannatta, "Cadmium Sulphide and Zinc Oxide Thin-Film Transducers," IEEE Transactions on Sonics and Ultrasonics, vol. 15, pp. 28-40, Jan. 1968

[16] H. F. Tiersten, "Linear Piezoelectric Plates Vibrations," New York: Plenum Press, 1969.

A new QCM Sensor Characterization Technique Based on the Phase/Mass Sensitivity Concept

Yeison Montagut[1], José.V. García[1], Yolanda Jiménez[1], Carmen March[2], Ángel Montoya[2], Antonio Arnau[1]
[1]Departamento de Ingeniería Electrónica
[2]Instituto Interuniversitario de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano
Universitat Politècnica de València
Valencia, Spain
yeimonfe@doctor.upv.es ; jogarnar@upvnet.upv.es

***Abstract*—Quartz crystal microbalance (QCM) technique is widely used in in-liquid biochemical applications. The main challenges remain on the improvement of sensitivity and limit of detection, as well as multianalysis capabilities and reliability. The improvement of sensitivity has been addressed in the last decades by increasing the sensor fundamental frequency; following the increment of the frequency/mass sensitivity with the square of frequency predicted by Sauerbrey. However, this sensitivity improvement has not been completely transferred in terms of limit of detection. A new concept of sensor characterization has been recently proposed based on the phase/mass sensitivity equation: $\Delta\varphi/\Delta m \approx -1/m_L$, where m_L is the liquid mass perturbed by the resonator in its oscillatory movement; this mass reduces proportionally with the frequency square root. The validation of the new concept is presented in this work. An immunosensor application for the detection of the very low molecular weight insecticide, carbaryl, has been chosen for the validation of the characterization technique.**

I. INTRODUCTION

Acoustic sensing has taken advantage of the progress made in the last decades in piezoelectric resonators for radio-frequency (RF) telecommunication technologies. Acoustic sensor based techniques combine their direct detection, real-time monitoring, high sensitivity and selectivity capabilities with a reduced cost in relation to other techniques. The so-called quartz crystal microbalance (QCM) technique, where the AT cut quartz crystal resonator is extensively used, has been classically based on the proportional change in the resonance frequency experimented by the resonator due to a surface mass density change on the sensor surface; the theoretical absolute mass sensitivity for this frequency shift is proportional to the square of the resonant frequency according to the expression [1]

$$S_a = \frac{\Delta f}{\Delta m} = -\frac{2}{\rho v}\frac{f_n^2}{n} \qquad (1)$$

where Δf is the frequency shift, Δm is the surface mass density change on the active sensor's surface, ρ is the quartz density, v the propagation velocity of the wave in the AT cut quartz crystal, f_n is the frequency of the selected harmonic resonant mode and n is the harmonic number (n=1 for the fundamental mode). In QCM, acoustic waves are excited by a sinusoidal voltage applied to an electrode structure where the quartz crystal is sandwiched; shear waves are excited which makes the operation in liquids viable [2]. This has opened a great deal of applications in bio-chemical sensing in liquid media.

II. PHASE-MASS CHARACTERIZATION FOR IN-LIQUID QCM APPLICATIONS

According to (1), the improvement of sensitivity has been addressed in the last decades by increasing the sensor fundamental frequency. However, this sensitivity improvement has not been completely transferred in improvements of the resolution or limit of detection. The decrease on frequency stability in the measurement due to the increase in phase noise and unwanted frequency jumps to spurious inharmonic modes, particularly in the case of oscillators, seems to be the reason behind the difficulty to reach the expected maximum resolution [3-5].

In QCM based biosensors the experimental frequency shifts expected are usually small, in the order of tens of Hertz. Therefore, the great efforts performed to improve the sensitivity of the sensor by increasing the frequency, only would be useful if they were accompanied with an increase in the limit of detection [6]. Once external perturbations on the sensor have been minimized (temperature and pressure variations, external vibrations, etc.), resolution on the measurement relies directly on the interface characterization system, its noise and stability and the measurement uncertainty associated to it. In oscillators, a classical electronic interface system for QCM sensors, the origin of frequency instability is the phase instability [7]; the high quality factor of the crystal resonator in RF oscillators allows absorbing the phase instabilities in the circuitry with very small shifts in the

978-1-61284-110-6/11/$26.00 ©2011 IEEE

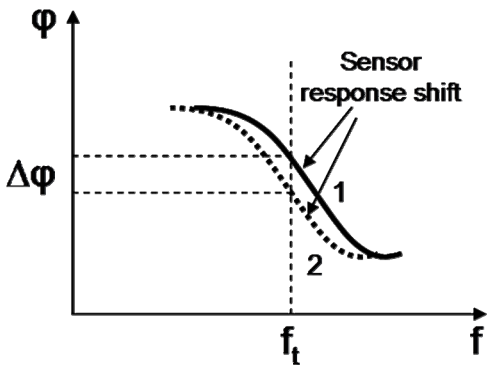


Figure 1. Phase-shift concept representation

oscillation frequency of the system. But in QCM sensor oscillators the resonator is not included in the circuitry with the main purpose of stabilizing the frequency of the oscillating system, although effectively it does it. Consequently, the oscillator circuitry should be as ideal as possible for not influencing the sensor phase shifts due to the processes to be monitored. Unfortunately to implement a very stable oscillator for high frequency QCM sensors, keeping in mind the low quality factors reached by these sensors in liquid conditions is not an easy and straightforward task.

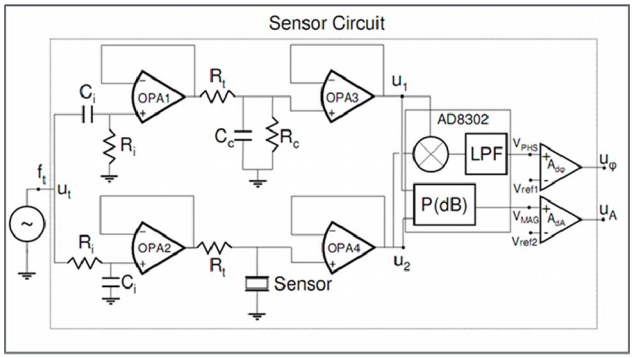
Having this idea in mind, a new concept of characterization, where the sensor is interrogated with a signal at constant frequency, has been recently proposed, based on the following phase-mass sensitivity equation for the sensor in liquid media [8]:

$$\Delta\varphi = -\frac{\Delta m_c}{m_q + m_L} \tag{2}$$

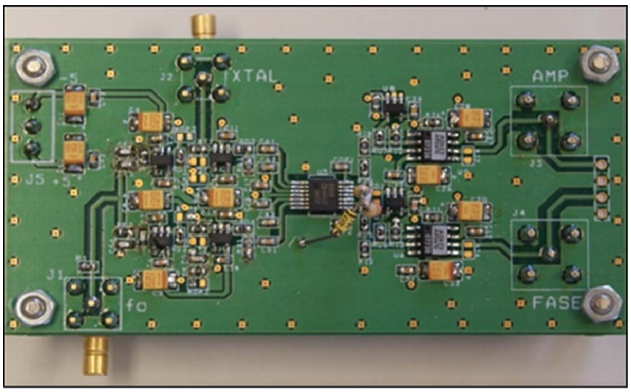
where m_c is the surface mass density of the coating, $m_q=\eta_q\pi/2v_q$, being η_q the effective quartz viscosity, v_q the wave propagation speed in the quartz and $m_L=\rho_L\delta_L/2$, where ρ_L and δ_L are, respectively, the liquid density and the wave penetration depth of the acoustic wave in the liquid. This last term, m_L, is the liquid mass surface density perturbed by the resonator in its oscillatory movement and reduces proportionally to the frequency square root due to the decrease of the penetration depth. As it was also described [8], when $m_L >> m_q$, (2) can be expressed in a simpler form,

$$\Delta\varphi \approx -\Delta m_c / m_L \,. \tag{3}$$

Taking into account that the expected frequency shifts in high resolution QCM applications are very small and assuming no variations on the liquid properties (conditions usually met in biosensor applications), it is possible to interrogate the sensor with a test signal of the appropriate constant frequency, in the resonance bandwidth, and measure the change in the phase response of the sensor due to the experimental process to be monitored, without leaving the maximum sensitivity region of phase response governed by (3). Fig.1 depicts this idea.



(a)



(b)

Figure 2. Characterization interface for in-liquid qcm sensor. (a) Schematics and (b) printed circuit board implemented

III. SYSTEM DESCRIPTION

A sensor interface circuit, which was theoretically proposed in [8], has been implemented to validate this characterization concept experimentally. Fig.2 shows the schematic diagram and the final implementation of the sensor interface. The circuit is mainly composed of two parallel and quite similar branches in order to form a differential circuit. Because the testing signal u_t has constant frequency, f_t, the only element in the circuit which contributes to a change in the phase difference between the reference signal u_1 and the signal u_2 is the change in the phase-frequency response of the sensor. This phase difference can be easily and continuously monitored by a phase detector, like the one found in the IC AD8302 from Analog Devices. Being a multiplier type phase detector, linearity and error in the measurement are maximum and minimum, respectively, for small phase differences around 90º between signals at its input. The networks formed by R_i and C_i at the input of the sensor circuit have been included to obtain this initial phase difference of 90º between signals at each branch of the circuit. These phase-shifting networks must be designed coherently with the resonant frequency of the sensor in order to obtain two signals 90º phase-shifted and of similar amplitude. In the case of the AD8302, the detector response is centered on 900mV, with a phase slope of approximately 10mV/º. A conditioning circuit, comprising a low noise differential amplifier and an adjustable reference voltage, is used to re-center and amplify the detector response around 0V with a response sensitivity of 100mV/º.

Figure 3. Experimental Setup

AD8302 also includes a matched pair of logarithmic demodulating amplifiers allowing measuring the magnitude ratio of the signals connected to it, with a response of 30mV/dB. The magnitude ratio voltage is also re-centered and amplified with similar conditioning circuits. Wide bandwidth operational amplifiers OPA1-4 are used to isolate the sensor and the reference network R_c-C_c from the rest of the circuit. At motional series resonance frequency (MSRF), the sensor reduces to a motional resistance, R_m, in parallel with the so-called static capacitance C_o. Therefore for optimum operation it is convenient to select R_c and C_c similar to R_m and C_0, respectively. Under these conditions and interrogating the sensor with a test signal of a constant frequency f_t around the MSRF, the voltage u_φ corresponding to the phase-shift measured should be the corresponding to the initial phase difference, 0mV, after proper reference adjustment.

IV. Experimental Setup

A real application based on an immunosensor for the detection of a low molecular weight pollutant, the insecticide carbaryl [9], has been chosen as validation model. The proposed system was compared with a QCM system based on an improved oscillator configuration [10] using 10MHz AT cut QCM sensors (International Crystal Manufacturing), the same validation protocol described elsewhere [9] and under the same experimental conditions. BSA-CNH carbaryl hapten conjugate was immobilized on the sensor surface through the formation of a thioctic acid self-assembled monolayer. The crystal was placed in a custom-made flow-cell and included in a fluidic system comprising also a peristaltic pump (Minipulse 3, Gilson), and loop injection and solution selection manual valves (Supelco). The sensor characterization circuit and the whole fluidic system, but the peristaltic pump, were placed in a thermostatic chamber. This thermostatic chamber was placed on an anti-vibration table in order to avoid external disturbances. The test signal was generated with a RF signal generation model Agilent-HP8664A (Agilent) and the voltage was measured with a multimeter HP34401A (Agilent) connected via GPIB bus to a PC running the NI-Visa Server Software (National Instruments) allowing a remote connection to the test equipment via Ethernet. Client software was developed to acquire and represent the data in a remote PC. This experimental setup is depicted in Fig. 3.

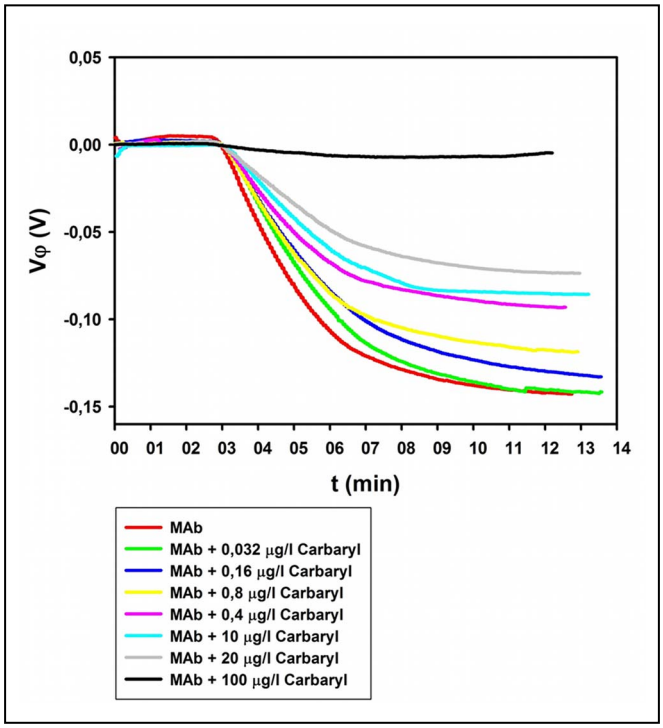


Figure 4. Piezoelectric immunosensor responses to different concentrations of analyte

The immunoassay developed was an inhibition test based on the conjugate coated format. Different standard concentrations of carbaryl were prepared and mixed with a fixed concentration of the monoclonal antibody LIB-CNH45 (I3BH-UPV) [11] in phosphate buffered saline (PBS). Since the analyte inhibits antibody binding to the immobilized conjugate, increasing concentrations of analyte will reduce the magnitude of the phase shift induced on the sensor. The phase shift was monitored in real-time for each sample injection during 12 min. A regeneration stage was performed after this reaction time by consecutive injections of hydrochloric acid (HCl) and phosphate buffered saline - tween 20 (PBST) at appropriate concentrations, breaking the antibody-hapten linkage and reaching the initial base-line.

V. Results

Fig. 4 shows the signals measured with the characterization system proposed for the injection of different standard concentrations of carbaryl. The voltage u_φ associated with the phase-shift, decays exponentially as soon as the molecular interaction occurs after each sample injection. As we can see, results show the typical response of a competitive immunoassay. The voltage decrease is lower as we increase the concentration of the analyte, due to the lower amount of antibody available in the sample being able to bind with the carbaryl hapten conjugate immobilized in the sensor surface.

These results validate the new characterization concept and the implemented interface. A similar behavior is observed when the resonant frequency shift is monitored. Moreover, a reduction of the noise in the new system was observed compared with the oscillator technique, under the same experimental conditions. The noise level for this phase-shift

interface was of 1mV for maximum signals that reached 200mV. Furthermore, it is important to notice that this improvement has been got even with relative low frequency sensors (10MHz), where electronic components and circuits have a very good performance in both, the oscillator and the phase-detector system. Therefore a significant improvement is expected to be found with very high fundamental frequency resonators.

VI. CONCLUSIONS

The new characterization concept, particularly useful for biosensor applications, has been experimentally validated in a real application, an immunosensor for the detection of the low weight pollutant carbaryl. The sensor is interrogated passively with an external source, which can be designed with high frequency stability and very low phase noise, even at very high frequencies. The interface is very simple and allows a straightforward noise analysis, simplifying the electronic design and implementation. This simplicity and its high integration capability would allow its use with multiple sensors working at the same fundamental resonance frequency, which could be characterized, in principle, with only one source, only repeating the sensor circuit, opening the possibility of working with sensor arrays for multianalysis detection.

ACKNOWLEDGMENTS

This work was developed under the contract reference AGL2009-13511. J.V. Garcia thanks the Spanish “Ministerio de Educación” for his PhD student fellowship, under the F.P.U. program. Authors are grateful to the company Advanced Wave Sensors S.L. (www.awsensors.com) for the help provided in the development of some parts of this work.

REFERENCES

[1] G. Sauerbrey, "Verwendung von schwingquarzen zur wägung dünner schichten und zur mikrowägung", Zeitschrift Fuer Physik, vol. 155, (2), pp. 206-222, April 1959.

[2] K. K. Kanazawa, J.G. Gordon II, "The oscillation frequency of a quartz resonator in contact with a liquid", Anal. Chim. Acta, vol. 175, pp. 99-105, 1985.

[3] B. Zimmermann, R. Lucklum, P. Hauptmann, J. Rabe, and S. Büttgenbach, "Electrical characterisation of high-frequency thickness-shear-mode resonators by impedance analysis", Sens. Actuators B, vol. 76, (1-3), pp. 47-57, June 2001.

[4] E. Uttenthaler, M. Schräml, J. Mandel, S. Drost, "Ultrasensitive quartz crystal microbalance sensors for detection of M13-Phages in liquids", Biosens. Bioelectron., vol. 16, (9-12), pp. 735-743, December 2001.

[5] L. Rodriguez-Pardo, J. Fariña, C. Gabrielli, H. Perrot, and R. Brendel, "Resolution in quartz crystal oscillator circuits for high sensitivity microbalance sensors in damping media", Sens. Actuators B, vol. 103, (1-2), pp. 318-324, September 2004.

[6] J.R. Vig, F. L. Walls, "A Review of sensor sensitivity and stability", Proc. 2000 IEEE/EIA Int. Frequency Control Symp. Exhibition, pp. 30 – 33, 2000.

[7] A. Arnau, V. Ferrari, D. Soares, and H. Perrot, “Interface Electronic Systems for AT-Cut QCM Sensors: A comprehensive review,” in Piezoelectric Transducers and Applications, 2nd ed., ch.5, pp. 117-186, A Arnau ed., Springer Verlag Berlin Heidelberg, 2008.

[8] A. Arnau, Y. Montagut, J. V. García, and Y. Jiménez, "A different point of view on the sensitivity of quartz crystal microbalance sensors", Meas. Sci. Technol., vol. 20, p.124004, December 2009.

[9] C. March, J. J. Manclús, Y. Jiménez, A. Arnau, and A. Montoya, "A piezoelectric immunosensor for the determination of pesticide residues and metabolites in fruit juices”, Talanta, vol. 78, (3), pp. 827-833, May 2009.

[10] Y. Montagut, J.V. García, Y. Jiménez, C. March, A. Montoya, and A. Arnau, “Frequency-Shift vs Phase-Shift Characterization of in-liquid Quartz Crystal Microbalance Applications”, Rev. Sci. Inst., in press.

[11] A. Abad, J. Primo, and A. Montoya, “Development of an enzyme-linked immunosorbent assay to carbaryl. 1. Antibody production from several haptens and characterization in different immunoassay formats”. J. Agric. Food Chem., vol. 45, 1486-1494, April 1997.

Characterization of Rotational Mode Disk Resonator Quality Factors in Liquid

Amir Rahafrooz and Siavash Pourkamali
Department of Electrical and Computer Engineering
University of Denver
Denver, CO, USA
E-mail: amir.rahafrooz@du.edu

***Abstract*— This work presents several variations of microscale single and multiple coupled thermally actuated rotational mode disk resonators capable of high quality factor (Q) operation in liquid. All surfaces of a rotational mode disk resonator slide in parallel to the liquid interface (rather than stroking against it). This minimizes interactions with the surrounding fluid and therefore viscous damping. Quality factors as high as 304 have been demonstrated for rotational mode disk resonators in liquid, which is more than 3X higher than the highest values previously demonstrated for other types of microscale resonators. Based on the measurement results it is suggested that most of the energy loss into the surrounding liquid is through the sidewalls of the disks. This could be blamed on the sidewall roughness caused by deep reactive ion etching, off-center rotation of the disks due to fabrication induced asymmetries, and slight bending of support beams during resonance. Much higher Q values are expected to be achievable by minimizing such limiting factors.**

I. Introduction

As highly sensitive mass sensors, micromechanical resonators can potentially open up a wide range of new opportunities in biomedical and chemical sensing applications leading to more compact low cost instruments with real-time sensing capabilities. In addition, due to their small size, micro resonators can be integrated in arrays of hundreds to thousands sensing a variety of analytes simultaneously. Most of the biosensing applications require detection and measurement of certain molecules in a liquid solution (typically aqueous). However, the resonant behavior of microscale resonators drastically deteriorates in liquid lowering their sensitivity and mass resolution. This is mainly due to the larger surface to volume ratio for such small scale devices and consequently large viscous damping to elastic energy ratio in them.

Over the past decade efforts have been made to operate micromechanical resonators inside liquid. Magnetically actuated cantilever beams resonating in their out-of-plane flexural modes have been used for mass sensing applications in aqueous environments [1-4]. It is not hard to imagine that such a paddling motion faces a lot of resistance from the surrounding liquid leading to excessive viscous damping. This reduces the quality factor of such devices to values as low as 5[1], 23[2], 2[3] and 10[4] in water. Viscous damping can be reduced by operating cantilever beams in their in-plane flexural resonant mode. Quality factors as high as 67 have been demonstrated for such devices in liquid [5]. Higher quality factor of the in-plane resonant modes for such thin structures can be attributed to having a much smaller portion of their surface area stroking against the liquid interface compared to their out of plane modes. In an improved design by Seo, et al [6], a structure comprised of two half-disks resonating in an in-plane quasi rotational mode, demonstrated quality factors as high as 94 in water. Despite having improved quality factors, similar to the in-plane flexural modes of microcantilevers, the straight sides of the half disks in such structures still have a perpendicular motion against the liquid interface. In [7] we demonstrated full disk structures with two tangential support beams thermally actuated in their rotational resonant modes. In such structures all the surfaces slide in parallel to the liquid interface exhibiting Qs as high as ~200 in water.

In this work, five different variations of single and multiple-coupled rotational mode disk resonators are studied to further understand the loss mechanism and optimize the performance of such structures.

II. Device Operation

Figure 1 shows the schematic view of a single-disk with straight tangential support beams as well as the electrical connections used to operate such device. The support beams act as both thermal actuators and piezoresistive stress sensors simultaneously [7]. Thermal actuation occurs by passing a current between the two pads on the two sides of the structure. By passing a combination of a DC and an AC current, the ohmic power loss will have a component at the same frequency as the applied AC current: $P_{ac} = 2R_e I_{dc} i_{ac}$, where R_e is the electrical resistance between the two pads and I_{dc} and i_{ac} are the applied DC and AC currents respectively. Due to their higher electrical resistance, most of the ohmic loss and therefore heat generation is concentrated in the narrow support beams. The applied AC power produces

This work was supported by National Science Foundation under grants MRI-RAPID #1061489 and MRI #0923518.

periodic temperature fluctuations in the thermal actuators (Fig. 1) which in turn causes alternating stress and strain in the support beams, actuating the disk in its rotational resonance mode (periodically rotating back and forth). As the resonator vibrates the resulting periodic stress changes result in fluctuations in the electrical resistance of the resonator due to the piezoresistive effect. This modulates the DC current passing through the resonator leading to an AC output motional current component that can be used to monitor the resonator vibration amplitude.

While the disk is vibrating in its rotational mode, the support beams vibrate in their extensional mode (periodically elongate and contract). As a result, all the surfaces of both the disk and its support beams move in parallel to the liquid interface, minimizing the energy loss to the surrounding liquid. Figure 2a shows the COMSOL modal analysis results showing the mode shape for a 100µm diameter, 20µm thick disk with 4µm wide, 26µm long support beams.

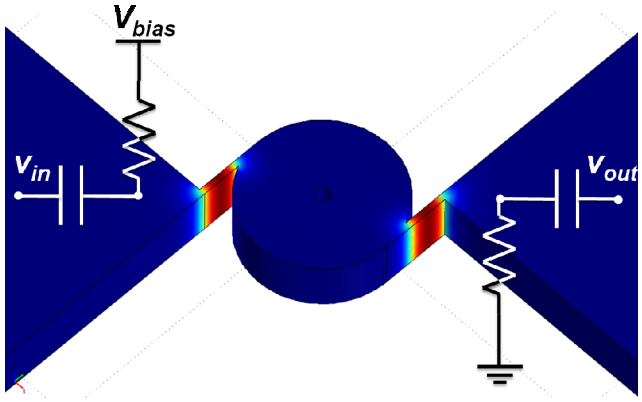


Fig. 1. Schematic view of a thermally actuated disk resonator showing the qualitative distribution of AC temperature fluctuation amplitude in the resonator (red being the maximum and blue minimum). The electrical connections required for one-port operation of the resonator are also shown.

III. Fabrication and Measurement Results

In this work, five different variations of single and multiple-coupled rotational mode disk resonators were fabricated and characterized (Fig. 2); 1. Single disks with straight tangential support beams, 2. Single disks with rounded support beams, 3. Parallel dual-disks, 4. Series dual-disks, and 5. Interconnected quad-disks. They were fabricated using the standard single mask SOI-MEMS process [8]. The fabrication process includes silicon DRIE to form the structures out of the silicon device layer, and releasing them by etching the underlying buried oxide (BOX) layer in hydrofluoric acid (HF). Resonators were fabricated on low resistivity N-type substrates with different device layer and BOX thicknesses. To optimize resonator electromechanical transduction, the support beams are aligned to the <100> crystalline direction where the longitudinal piezoresistive coefficients are maximum [8]. The holes in the middle of the disks are to reduce the undercut time in HF. Due to circular shape of the holes and relatively small vibration amplitude in the center of the disks such holes are expected to have negligible effect on viscous damping of the resonator.

The resonators were tested in a one-port configuration as shown in Fig. 1. The liquid tests were performed in heptane to provide better electrical isolation and avoid contamination resulting from electrolysis of water. Figure 3 shows the measured frequency responses for the 100µm single disk

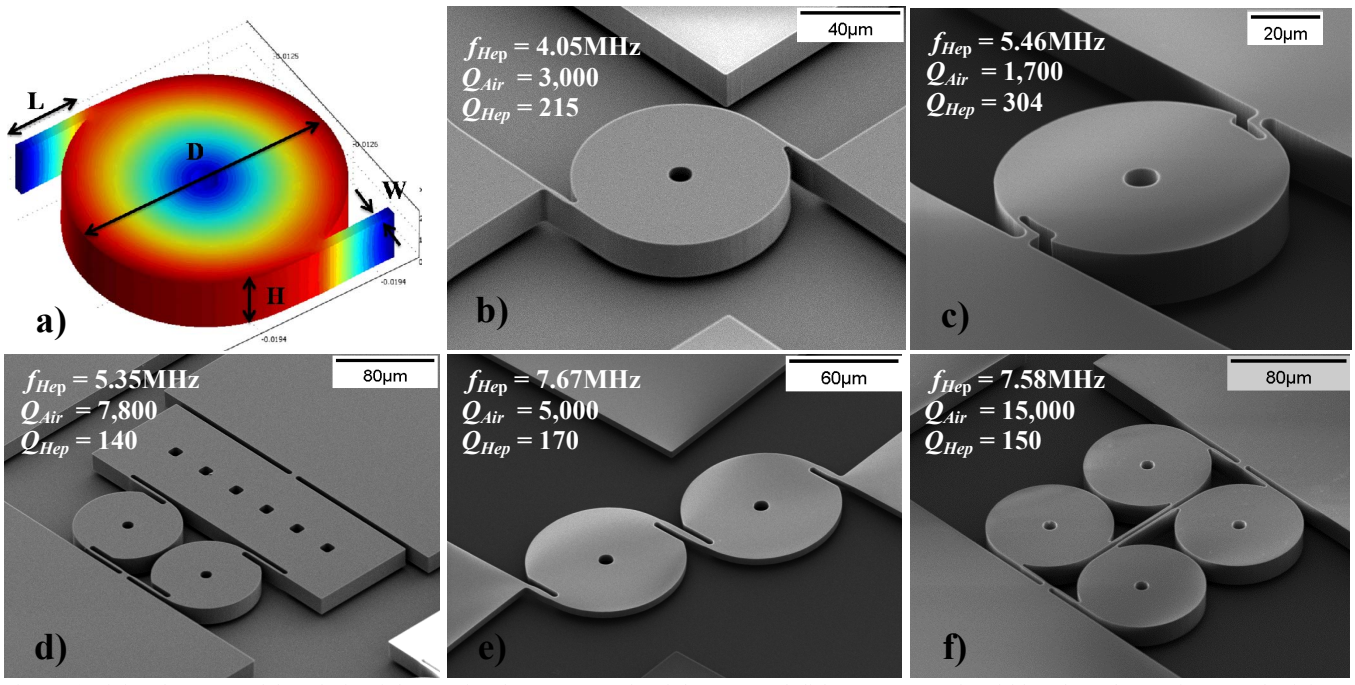


Fig. 2. a) Finite element modal analysis of the rotational resonant mode of the proposed disk structure. Red and blue show the maximum and minimum vibration amplitudes respectively. SEM view of fabricated devices along with their measured Q in both air and liquid for: b) single disk with straight tangential support beam, c) single disk with rounded support beam, d) parallel dual-disk, e) series dual-disk, f) interconnected quad-disk. The thermal actuators are along <100> crystalline direction for optimized transduction.

resonator of Fig. 2c with rounded thermal actuators both in air and heptane. As expected, the resonator has relatively low quality factors in air (due to excessive support loss), however, unprecedented quality factor of 304 was measured in heptane. Such high Q values in heptane confirm our hypothesis that elimination of the stroking surfaces from the mode shape can significantly improve resonator Q in liquid.

Table I summarizes the measurement results for a variety of resonators of different types with different dimensions in both air and liquid. It also shows the finite element modal analysis for different structures demonstrating the vibration amplitude at different parts of the structure. The modal analysis results show that unlike the other structures the interconnected quad-disks have an off-center rotation which justifies their lower Q in liquid despite their relatively high Q in Air.

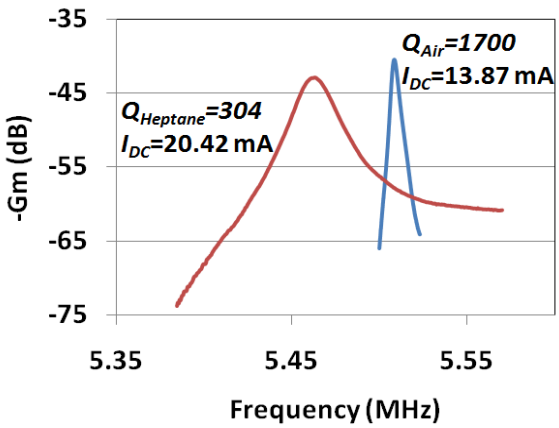
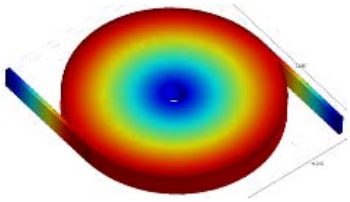
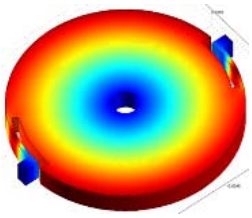
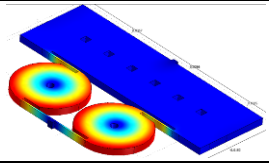
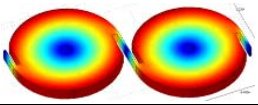
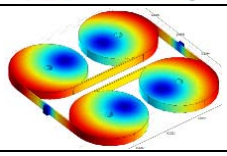


Fig. 3. Measured frequency response for the 100µm diameter 20µm thick disk resonator shown in Fig. 2c with rounded support beams in air and heptanes.

TABLE I MEASUREMENT RESULTS, AND MODE SHAPES FOR DIFFERENT TYPES OF SINGLE AND MULTIPLE DISK ROTATIONAL MODE RESONATORS. MEASUREMENT RESULTS IN BOTH AIR AND LIQUID (HEPTANES) ARE PRESENTED (W=2µ METER FOR ALL THE SUPPORT BEAMS).

Type	Finite Element Modal Analysis Showing Resonant Mode Shape	Resonator Dimensions (µm)			Freq.$_{Air}$ (MHz)	Q.$_{Air}$	Freq.$_{Hep.}$ (MHz)	Q.$_{Hep.}$
		D	L_{th}	*H*				
Single Disk with Straight tangential support beam		100	42	5	5.500	3500	5.500	145
				10	5.144	1900	5.144	175
				20	4.052	3000	4.052	215
			17	5	6.827	2000	6.827	155
				10	6.427	1100	6.427	180
				20	5.356	1100	5.356	180
		200	55.5	5	2.586	4200	2.586	90
				10	2.383	700	2.383	170
				20	1.767	11700	1.767	110
			35	5	3.368	2800	3.368	100
				10	2.849	3000	2.849	180
				20	3.056	1400	3.056	220
Single Disk with Rounded support beam		100	24	10	7.174	3500	7.174	235
				20	4.152	3800	4.152	230
			11	5	7.366	1400	7.366	185
				10	8.406	600	8.406	195
				20	5.463	1700	5.463	304
		200	40	5	2.781	4900	2.781	95
				10	2.911	2600	2.911	230
				20	1.851	8500	1.851	160
			24	5	3.394	4000	3.394	120
				10	3.513	1100	3.513	150
				20	2.616	5500	2.616	220
Parallel Dual-Disk		100	114	5	5.594	8000	5.594	125
				10	5.627	6500	5.627	105
				20	3.973	2000	3.973	180
			103	10	5.350	7800	5.350	140
				20	3.752	2300	3.752	140
Series Dual-Disk		100	42	5	7.673	5000	7.673	170
				10	7.396	5000	7.396	180
				20	5.632	3200	5.632	180
Interconnected Quad-Disk		100	84	5	7.517	13500	7.517	155
				10	7.581	15000	7.581	150
				20	5.386	4800	5.386	50

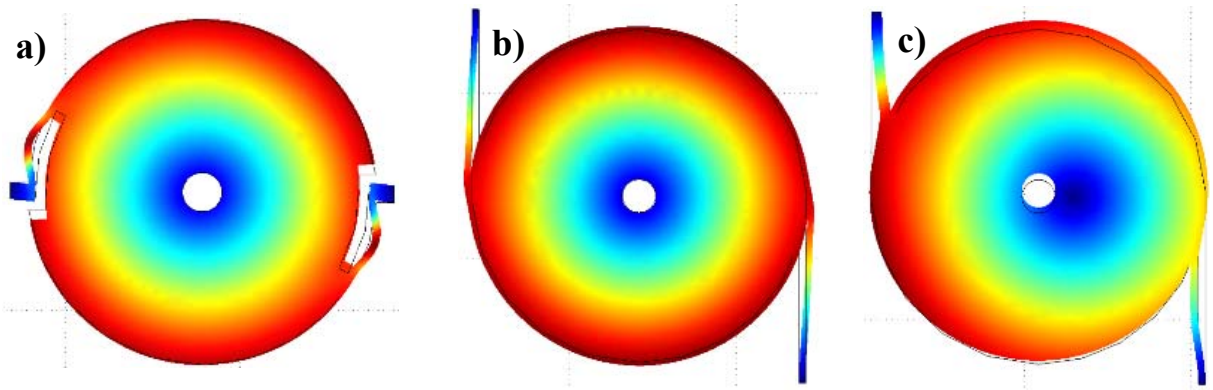


Fig. 4. Finite element modal analysis showing different effects potentially limiting resonator quality factors in liquid: a,b) bending support beams stroking against the surrounding liquid, c) mismatched support beams causing off-center resonance. In this extreme case, the support beams have widths of 2µm and 3µm and equal lengths.

The axial stress in all the support beams of the single disk designs and half of the support beams in dual-disks is directly transferred to their clamping points and consequently into the substrate. Therefore, in air, where Q is mainly limited by support loss, the Q of single disk designs is generally smaller than that of the dual-disks, which is in turn smaller than that of the quad-disks. In liquid however, viscous damping becomes the dominant source of loss and the Q for all designs drops to values in the 100-300 range. Although the highest Q of 304 in heptane obtained from the single disk structure with rounded supports of Fig. 2c, not much of a difference was observed between the performance of single disks with straight and rounded thermal actuators. Although higher quality factors are expected from the thicker resonators due to their smaller surface to volume ratio, measurement results do not show any significant improvement in quality factor of 20µm thick devices compared to 5µm thick devices. This suggests that most of the viscous damping is associated with the resonator sidewalls rather than their upper and lower surfaces. Resonator sidewall area is proportional to the resonator thickness and consequently the amount of elastic energy stored in the resonator. Therefore, if the sidewalls are the main source of loss no significant dependence of resonator quality factor to its thickness is expected. This could be blamed on the sidewall roughness caused by deep reactive ion etching, slight bending of support beams during resonance and off-center rotation of the disks due to fabrication induced asymmetries. Figure 4a and b demonstrate the bending of the rounded and straight support beams during the rotational vibration of the disks and Fig. 4c shows the effect of a fabrication induced asymmetry on the resonance mode of single disks which causes off-center rotation and therefore stroking against the liquid interface.

IV. Conclusions and Future works

Five different variations of single and multiple coupled disk resonators were fabricated and successfully operated in heptane with relatively high quality factors. The highest Q of 304 in liquid was obtained from a single disk structure with rounded support beams, which is 3X larger than the highest values previously reported for other types of micro-resonators. In addition, by studying the performance of different disk resonators with different structures and dimensions, the main energy loss mechanisms in liquid were further investigated. No significant improvement in the quality factor values for rotational mode disk resonators was observed by increasing their structural thickness. This suggests that such resonators mainly lose their energy through their sidewalls rather than upper and lower surfaces..

Micromechanical resonators with such high quality factors in liquid can enable direct high precision sensing of different analytes in liquid media paving the way towards label-free real-time biochemical sensors with fully electronic readout.

Future work includes further characterization and design optimization of the disk resonators to lower their energy loss in liquid. For instance, the structure of supporting beams can be optimized to reduce their stroking against the liquid.

Acknowledgment

Authors would like to thank staffs at Georgia Tech nanotechnology research center and also Dr. Ashwin Samarao for their help with silicon deep reactive ion etching.

References

[1] A. Vidic, D. Then, and C. Ziegler, “A new cantilever system for gas and liquid sensing,” Ultramicroscopy, vol. 97, no. 1, pp. 407–416, Oct. 2003.

[2] Y. Li, C. Vancura, et al, “Very high Q-factor in water achieved by monolithic, resonant cantilever sensor with fully integrated feedback,” in Proc. IEEE Sensors Conf., 2003, pp. 809–813.

[3] J. Tamayo, A.D.L. Humphris, A.M. Malloy, M.J. Miles, “Chemical sensors and biosensors in liquid environment based on microcantilevers with amplied quality factor,” Ultramicroscopy, vol. 86, no 1-2, pp.167-173, Jan 2001.

[4] C. Vancura, Y. Li, J. Lichtenberg, K.U. Kirstein and A. Hierlemann, “Liquid-phase chemical and biochemical detection using fully integrated magnetically actuated complementary metal oxide semiconductor resonant cantilever sensor systems,” Analytical Chemistry, vol. 79, no. 4, pp. 1646-1654, Feb. 2007.

[5] L.A. Beardslee, K.S. Demirci, Y. Luzinova, J.J. Su, B. Mizaikoff, S. Heinrich, F. Josse , O. Brand, “In-plane mode resonant cantilevers as liquid phase chemical sensors with ppb range limits of detection,” in Tech. Dig. Solid-State Sens.,Actuator Microsyst. Workshop, Hilton Head Island, SC, Jun. 2010, pp. 23-26.

[6] J. H. Seo and O. Brand, “High Q-Factor In-Plane-Mode Resonant Microsensor Platform for Gaseous/Liquid Environment,” JMEMS 2008, Vol. 17, issue 2, pp. 483-493.

[7] A. Rahafrooz, and S. Pourkamali, “Rotational mode disk resonators for high-Q operation in liquid,” proceedings, IEEE Sensors conference, 2010, pp. 1071-1074.

[8] A. Rahafrooz, and S. Pourkamali, “High-frequency thermally actuated electromechanical resonators with piezoresistive readout,” IEEE Trans. On Electron Device, 2011, Vol. 58, issue 4, pp. 1205-1214

Relationship between time interval measurement and frequency sources comparison and its application

Miao Miao, Zhou Wei
Department of Measurement and Instrument
Xidian University
Xi'an, P.R. China
mmiao@mail.xidian.edu.cn

Qu Bayi
Department of Information Engineering
Chang'an University
Xi'an, P.R. China

***Abstract*—Comparison of 1pps signals has been applied in the relevant fields, such as time transferring, time keeping, and communication. Time comparison can be carried out in many methods with different precision and the time interval universal instrument is used usually. However, 1pps signals are generated from frequency standard sources so there is the relationship of multi-periods synchronization between 1pps signals comparison and frequency sources comparison, which restricts the improvement of the resolution of time comparison.**

The paper analyzes the relationship between 1pps signals comparison and frequency sources comparison. The precision of frequency sources comparison is far better than the direct comparison of 1pps signals, reaching ps. The indirect scheme of time comparison is proposed using frequency sources comparison. Utilizing the high precision of frequency sources comparison, the scheme has get round many disturbance factors from acquiring 1pps signals. Then, by the way of computation to get the comparison results, the better precision comparison can be obtained. The scheme has been tested by lots of experiments.

The scheme has simpler construction and higher accuracy. The comparison of atomic clocks is similar with the comparison of 1pps signals and the scheme also applies to it. Furthermore in time transferring, due to the definite phase relations between the local time signal and the master clock in Astronomical Observatory, the user can also acquire high accurate time transferring utilizing indirect phase comparison. The comparison procedure should be done with determining the initial state and continuous measuring, or the error will be increased. The further research is required for the scheme because the experiment model has a bit idealization and simplicity.

I. INTRODUCTION

At present, time comparisons are usually accomplished through the comparison between 1pps signals with high-precision time internal measurement. 1pps signals are generated from standard frequency sources, so there is the relationship of multi-periods synchronization between 1pps signals comparison and frequency sources comparison. Especially, for high-accurate time keeping, many technologies, such as atom clocks group, backup clocks, phase fine-tuning and international regular comparison and so on, are applied[1][2].

With the continuous development of time transferring technologies, the requirement of time comparison become more complex and more precision. In recent years, high-resolution time transferring needs higher performance for the relation with high-resolution positioning. The accuracy of time interval measurement should be concerned, while the error is less than ±100ps [3~5]. Many high-precision time interval measurement instruments mainly adopt analog interpolation, time-voltage conversion and delay quantization to reduce measuring error, even reaching the order of 10ps [3] [4] [6]. Compared with the above the comparison between sources has lower error, such as the comparison in 10s has less than ps error [7] [8]. On the base of the circuit stability, the period signal processing greatly reduces trigger error to reach higher resolution [9] [10].The paper mainly discusses how to utilize the relation of source comparison and 1pps signal comparison in order to get a high-precision time comparison method. Experimental results agree with theory analysis.

II. ANALYSIS OF TWO KINDS OF COMPARISON

Figure.1 shows the relation between comparison of frequency sources and comparison of 1pps signals.

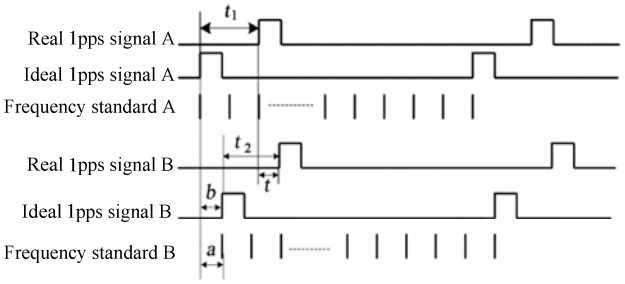


Figure1. Diagram of the relation between comparison of frequency sources and comparison of 1pps signals

In Figure.1, the 1pps signal is generated from counter's frequency dividing. a is the comparison result; b is the time difference of 1pps signals in ideal case; t_1 is the delay of

Project Supported by the National Natural Science Foundation of China (Grant No.60772135 and 10978017), the Fundamental Research Funds for the Central Universities China (Grant No.JY10000905015)

frequency standard A; t_2 is the delay of frequency standard B; t is the real time difference of 1pps signals. The comparison result of 1pps signals equals to the time internal of rising edges or falling edge.

Under ideal circumstance, without consideration of the delay of frequency divider, the comparison result just is the comparison of 1pps signals, shown as Eq.1.

$$a = b \tag{1}$$

Taking the delay of frequency divider into account, the edges of the frequency source and the 1pps signal have a difference, shown as Eq.2.

$$a - t = t_1 - t_2 \tag{2}$$

|t1 - t2| is the delay difference of two dividers. If |t1 - t2| is given, the comparison result can be obtained through analysis and processing of the phase comparison. Fig.1 shows a certain state, under which the states of two dividers are relevant. That is, they are electricity at the same time and generate a pulse after counting 10M to avoid the whole period error. Generally, the counters' procedure is closely related with its working state and starting time, while two counters are irrelevant. The difference of 1pps signals equals to phase difference adding the revision of integral units of periods. The revision value is a constant when the system is working continually.

III. EXPERIMENTS

A. Experiment Schemes

Through the experiment, the paper will show the close relation between the comparison of frequency sources and the comparison of 1pps signals. Under some certain condition, the former can be calculated from the later. As shown in Figure.2. 5585B is a high-accuracy Cs atom clock, with a 1pps signal output. SR620 is a high-accurate time interval measuring instrument. Experiment adopts EPM7032 to achieve 1pps signal based on frequency dividing counter.

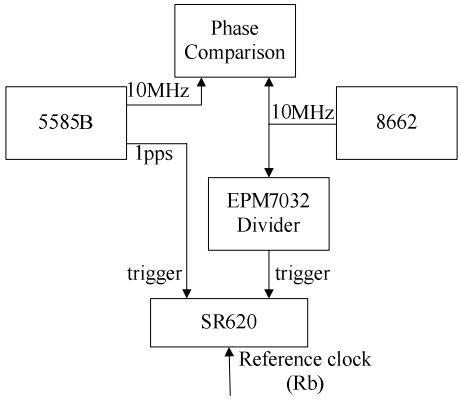


Figure2. Experiment system block

High-precision phase comparison instrument is used to measure the phase difference and record measuring data, while the time interval measuring instrument to compare 1pps signals and record data. Phase comparison method is utilized to measure the phase difference, with one data per 10 seconds. There are 116 data value in 150 seconds. Meanwhile, SR620 is used to measure the phase difference and record them, with one data per 10 seconds.

B. Experiment result analysis

Fig.3 is the curve of the output of the phase comparison. Experiment data is revised as follows: the data of the second period is subtracted by 100ns and the data of the third period is subtracted by 200ns, etc. 100ns is the whole period of 10MHz signals comparison.

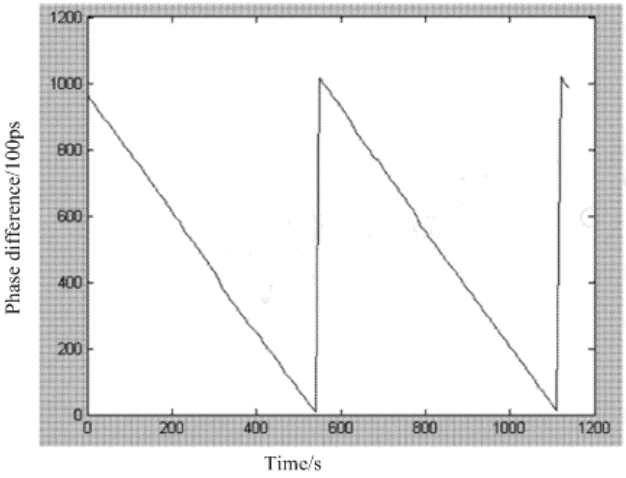


Figure3. Original result of output of phase comparison

For Figure.3, the phase difference in a period is reducing continuously, so the real data in the second period should be the measuring data subtracts the whole period, etc. Then the curve in Figure.3 becomes the middle line in Figure.4.

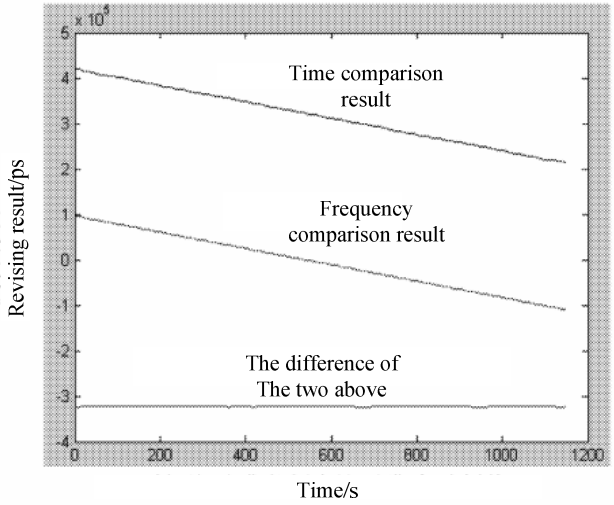


Figure4. Contrast diagram of time comparison and frequency comparison

Let w to be the difference of the phase difference and the 1pps difference. So Figure.4 is the curve of w. Combining with the definition of the 1pps signals difference, the delay difference of dividing circuit, w', can be derived from w.

Take the experiment for instance, w is revised for w'. The middle value of w' is 18.4ns, the relative change is about 8%. 18.4ns is the delay difference of dividing circuit, while 8% is

due to the instability of circuits. Experiments show that the delay of 1pps signals from EPM7032 is 34.77ns, while the delay from 5585B is 16.32ns. The difference is 18.45ns. Comparing with 18.4ns, the difference 0.05ns is caused by error. Figure.5 is the difference result of phase comparison result subtracting fitting line. Figure.6 is the difference result of time interval measured subtracting fitting line. The measuring result of time interval has poorer accuracy and stability. In the experiment, the phase comparison instrument used can reach the order of ps, simple construction, easy realization, high accuracy, comparing with SR620 [5].

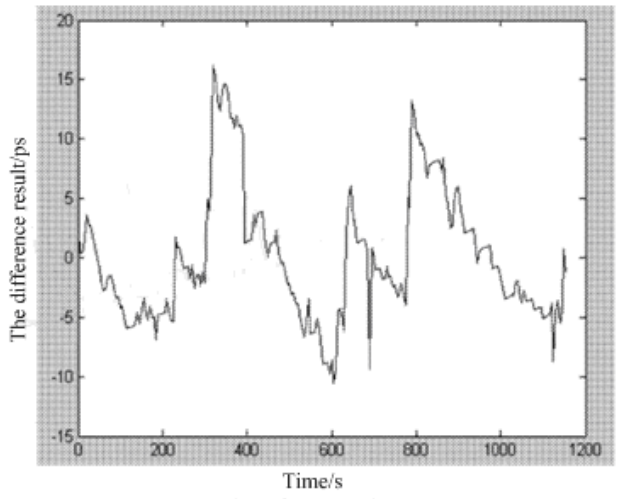


Figure5. Difference result of phase comparison result subtracting fitting line

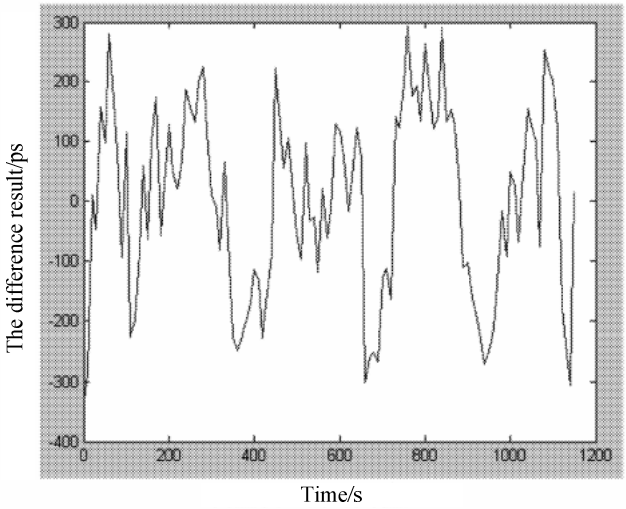


Figure 6. Difference result of time interval measured subtracting fitting line

IV. CHARACTERISTIC FEATURES

Based on the above analysis, combination with general-accurate time interval measuring instrument and high-precision phase comparison instrument can get a measuring result as a high-accurate time interval measuring instrument. If the delay difference of dividing circuit and the whole period difference, we also can use a high-accurate phase comparison. The method need to be developed furthermore because the experiment model is simpler and more ideal.

The atom clock comparison is similar as the above. As for time transferring, making use of the definite phase relation between the standard time signal and the main clock, high-precision can be obtained. The characteristic features are as follows:

A. Comparing with the traditional direct comparison method, the phase comparison is a indirect method. Its construct is simple and the accuracy is better.

B. The method can reflect the fine phase change between two 1pps signals, but the premise is determining high-precisely the original state and can measure continuously.

Breaking measurement may increase the whole period error and the initial state can be got by time interval measurement. The initial states are the whole period error and delays of dividing circuits.

V. CONCLUSION

Because 1pps signals are generated from frequency standard sources, the phase comparison of sources has a certain relation with the comparison of 1pps signals. Based on the relationship, the high-precision phase comparison can be attained through phase comparison. We can realize the phase comparison instrument with simple construction, relative cheap price. So, the combination of common time interval measurement and high-resolution phase comparison instrument can reach high-precision time comparison. Furthermore, if the delay of dividing circuits and the initial whole period error, etc, is given, a high-precision phase comparison instrument can even get the result of time comparison.

REFERENCES

[1] Z. M. Wang, J. F. Gao, " The Progress on Precise Time Transfer and Comparison between Remote Sites," Progress in Astronomy, 2000, 18(3), pp.181-190.

[2] L. Liu, C. H. Han, "Two Way Satellite Time Transfer and its Error Analysis," Progress in Astronomy, 2004, 22(3), pp.219-226.

[3] L. Li, W. Zhou, F. W. Wang, X. J. Ou, N. Ding, C. Z. Zou, et al, Frequency Control Symposium, 2007 Joint with the 21st European Frequency and Time Forum, IEEE International, pp.267-272.

[4] M. Miao, W. Zhou, "Comparison between Analog and Digital Time and Frequency Measurement Techniques," Frequency Control Symposium, 2007 Joint with the 21st European Frequency and Time Forum, IEEE International, pp.801-804.

[5] B. Y. Qu, L. L. WANG, Z. Q. Li, et al. A phase comparator with 10 picosecond resolution based on phase coincidence phenomena. Chinese Journal of Scientific Instrument, 2007, 28 (8), pp.250-252.

[6] B. Y. Qu, W. Zhou, M. Miao, D. JIN, "Time-to-digital Technology Based on Transmission Delay, Journal of astronautic metrology and measurement," 2009, 29 (4), pp. 45-49.

[7] Z. Huang, B. Liu. New Method to Measure the Time of Flight in Pulse Laser Ranging. Journal of Op to electronics, Laser, 2006, 17 (9), pp.1153-1155.

[8] W. Zhou, Z. Q. Li, et al. A Novel Frequency Measurement Method Suitable for a Large Frequency Ratio Condition. Chinese Physics Letters, 2004, 21 (5), pp.786-788.

[9] H. Zhou, W. Zhou, A Time and Frequency Measurement technique Based on Length Vernier. Proceedings of the 2006 IEEE International Frequency Control Symposium, 2006, pp.267-272.

[10] P. Chen, C. C. Chen, J. C. Zheng, et al. A PVT Insensitive Vernier Based Time to Digital Converter With Extended Input Range and High Accuracy. IEEE transactions on nuclear science, 2007, 54 (2), pp.294-301.

Improvement of the Asia-Pacific TWSTFT Network Performance Utilizing DPN results

H.T. Lin, Y.J. Huang, C.S. Liao, F.D. Chu, and W.H. Tseng

National Time and Frequency Standard Lab.
Telecommunication Laboratories (TL), Chunghwa Telecom
Taoyuan, Taiwan
linht@cht.com.tw

Wei-Chih Hsu

Department of Computer and Communication Engineering,
National Kaohsiung First University of Science and Technology,
Kaohsiung 811, Taiwan
weichih@nkfust.edu.tw

***Abstract*—Two-way satellite time and frequency transfer is a precise time transfer method and has being used to generate the TAI since 1999. Nowadays, TWSTFT links have formed a worldwide network and the utilization of the highly redundant TWSTFT data become an important topic. In our pervious study, a feasible method has been proposed, with full time transfer network data, to improve the results of TWSTFT network.**

Besides, NICT has recently developed the software-defined radio based two-way time transfer modems using the dual pseudo random noise (DPN) signal. The first DPN two-way time transfer experiments, using these modems, were performed between NICT(Japan) and TL(Taiwan) from March to December 2010 and its ability in improving the time transfer precision was approved on the Japan-Taiwan baseline. In compare with the conventional NICT-TL TWSTFT link, the DPN time transfer results have higher precision and less diurnals.

Since the DPN results show better performance than the conventional TWSTFT time transfer, we would adopt the DPN results for the NICT-TL link, and solve the TWSTFT+DPN network solution by using our proposed method. The numerical results about how the DPN data can benefit the performance of TWSTFT network will be illustrated in the paper.

I. INTRODUCTION

Two-way satellite time and frequency transfer (TWSTFT) [1] is a precise time transfer method and has developed into a widely used technique in those laboratories contribute to the generation of International Atomic Time (TAI) and realization of UTC [2], [3]. Rules of data taking, processing, and reporting, for TWSTFT direct-link, are laid down in the recommendation ITU-R TF.1153-2. However, the TWSTFT were usually measured as a network but its data were used as single base lines, so a lot of measured data may be wasted.

In recent years, the number of point-to-point two-way time transfer links is growing, so the concept of network time transfer has been proposed [4]–[6] and the utilization of the redundant data of each indirect link in the TWSTFT network becomes an important topic. Besides, combining different time transfer results techniques is also a solution for the high redundancy in the TAI network so as to fully use the total potential of different time link data. Because more and more new techniques are coming, e.g., Glonass, Galileo, Compass, etc, the question of how to combine different techniques is inevitable and we cannot evade this trend. For example, Jiang *et al.* [7] has proposed an idea, combining the advantages of both TW & PPP methods, to improve the time transfer results. By using the TWSTFT+PPP method, the diurnals in TWSTFT results can be reduced.

Moreover, National Institute of Information and Communications Technology (NICT) has recently developed the software-defined radio based two-way time transfer modems [8]-[10] and their ability in improving the time transfer precision by using the dual pseudo random noise (DPN) signal was approved on the Japan-Taiwan baseline [11]. The DPN two-way time transfer experiments, using these modems, were performed between NICT(Japan) and TL(Taiwan) from March to December in 2010. During this period, the Asia-Pacific rim region TWSTFT network performed continuously. Numerical analyses show that the DPN time transfer result of NICT-TL experiment has higher precision and less diurnals in compare with the conventional TWSTFT experiments.

In our previous study, a feasible method was proposed utilizing full data of the TWSTFT network for data processing [12]. The numerical results show that better short-term stabilities and uncertainties of most links in the network can be obtained. Therefore, we are curious about how the results of a single DPN link can benefit the TWSTFT time transfer network.

In this study, we would try to include the time transfer data of DPN experiment in solving the TWSTFT network solution

This work is support in part by the Bureau of Standards, Metrology & Inspection, Ministry of Economic Affairs, Taiwan under Grant number BSMI 100-1403-05-05-13.

with our proposed method. One will see that, for most links in the "TWSTFT+DPN" network, not only the short term stabilities but also the diurnals are effectively improved.

II. Theoretical Background and Method

A. Solving the Time-Transfer Network Solutions

First of all, we would briefly describe our previous proposed method for solving time-transfer network solutions. To calculate the time difference T_1-T_2 between two stations 1 and 2, the time interval readings at stations 1 and 2, i.e., TIC_{12} and TIC_{21}, can be applied. The related equation for TWSTFT is as the following

$$\begin{aligned} T_{12}&=(T_1-T_2)=N_{12},\\ \text{where } N_{12}&=\tfrac{1}{2}(TIC_{21}-TIC_{12})\\ &\quad+\tfrac{1}{2}[(TU_1-TD_1)-(TU_2-TD_2)]\\ &\quad+\tfrac{1}{2}(\tau_{12}-\tau_{21})\\ &\quad+\tfrac{1}{2}[(\tau_1^{TX}-\tau_1^{RX})-(\tau_2^{TX}-\tau_2^{RX})]+\Delta\tau_{R12} \end{aligned} \tag{1}$$

Equation (1) is the formula for calculating the time difference between each pair of TWSTFT stations. In equation (1), TU_i and TD_i denote satellite propagation delays of up-link from station i and down-link to station i. The third term $1/2(\tau_{12}-\tau_{21})$ is the non-reciprocity due the satellite equipment delays. The fourth term is the difference of the differential delays of the transmitting and receiving parts (equipment delays) of earth stations 1 and 2. The last term $\triangle\tau_{R12}$ is a correction of Sagnac effect.

For the case of 4-station TWSTFT network, we adopt the condition of zero closure sums and get the following linear equations

$$\begin{cases} T_{12}=N_{12}=Y_1\\ T_{13}=N_{13}=Y_2\\ T_{14}=N_{14}=Y_3\\ T_{23}=T_{13}-T_{12}=N_{23}=Y_4\\ T_{24}=T_{14}-T_{12}=N_{24}=Y_5\\ T_{34}=T_{14}-T_{13}=N_{34}=Y_6 \end{cases}, \tag{2}$$

with $N_{ij}=Y_k$ just for simplifying the subscript. Equation (2) can be written in matrix notation [13] as

$$X\,T=Y, \tag{3}$$

with $X=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\-1&1&0\\-1&0&1\\0&-1&1\end{bmatrix}$; $T=\begin{bmatrix}T_{12}\\T_{13}\\T_{14}\end{bmatrix}$;

$$Y=\begin{bmatrix}Y_1\\Y_2\\Y_3\\Y_4\\Y_5\\Y_6\end{bmatrix}=\begin{bmatrix}N_{12}\\N_{13}\\N_{14}\\N_{23}\\N_{24}\\N_{34}\end{bmatrix}$$

It is well known that the performances (e.g., stability) of every TWSTFT links are different. According to the TDEV value of each link Y_i at averaging times of 1 hour, we determine the weighting value as

$$w_i=\frac{1/\sigma_i^2}{\sum_{k=1}^{N}\left(1/\sigma_k^2\right)}, \tag{4}$$

where $\sigma_i = TDEV\ [Y_i]$ and $N = n(n-1)/2$.

Applying the weighting to (3), we get

$$WX\,T=WY, \tag{5}$$

where $W=\begin{bmatrix}w_1&0&0&0\\0&w_2&0&0\\0&0&\ddots&0\\0&0&0&w_{n(n-1)/2}\end{bmatrix}$.

Equation (8) can be solved by the weighted least squares method：

$$T=(X^T WX)^{-1}X^T WY. \tag{6}$$

B. Utilization of new time transfer data

To improve the two-way time transfer precision, NICT has developed a novel modem with a new timing signal. A pair of pseudo random noises (DPN) signal, based on the binary offset carrier (BOC) [1], separate in frequency domain is introduced. Since the experiment results of DPN show better short-term stability and less diurnal, it will be helpful in improving the network performance if we include the DPN results in the analysis of TWSTFT network solution.

In the following numerical analysis, we would not combine the numerical results of the TW and DPN time transfer links. We calculate the average value of the differences between TW link and DPN results, and shift DPN data by this value to align with the regular TWSTFT data. Then direct TW and shifted DPN results of NICT-TL link can be used separately in solving the network solutions of simple TWSTFT network and TWSTFT+DPN network (see case (a) and case (b) in Fig. 1.). By doing this, we will be able to distinguish the improvement effects caused by simple TWSTFT network analyses or including a DPN link in it.

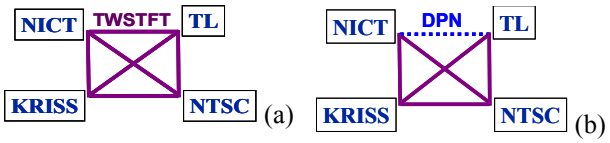


Fig. 1. The TWSTFT network connection among four stations. (a) simple TWSTFT network (b) TWSTFT+DPN network.

III. Numerical Results

The NICT-TL DPN experiment was performed from March 2010 to Dec. 2010. In the forepart of this period, some discontinuities exist in the measured results due to the power blackout, rain fade, and some other unexpected events. Since DPN is a novel technology to TWSTFT and this is the first international DPN link, many parameters (e.g. the transmitting power level) were not properly adjusted, and the performance of the current DPN results was not optimal. However, according to the observation and numerical analyses, we think DPN method is very potential and comparable to GPS PPP.

Let us consider the network constructed by links among four stations, including NICT, TL, the National Time Service Center of China (NTSC) and the Korea Research Institute of Standards and Science of Korea (KRISS). We select the results of the latter part (MJD 55515 to MJD 55532) of the DPN experiment for data analyses, because the DPN data continuity in this period is better. One can see that the TDEV values of all direct TWSTFT links (see the second row marked as "Direct" in Table-I) are small which means all the TWSTFT links have good stabilities.

The TDEV values of all direct TW link data, TWSTFT network and TWSTFT+DPN network solutions, are calculated and summarized in Table I. TWSTFT+DPN network mean that DPN results, instead of regular TWSTFT data, of NICT-TL link are used to solve the network solutions.

TABLE I
TDEV VALUES OF ASIA-PACIFIC TWSTFT DIRECT LINK, TWSTFT NETWORK RESULTS, AND TWSTFT+DPN NETWORK RESULTS. RELATED GAIN VALUES ARE ALSO SHOWN.

Link / TDEV (/1hr)(ps)	NICT-TL	NICT-NTSC	NICT-KRISS	TL-NTSC	TL-KRISS	NTSC-KRISS
Direct	104	157	179	140	84	125
TWSTFT Network	104	138	120	119	89.5	129
Gain(Net)	0%	12.1%	33%	15%	-6.5%	-3%
TWSTFT +DPN	47	122	94	116	88	129
Gain(DPN)	54.8%	22.3%	47.5%	17.1%	-4.8%	-3%

* TDEV value of NICT-TL DPN link is 48 ps in this case.

The Gain values of TWSTFT network and TWSTFT+DPN solutions at one-hour averaging time are also illustrated. One can see that the range of positive gains in TWSTFT network results is from 0% to 33%. Because all links are stable and their TDEV values are already small, the improvements upon their TDEV values through network analyses are limited, and two negative gains (-6.5% for TL-KRISS link and -3% for NTSC-KRISS link) are obtained. Besides, the Gain of network analyses for NICT-TL is zero.

For the TWSTFT+DPN networks, the Gain values for most links are obviously increased and the range is extended to from 17.1% to 54.8%. Actually, the NICT-TL DPN link already shows better stability (TDEV (/1hr) value is about 48 ps) than the regular NICT-TL link. The improvement caused by including the DPN results in network analyses is very obvious. Due to the equally stable performance of all links, we still have two negative Gains, but the values are reduced a bit.

Fig. 2 shows the TDEV comparisons among direct TW, TWSTFT network and TWSTFT+DPN network results of the NICT-TL link. The network solution of NICT-TL link is almost the same with the direct link (i.e. Gain = 0%). This is reasonable, as we have mentioned that all these links are stable, so limits the improvement to the results through TWSTFT network analyses.

After including the DPN results into the network analysis, the improvement of its short-term stabilities is obvious. Moreover, the diurnal effect is also reduced.

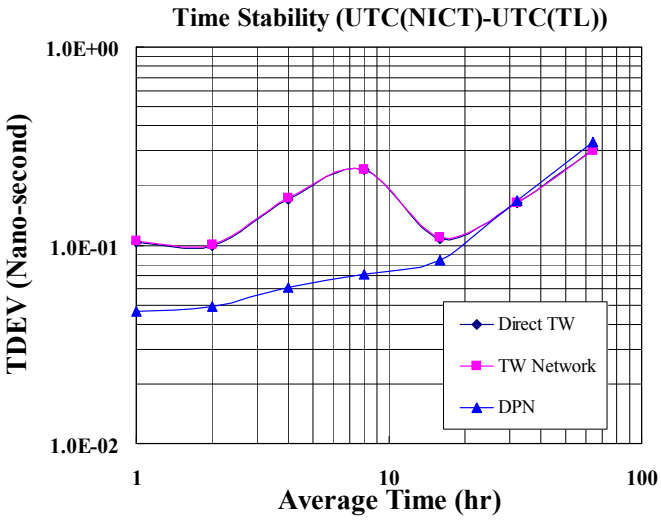


Fig. 2. TDEV values of the direct link data, and TWSTFT network and TWSTFT+DPN network solutions of NICT-TL link.

Now we check another time transfer link, say NICT-NTSC link, with obvious diurnal. It's a regular TWSTFT link and is included in our case study. The results of the direct link, TWSTFT network and TWSTFT+DPN network analyses for NICT-NTSC link are shown in Fig.3. The network result has better short-term stability than that of direct link and the Gain value is 12.1% (see the row "Gain(Net)" in Table I). By using the TWSTFT+DPN network analyses, the Gain value increases to 22.3% (see the row "TWSTFT+DPN" in Table I). In Fig. 3, we can also see that the diurnal effect is reduced. Therefore, including the DPN link for the TWSTFT+DPN network analyses will benefit to the reduction of diurnal.

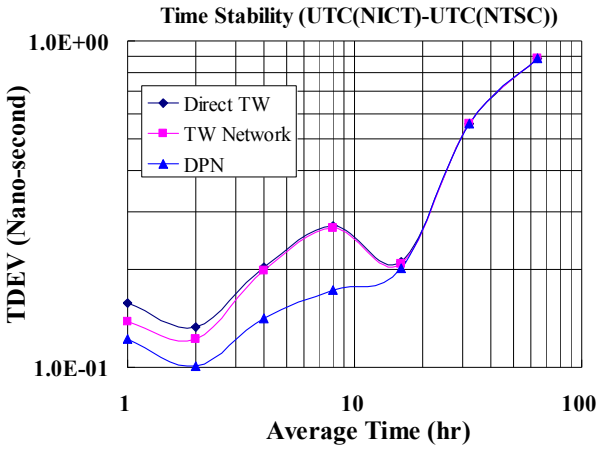


Fig. 3. TDEV values of the direct link data, and TWSTFT network and TWSTFT+DPN network solutions of NICT-NTSC link.

IV. CONCLUSION

The first international DPN experiment was performed between NICT and TL from March 2010 to Dec. 2010. Its results show better performance, with smaller TDEV value and less diurnal, than regular NICT-TL TWSTFT link. So we are curious about how the DPN results can benefit the TWSTFT time transfer network.

In this paper, our previous proposed method, with full time transfer network data by using weighted least squares method, is used to solve the TWSTFT network solutions. Because the TDEV values of these regular TWSTFT links are already small. We get two small "negative Gains" in six links, and the maximal Gain is up to 33%.

Replacing the direct NICT-TL link results by the DPN data, the TWSTFT+DPN network solutions can be solved with our previous proposed method. In compare with the TWSTFT network results, the positive Gains are further increased and the negative Gains are decreased a bit in TWSTFT+DPN network. This confirms that including the single DPN link is helpful in improving the stabilities of the links in TWSTFT network. We also find that the DPN link has less diurnal effect and can help to reduce the diurnals of other links in TWSTFT+DPN network.

ACKNOWLEDGMENT

Special thanks to Mr. Tadahiro Gotoh, Jun Amagai, Thomas Hobiger and related staffs of NICT for developing the GPU-based DPN modem. We would also like to express our gratitude to NICT colleagues for their devotion to the Asia-Pacific TWSTFT network and the arrangements for the NICT-TL DPN experiment.

REFERENCES

[1] D. Kirchner, "Two-Way Time Transfer via Communication Satellites," *Prodeedings of the IEEE, 79*, pp.983-989, 1991.

[2] E. F. Arias, "The metrology of time," *Philosophical Transaction of the Royal Society A*, 363, (1834), 2289-2305, 2005.

[3] E. F. Arias and G. Petit, "BIPM Comparison of Time Transfer Techniques," *in Proc. of 2005 joint IEEE IFCS and PTTI,* Vancouver, Canada, pp.244-246, August 2005.

[4] Z. Jiang and G. Petit, "Redundancy in the TAI TWSTFT time transfer network," in *Proc. 20th European Frequency and Time Forum*, 2006, pp. 468–475.

[5] Z. Jiang, "Towards a TWSTFT network time transfer," *Metrologia*, vol. 45, pp. S6–S11, 2008.

[6] Wen-Hung Tseng, Shinn-Yan Lin, Kai-Ming Feng, Miho Fujieda, and Hideo Maeno, "Improving TWSTFT Short-Term Stability by Network Time Transfer," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 57, pp. 161-167, 2010.

[7] Z. Jiang, G. Petit and P. Defraigne, "Combination of GPS PPP and Two-Way Time Transfer for TAI Computation," *in Proc. of 39th Annual Precise Time and Time Interval (PTTI) Meeting*, pp. 195–202, 2007.

[8] J. Amagai, "Two-Way Satellite Time and Frequency Transfer using a Pair of Pseudo Random Noises," ATF 2006

[9] J. Amagai and T. Gotoh, "Current Status of Two-Way Time and Frequency Transfer via Satellite," ATF 2008

[10] T. Gotoh, J. Amagai, "Two-Way Time Transfer with Dual Pseudo-Random Noise Code," in Proc. 40th Precise Time and Time Interval (PTTI) System and Applications Meeting, 2008, pp.459-466.

[11] T. Gotoh, J. Amagai, T. Hobiger, M. Fujieda, and M. Aida, "Development of a GPU-Based Two-Way Time Transfer Modem," in *IEEE Transactions on Instrumentation and Measurement*, Issue 99, in press, 2011.

[12] H. T. Lin, C. S. Liao, F. D. Chu, Y. J. Huang, and W. H. Tseng, "Full Utilization of TWSTT Network Data for the Short-Term Stability and Uncertainty Improvement," in *IEEE Transactions on Instrumentation and Measurement*, Issue 99, in press, 2011.

[13] S. Hongwei, and M. Imae, "Improvement of precision on two-way satellite time and frequency transfer", *Electrics Letters*, 27th August 2009, 45(18), pp.945-946.

An Optical Fibre-based Frequency Dissemination Network for Australia

Yabai He, Magnus T. L. Hsu, Malcolm B. Gray, Michael J. Wouters and R. Bruce Warrington
National Measurement Institute
Sydney, Australia
yabai.he@measurement.gov.au

Andre N. Luiten
University of Western Australia
Perth, Australia

Kenneth G. H. Baldwin and Daniel A. Shaddock
Australian National University
Canberra, Australia

Brian J. Orr
Macquarie University
Sydney, Australia

Guido Aben
AARNet
Perth, Australia

***Abstract*—An optical-fibre network is being planned for distribution of optical and microwave reference frequencies in Australia. This network would connect a number of research facilities and the National Measurement Institute (NMI), allowing long-range comparison of frequency standards to high precision and providing reference signals for a wide range of applications. First results for transfer in the laboratory achieve a fractional frequency stability of 7×10^{-17} at an averaging time of 10^4 s for the distribution of an 80 MHz radio frequency signal. We present details of the technique used to cancel noise introduced by the fibre, describe the all-digital radio-frequency phasemeter used to characterise performance, and outline plans for the further development of the network.**

I. INTRODUCTION

Optical fibre links have long been used to distribute radio frequency reference signals across large facilities [1], and more recently over longer distances using dedicated dark fibre [2]. With the advent of optical frequency standards [3], optical fibre links have emerged as one of the few methods capable of transferring the very high levels of accuracy and stability of these standards in practical averaging times [4]. In consequence, large-scale projects are already underway in Europe demonstrating optical frequency transfer up to 900 km [5, 6], and combining transfer of reference frequencies with other data traffic on commercial networks [7].

Australia is well placed both to contribute to and to benefit from the continuing development of optical fibre frequency transfer. Within Australia, fibre links can be tested across a transcontinental baseline. The National Measurement Institute (NMI) in Sydney on the east coast and the University of Western Australia (UWA) on the west coast are separated by over 3000 km. Both institutions operate atomic frequency standards, ultra-stable microwave oscillators and optical frequency combs. NMI and UWA already participate in GPS and two-way satellite time transfer, and are preparing for the Atomic Clock Ensemble in Space (ACES) mission [8] which will characterise new transfer techniques. UWA leads an Australian collaboration awarded significant funding by the Australian Government for ACES ground-station infrastructure, including a high-performance station clock.

Australia also has significant fibre network infrastructure. The Australian Academic Research Network, operated by AARNet, provides a number of options to connect laboratories within a proposed national network, including short (NMI to Macquarie University in Sydney, less than 10 km), intermediate (NMI to the Australian National University in Canberra, about 300 km) and very long (NMI to UWA) links. Fibre is currently being laid for trial sites within a new National Broadband Network, planned to connect over 90% of Australian premises.

Finally, Australia has major potential applications for timing and frequency signals delivered over fibre, particularly in the area of radio astronomy. For example, the Australia Telescope National Facility operates three radio observatories across New South Wales including the 64-m Parkes telescope. Very Long Baseline Interferometry is conducted using these and other radiotelescopes across the country. Australia is also one of two candidate hosts for the Square Kilometre Array (SKA), one of the largest research facilities ever devised [9]. Development of the Australian SKA Pathfinder (ASKAP) telescope is already in progress at the Murchison site in

978-1-61284-110-6/11/$26.00 ©2011 IEEE

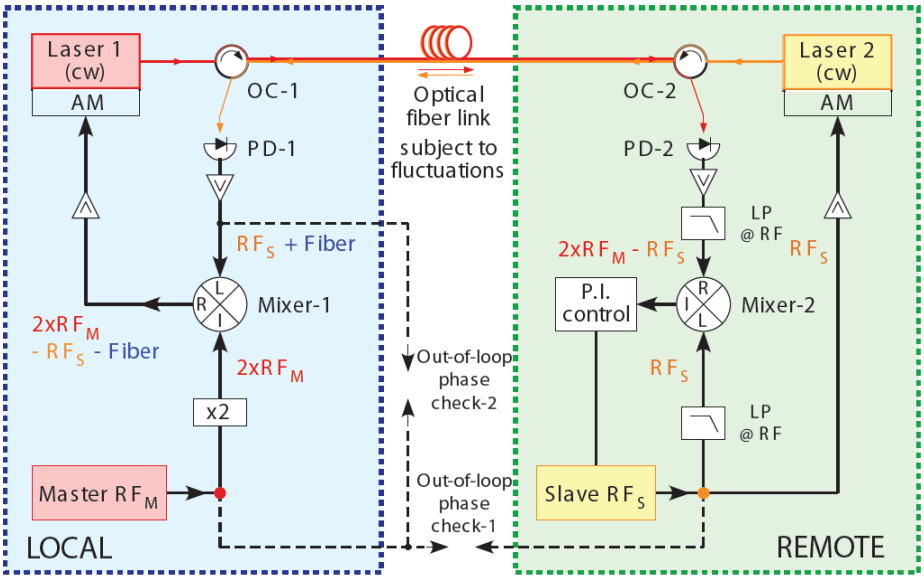


Figure 1. Simplified schematic of our fibre-based radio-frequency transfer system, with some components omitted for clarity. OC: optical circulator, AM: amplitude modulator, LP: low-pass filter, PD: photodetector, P.I. control: proportional-integral servo control, and ×2: frequency doubler.

Western Australia [10], and new fibre infrastructure to connect proposed SKA nodes is in preparation.

II. Apparatus

A. *Fibre-based radio-frequency distribution system*

As a first step towards planned long-range networks, we demonstrate transfer of radio-frequency signals over a fibre spool in the laboratory. Our system is illustrated in Fig. 1, with some components omitted for clarity to illustrate the basic design. The system uses commercial off-the-shelf components, and is designed to be relatively low-cost and scalable.

The local site has a stable master oscillator (in this case a hydrogen maser) which produces a reference radio frequency signal, RF_M, for dissemination to and stabilisation of a remote slave oscillator output, RF_S. The optical system consists of two RF-modulated lasers, one 'local' and the other 'remote', counter-propagating through the same optical fiber. Propagation through the fiber is subject to mechanical and thermal fluctuations, with the resulting phase noise suppressed as follows.

Commencing at the remote site, the remote signal RF_S is sent through the fibre acquiring the phase noise term *Fiber*. At the local site, the master oscillator output RF_M is frequency doubled and mixed down with the photodetector signal, $RF_S + Fiber$. The mixer output signal $2\,RF_M - RF_S - Fiber$ is then modulated onto the local laser and sent back through the fiber. The detected signal at the remote site is given by $2\,RF_M - RF_S$ only, since the additional *Fiber* noise term added on propagation cancels algebraically. The remote photodetector signal is then mixed with the slave oscillator RF_S to yield an error signal for feedback control.

As in all such fiber-noise cancellation schemes, cancellation requires that the fiber noise introduced is the same for the two counter-propagating beams. The practical performance of microwave components can also limit the effectiveness of cancellation; the design in Fig. 1 is chosen to improve isolation of unwanted sidebands, noting again that additional filters have been omitted for clarity. Finally, the bandwidth of the cancellation is limited by the fibre round-trip time, as in other schemes.

B. *All-digital radio-frequency phasemeter*

The design of the phasemeter has been presented previously [11]. As shown in Fig. 2, it uses front-end analog-to-digital conversion with dual phase-locked loops implemented digitally on a commercial field-programmable gate array (FPGA) platform. The digital implementation gives excellent timing coherence and signal isolation, and allows rapid prototyping and optimisation of control characteristics.

The noise floor of the phasemeter was measured by applying the same electronic signal to both reference and signal inputs in Fig. 2. The spectral density of phase fluctuations is well below 1 μrad/√Hz at frequencies above 0.1 Hz. The improvement over the earlier results [11] is due to improved synchronisation within the FPGA implementation, yielding greater rejection of common-mode noise.

III. Results

We measured the fractional frequency stability of our fiber-based radio frequency transfer system using the digital phasemeter. For this test, the remote oscillator was located in the same laboratory as the master oscillator. As shown in Fig. 1, the out-of-loop phase check-1 measures the stability of the master oscillator with respect to the slave oscillator.

Results for this measurement are shown in Fig. 3, together with the measured stability of typical reference oscillators for comparison. (The out-of-loop phase check-2 in Fig. 2 measures the fiber phase noise directly; however, this signal is

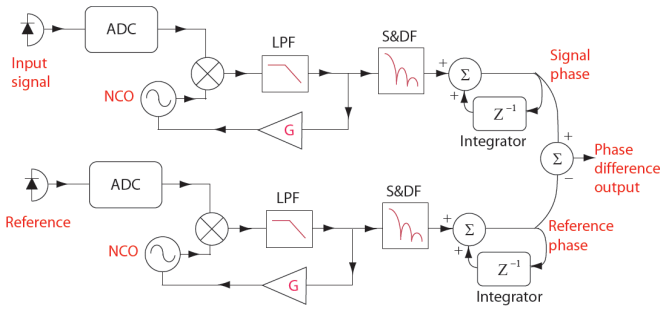


Figure 2. All-digital RF phasemeter [11], implemented on a commercial FPGA platform. NCO, numerically-controlled oscillator, LPF low-pass filter, and S&DF, sum-and-dump filter.

above the Nyquist limit for the phasemeter, and modifications to complete this measurement are currently under way.)

Operating at a carrier frequency of 80 MHz through a 20-km fiber spool, the transfer system achieves a fractional frequency stability of 7×10^{-17} (measured at 20 MHz) at an averaging time of 10^4 s. This is sufficient to transfer a reference signal from a hydrogen maser without degradation, as is evident from the direct comparison of two NMI hydrogen masers shown in Fig. 3. The noise floor of the transfer system was measured using a very short fiber link of 10 m, yielding a frequency stability of 4×10^{-17} at 10^4 s. We interpret these results as indicating that the system was operating close to the noise floor, for this length of spooled fibre and these averaging times.

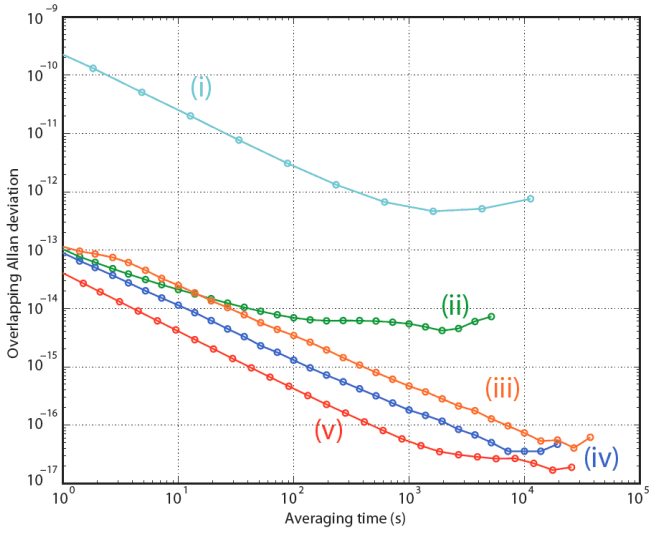


Figure 3. Fractional frequency stability results for (i) commercial rubidium and cesium oscillators, (ii) two hydrogen masers, fiber frequency distribution system through (iii) 20 km and (iv) 10 m of fiber, and (v) electronic noise floor of the phasemeter.

IV. Further Work

The initial work presented here is the first stage of a larger program, working towards a national network for transfer and dissemination of reference frequencies across Australia. The goals of the next stage of the program are to extend RF transfer to longer distances, in particular to connect existing Australian radio observatories, and to extend the transfer to optical frequencies.

At the time of writing first light has recently been transferred over the 10-km link between NMI and Macquarie University. A number of other fiber links have been identified for longer-range testing, with work already under way to establish access.

Acknowledgment

We acknowledge support from the Australian Research Council, and Macquarie University under the Linkage Seeding Program.

References

[1] S. M. Foreman, K. W. Holman, D. D. Hudson, D. J. Jones and J. Ye, "Remote transfer of ultrastable frequency references via fiber networks", Rev. Sci. Instrum. vol. 78, p. 021101, 2007.

[2] O. Lopez, A. Amy-Klein, C. Daussy, C. Chardonnet, F. Narbonneau and G. Santarelli, "86-km optical link with a resolution of 2×10^{-18} for rf frequency transfer", Eur. Phys. J. D vol. 48, pp. 35–41, 2008.

[3] H. S. Margolis, "Optical frequency standards and clocks", Contemporary Phys. vol. 51, pp. 37–58, 2009.

[4] D. Piester and H. Schnatz, "Hochgenaue Zeit- und Frequenzvergleiche über weite Strecken", PTB-Mitteilungen vol. 119, no. 2, pp. 131–143, 2009 (in German); "Novel techniques for remote time and frequency comparisons", PTB-Mitteilungen vol. 119, no. 2, pp. 33–44, 2009 (in English).

[5] H. Schnatz, O. Terra, K. Predehl, T. Feldmann, T. Legero, B. Lipphardt, U. Sterr, G. Grosche, R. Holzwarth, T. W. Hansch, T. Udem, Z. H. Lu, L. J. Wang, W. Ertmer, J. Friebe, A. Pape, E.-M. Rasel, M. Riedmann and T. Wubbena, "Phase-coherent frequency comparison of optical clocks using a telecommunication fiber link", IEEE Trans. Ult. Ferr. Freq. Cont. vol. 57, pp. 175–181, 2010.

[6] O. Terra, G. Grosche and H. Schnatz, "Brillouin amplification in phase coherent transfer of optical frequencies over 480 km fiber", Opt. Expr. vol. 18, pp. 16102–16111, 2010.

[7] O. Lopez, A. Haboucha, F. Kéfélian, H. Jiang, B. Chanteau, V. Roncin, C. Chardonnet, A. Amy-Klein and G. Santarelli "Cascaded multiplexed optical link on a telecommunication network for frequency dissemination", Opt. Expr. vol. 18, pp. 16849–16857, 2010.

[8] L. Cacciapuoti and Ch. Salomon, "Space clocks and fundamental tests: the ACES experiment", Eur. Phys. J. Special Topics vol. 172, pp. 57–68, 2009.

[9] The Square Kilometre Array, www.skatelescope.org (accessed 18 May 2011).

[10] Australia and New Zealand SKA project, www.ska.gov.au (accessed 18 May 2011).

[11] M. T. L. Hsu, I. C. M. Littler, D. A. Shaddock, J. Herrmann, R. B. Warrington and M. B. Gray, "Subpicometer length measurement using heterodyne laser interferometry and all-digital rf phase meters", Opt. Lett. vol. 35, pp. 4202–4204, 2010.

Improvement of the Reception Capability of Time Dissemination of Shortwave Broadcasting by Digital Modulation

S. H. Yang, Y. K. Lee, T. Y. Kwon, C. B. Lee
Center for Time & Frequency, Division of Physical Metrology
KRISS
Daejeon, South Korea
shyang@kriss.re.kr

S. J. Lee
Department of Electronics Engineering
Chungnam National University
Daejeon, South Korea
eesjl@cnu.ac.kr

Abstract— **KRISS (Korea Research Institute of Standards and Science) has been operating the standard time and frequency broadcasting station (call sign: HLA) since 1984. The station HLA broadcasts continuous time and frequency signals on 5 MHz carrier frequency and 2 kW transmit power. But the poor reception due to radio interference is decreasing users. Therefore to increase the radio reception, a digital modulation method and an increase of bit rate are introduced additionally to the present amplitude-modulated radio signal. We describe a digital modulation method for improving radio reception of the standard time and frequency by shortwave.**

I. INTRODUCTION

The radio reception of a shortwave broadcasting is low due to a signal fading easily and is susceptible to environment, while a shortwave broadcasting reaches to long distance owing to ionosphere reflection (a phenomenon known as skywave propagation). Shortwave radio is used for broadcasting of voice and music, and long-distance communication to ships and aircraft, or to remote areas out of reach of wired communication or other radio services. Amateur radio on these frequencies can provide hobby, educational and emergency communication [1].

But the reception is becoming inadequate by a bad environment such as complicated radio propagation, many tall buildings and so on. Thus it is necessary to develop the modulation method in order to enhance the radio reception against these conditions. KRISS has broadcasted the Korea Standard Time on 5 MHz carrier frequency with IRIG-H time code format [2] by amplitude modulation (AM). We introduce a digital modulation method to enhance the radio reception of the standard time and frequency broadcasting at KRISS, and carry out simulations to find out an optimal digital modulation method [3]. And we design a modulation method combined BPSK-DSSS (Direct Sequence Spread Spectrum) with AM for the existing AM users.

Goal of this study is to enhance the radio reception by application of a digital modulation method. And the Scope is as follows:

- Design an optimal digital modulation method for improving the radio reception by analysis of the conventional modulation method and suitability evaluation through the analysis of various digital modulation methods
- Verify the radio reception enhanced by digital modulation with a simulator by selection of the indexes for the comparison of reception performance and comparison with the conventional method
- Preparation of a digital modulation system and its performance test

II. ALTERATION OF BROADCSTING STATION AT KRISS

A. Current Station

The specifications of current station are as follows:

- Carrier frequency: 5 MHz
- Modulation: AM
- Time code format: IRIG-H
- Band width: 6 kHz
- Transmit power: 2 kW
- Antenna type: conical

The drawbacks in current system are vulnerable to fading and environmental noise, besides low power and outdated equipment such as vacuum tube amplifier. Therefore we have tried to replace with solid-state device (SSD) amplifier and will change the modulation method.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

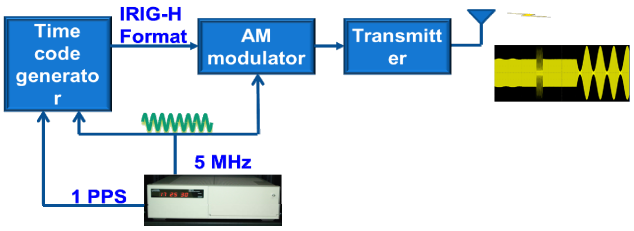


Figure 1. Block diagram of current system for the broadcasting timing signal.

B. *Modification of the Station*

For the modification of the station, as shown in Fig. 2, we adopt the BPSK digital modulation, and increase the bit rate by IRIG-B time code format and the transmit power up to 5 kW, and extend the bandwidth to 10 kHz. At preset we have equipped new timing system such as SSD power amps, have been making the time code generators and digital modulators combined analog modulation and so on.

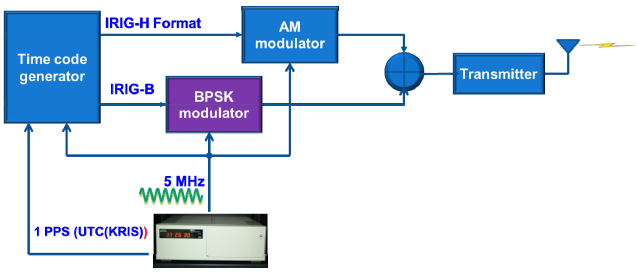


Figure 2. Block diagram of transmission system combined AM and BPSK-DSSS.

III. VERIFICATION OF TX AND RX BY SIMULATION

We carried out the simulation so as to select an optimal modulation method as shown Fig. 3. In order to analyze reception performance corresponding to the addition of the digital modulation scheme, we designed a simulation program. The program is divided into two parts as the transmitting part and the receiving part. In the transmitting part, the transmitting signals are generated by combining the AM modulated with the IRIG-H time code and the BPSK-DSSS modulated with the IRIG-B time code signals. In the receiving part, the demodulations are performed for the AM and BPSK-DSSS signals in which the additive white noise is added generated from the channel. Fig. 4 and 5 show the results of the monitored spectrums in the transmitting and the receiving parts, respectively.

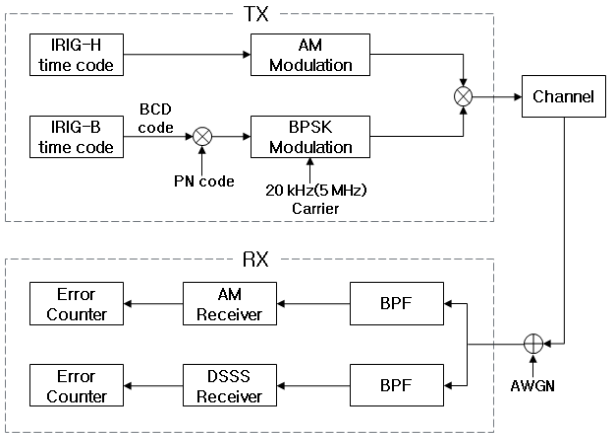


Figure 3. Simulation design of transmit and reception with DSSS

In the case of employing the bit error rate (BER) as the reference of the receiving performance, tremendous amount of symbol data are required in order to get reliable results. The carrier frequency of the standard time and frequency broadcast signals is 5 MHz, however, a tremendous amount of time is required to generate the large symbol data with the 5 MHz carrier frequency. Thus, we set the carrier frequency as 20 kHz instead of 5 MHz and sampling frequency as 80 kHz for efficient tests the in the simulations.

There are many referencing methods for comparing the performances of various digital modulation schemes. Among them, the BER and the bandwidth efficiency are generally used. Therefore, we use them for comparing the theoretical performances of the modulation scheme [4].

Fig. 6 shows the demodulated result when the AM modulated and the BPSK-DSSS modulated signals are transmitted together. From the results, it is observed that the required E_b/N_0 of the AM modulated signals is about 31 dB and that of the BPSK-DSSS modulated signals is about 7.5 dB when the BER is $1x10^{-2}$. This means that the required power of the BPSK modulation scheme is lower than that of the AM modulation scheme to get the same BER, and in other words the BER performance of the former is better than that of the latter under the same E_b/N_0 environment.

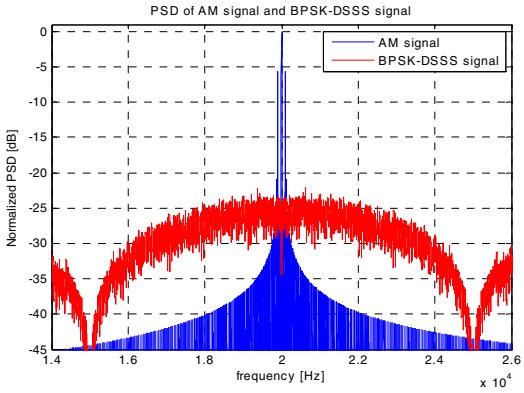


Figure 4. Power spectrums of AM and BPSK- DSSS at transmitting part in figure 3.

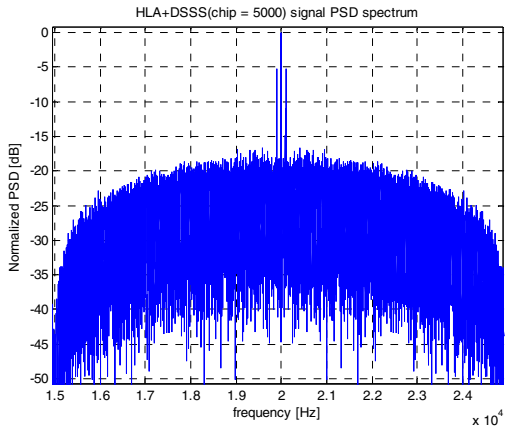


Figure 5. Simulation of the combinaton of AM and BPSK- DSSS at receiver input in figure 3.

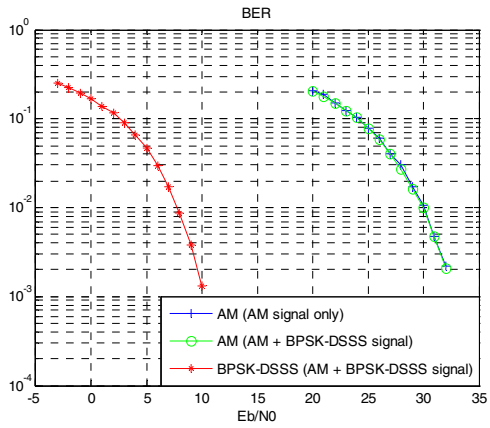


Figure 6. BERs of AM and BPSK-DSSS at each user receiver by simulation.

IV. CONCLUSIONS

We introduced a digital modulation method to enhance the radio reception of the standard time and frequency broadcasting at KRISS. And we carried out simulations to find out an optimal digital modulation method. And we design a modulation method combined BPSK-DSSS with AM for the existing AM users, and the IRIG-B and the IRIB-H time code format will be applied simultaneously to digital modulation and amplitude modulation respectively.

A modulator combining digital and analog signal will turns out in the near future. And we will fabricate a receiver for digitally modulated signal and verify the improvement by the comparison of the radio reception between an analog and a digital receiver at the field.

REFERENCES

[1] http://en.wikipedia.org/wiki/shortwave

[2] Telecommunications and Timing Group, IRIG Standard 200-04: IRIG Serial Time Code Format, Range Commanders Council, 2004.

[3] Ziemer, R, and W.Tranter, Principles of Commumications: Systems, Modulation, and Noise, 4th ed. Wiley, New York, 2002

[4] James E. Gilly, "Bit-Error-Rate Simulation Using Matlab," Transctypt International Inc., 2003

Impact of the transponder configuration on the Asia-Europe TWSTFT network

M. Fujieda*, H. Maeno*, D. Piester†, A. Bauch†, S. H. Yang‡, T. Suzuyama§, W. H. Tseng¶, L. Huanxin‖, Y. Gao**, J. Achkar††, and D. Rovera††
*National Institute of Information and Communications Technology, Tokyo, Japan
†Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
‡Korea Research Institute of Standards and Science, Daejeong, Korea
§National Institute of Advanced Industrial Science and Technology, Ibaraki, Japan
¶Telecommunication Laboratories, Taoyuan, Taiwan
‖National Time Service Center, Shaanxi, China
**National Institute of Metrology, Beijing, China
††Observatoire de Paris, Paris, France

***Abstract*— Two-way satellite time and frequency transfer between Asia and Europe has been performed for many years first by a two-transponder and later by a one-transponder configuration on the IS-4 and the AM-2 satellites, respectively. A diurnal phase instability was observed in links established through IS-4, while such instability has not been observed when AM-2 was used. It could be shown that the diurnal variation are site-independent. An on-board path asymmetry due to the use of two transponders seems thus to cause it.**

I. Introduction

Today two-way satellite time and frequency transfer (TWSTFT) is one of the essential tools for time transfer made in context with International Atomic Time (TAI) calculation [1], [2]. In TWSTFT, paired earth stations transmit and receive pseudo-random-noise coded signals via a communication satellite to exchange timing information [3]. Because the transmission and reception paths are almost the same, the atmospheric and geometrical delays between earth station and satellite are generally canceled. Today's operational TWSTFT consequently enables precise time and frequency transfer down to the 10^{-15} uncertainty level for frequency comparisons and 1 ns for time transfers [4].

The first TWSTFT experiment connecting Europe (Eu) and Asia was between Telecommunication Laboratories (TL) and NMi Van Swinden Laboratorium B.V. (VSL) [5]. Subsequently a link between National Institute of Information and Communications Technology (NICT) and Physikalisch-Technische Bundesanstalt (PTB) was established using the geostationary telecommunication satellite Intelsat-4 (IS-4) at 72°E in July of 2005 [6]. The Korea Research Institute of Standards and Science (KRISS) joined the link in November of 2005 and the National Metrology Institute of Japan (NMIJ) joined in September of 2007. The number of the participants sharply increased after that: TL, National Time Service Center (NTSC), and National Institute of Metrology (NIM), the latter two from P.R. China, joined in March of 2008, May of 2008 and December of 2008, respectively. In March of 2009, Observatoire de Paris (OP) came in as the second station in Europe. The Asia-Eu link using IS-4 became one of the most active TWSTFT networks. Additionally the direct comparison of Cs atomic fountains was performed between NICT and PTB and an uncertainty of frequency transfer of 7×10^{-16} was achieved during a 15-days campaign [7].

In the intercontinental TWSTFT link, two transponders in the Ku band are typically used to cover a wide area. Diurnal phase variations were observed in the links with such configuration [8]. So far, the effects due to the atmosphere, instrumental delays and satellite motion have been studied [9], [10], [11]. The Asia-Eu TWSTFT network using IS-4 stopped working because of a satellite malfunction on February 1, 2010. After the switch to the satellite AM-2 at 80°E, the link was resumed on October 1, 2010. Since the AM-2 satellite has one wide beam, the link is now established by one transponder. The clear diurnals have no longer been observed in the new link. In this report, the link performances before and after the satellite switch are shown, with particular view on the different transponder configurations.

II. The Asia-Eu TWSTFT link

The Asia-Eu link was established using IS-4 and AM-2 from November 2005 to February 2010 and from October 2010 to the present, respectively. Table I shows the locations and antenna diameters of the participating stations. The used frequency bandwidth was 2.5 MHz in the Asia-Eu link. To transmit a 2.5-Mcp/s modulated signal, SAW filters whose 3-dB bandwidth was 2.3 MHz were applied in the transmission path to limit the bandwidth occupied by the signal. In the earth stations the same equipment was used in the IS-4 as well as later in the AM-2 link. Two Ku-band transponders were used to establish the link via IS-4. In other words, different uplink and downlink frequencies were assigned for Eu and Asia stations. Time transfer was performed once every hour. After the satellite switch to AM-2, the link with OP stopped because OP is out of the coverage of the AM-2 footprint. Time transfer among PTB, TL, NTSC and NIM is performed once every hour with a duration of 5-minutes per session between

978-1-61284-110-6/11/$26.00 ©2011 IEEE

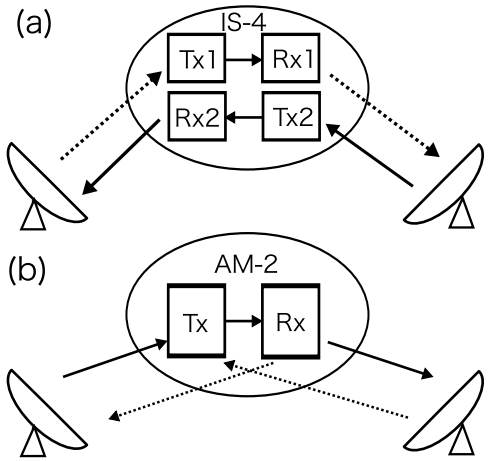


Fig. 1. Transponder configurations:(a) IS-4 link, (b) AM-2 link.

13 h UTC and 22 h UTC. The operation time is due to the limited working time of the satellite transponder. Only one Ku-band transponder is used to establish the link via AM-2 and thus the uplink and downlink frequencies are common for all stations. Fig. 1 shows the transponder configurations in the cases of IS-4 and AM-2. When two transponders are used, the signal paths involved in the on-board frequency conversion are different for the signals transmitted from Asia and Eu and vice versa, and a path asymmetry is induced accordingly.

TABLE I
LOCATION AND ANTENNA DIAMETER.

Abbr.	Location (LA, LO)	Antenna diameter[m]
NICT	(N 35.71, E 139.49)	2.4
KRISS	(N 36.39, E 127.37)	2.4
NMIJ	(N 36.07, E 140.13)	2.4
TL	(N 24.95, E 121.16)	2.4
NTSC	(N 34.37, E 109.22)	2.4
NIM	(N 39.96, E 116.41)	1.8
PTB	(N 52.30, E 10.46)	2.4
OP	(N 48.84, E 2.33)	2.4

III. LINK PERFORMANCE

A. via IS-4

Fig. 2 shows the frequency instabilities in the IS-4 links expressed by the Modified Allan Deviation (MDEV) between H-masers at PTB, OP and the reference signals in the Asian laboratories. The data period is from MJD 55120 to MJD 55165. The reference signals of the Asian laboratories except NIM were local time scales generated from H-masers. A commercial Cs clock served as the reference clock for the NIM station. The instabilities seen in links to NIM were governed by the performance of the Cs clock. Other long-term instabilities can be attributed to the frequency adjustments of the local time scales. The short-term instabilities of PTB-NICT, PTB-TL and OP-NICT links reached the mid 10^{-14} level at an averaging time of 1 hour. Since small phase jumps were seen in the OP-TL link, the short-term instability became worse. In the links with NTSC, the deviation of the received data at PTB and OP were larger than those received from TL, NIM and NICT. It is thought that the SAW filter used in the Tx path of NTSC station caused it. Almost all frequency instabilities show a significant bump around an averaging time of half a day, indicating an instability with a diurnal pattern. The reason is discussed in the next section. The frequency stabilities of the PTB-TL, PTB-NICT and OP-NICT links reach the low 10^{-15} level at an averaging time of 1 day, nevertheless. It allowed us to perform meaningful frequency comparison between Cs atomic fountains [4].

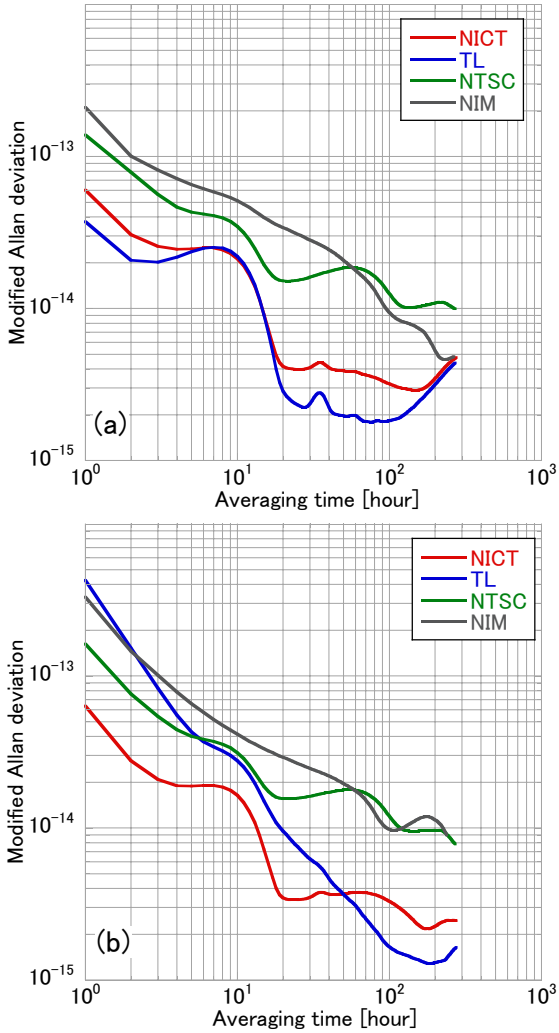


Fig. 2. Frequency instabilities with H-masers at PTB (a) and OP (b).

B. via AM-2

Fig. 3 shows frequency instabilities in the AM-2 links with PTB (a) and among Asian stations (b). The data period was from MJD 55617 to MJD 55630. The reference signals were local time scales generated from H-masers at PTB, TL and NTSC stations. Now the signal from a H-maser was also used

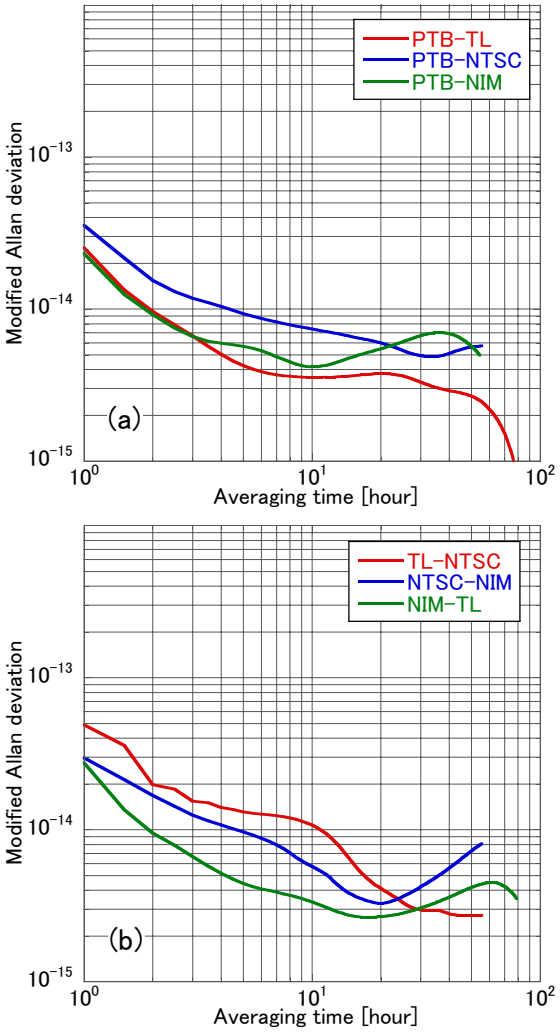


Fig. 3. Frequency instabilities in the AM-2 link with PTB (a) and among Asian stations (b).

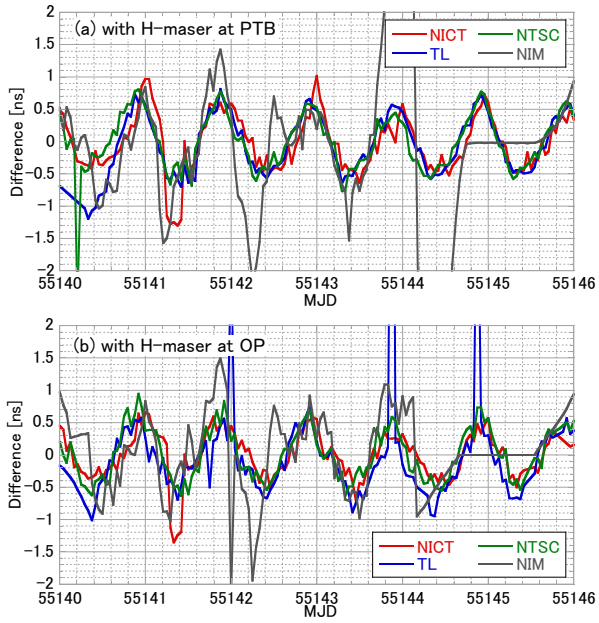


Fig. 4. Diurnal instabilities with H-masers at PTB (a) and OP (b).

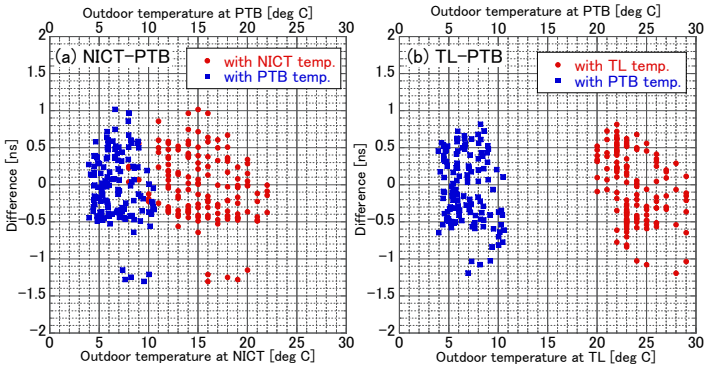


Fig. 5. Scattering plots between diurnal instabilities and outdoor temperatures.

at the NIM station. A bump indicating a diurnal instability is not seen in Fig. 3. The short-term instabilities are improved in comparison to those in the IS-4 link. The stabilities with NTSC look still worse because of the effect of the SAW filter, however. From the results shown in Fig. 3 (a) and (b), the frequency instabilities were similar in the links with PTB and among Asian stations. Although the elevation angle at PTB was as low as 3.7°, a detrimental effect has not appeared so far. The link performance was limited at first due to the limited (regularly interrupted) working time of the satellite transponder on AM-2. The regular interruption between the working times was linearly interpolated and MDEV was calculated. The frequency instability at an averaging time of one day was comparable to that of the IS-4 link. However, in this data procedure a bias is inserted to the MDEV, resulting that it may look better than it actually is.

IV. Diurnal Variation

A. via IS-4

1) Two transponders: To cover a wide area, multiple transponders are typically used in a communication satellite. Such configuration is shared in the intercontinental TWSTFT networks and causes an asymmetry in the signal paths between paired earth stations. It seems that this asymmetry causes an instability with diurnal pattern. The Asia-Europe TWSTFT network was established using two transponders on IS-4 as shown in Fig. 1 (a). Different uplink and downlink frequencies were assigned for Asia and Europe. From the beginning of the NICT-PTB link diurnal variations have been observed. The possible instability sources, as listed in Table II, have been discussed previously [10], [11]. The delays due to ionosphere, troposphere and the Sagnac effects were evaluated. Since their amplitude and phases were not synchronized with the observed instabilities, it was concluded that they were not the main sources of the instabilities. Fig. 4 shows the clock comparisons with H-masers at PTB (a) and OP (b) which are presented

in the form of the difference between raw data and moving-average data for 1 day to emphasize the diurnal pattern. Unfortunately some outliers are seen in the data with NIM and TL. It is apparent that the amplitudes, periods and their peak locations are almost the same and they look independent of the site. It is know that the instability due to the instrumental delay is highly related to the local temperature variation [13], [14], [15]. We thus computed the correlation factors between the outdoor temperatures and diurnal variations plotted in Fig. 4 (a). As shown in Fig. 5, the correlation was not larger than 0.2, and no systematic relation is visible. Consequently the diurnal patterns may be caused by a site-independent source, that is, the path asymmetry due to the transponder configuration, rather than being related with the instrument delays in each station.

Geostationary satellites orbit with a period of 23 hours 56 minutes 4 seconds together with the earth with respect to the stars. On the other hand, it takes 24 hours to return to the same position relative to the sun. If the diurnal instability would correspond to the former period, the phase of the variations would be shifted by 180° after half a year. The diurnal instabilities from MJD 54951 to MJD 54957, and from MJD 55131 to MJD 55137 on the TL-PTB link are shown in Fig. 6. The phase of the instability did not change, only the amplitudes slightly changed. One can conclude from this result that the period is likely determined from the relation between the sun and the satellite.

The NICT-PTB link was operational using IS-4 for about four years. The diurnal variations have been observed since the beginning of the link. We computed MDEV for each month and checked the amplitudes and found no seasonal variation of the diurnal peak. The mean value of MDEV was 2.0×10^{-14} at an averaging time of half a day, which corresponds to the diurnal amplitude in phase of 860 ps.

So far it was reported in [8] that a time step was observed in a transatlantic TWSTFT link with two-transponder configuration and it is thought that the step was induced by the delay variation on the transponder. Fig. 4 may be an example that the path asymmetry due to the transponder configuration induced the diurnal instabilities.

TABLE II
POSSIBLE INSTABILITY SOURCES.

Instability source	Amplitude [ps]	Site dependency
Instrument delay	~ 500	Yes
Ionosphere	~ 150	Yes
Troposphere	< 4	Yes
Sagnac effect	$50 \sim 650$	Yes
Transponder instability	~ 1000*	No

*This work: estimated to be slightly below 1 ns for Asia-Eu link via IS-4.

2) One transponder: To compare the link performances in the cases of two transponders and one transponder, respectively, inner-Asian time transfer using only one transponder on IS-4 was performed between NICT and KRISS from MJD 53975 and MJD 53982. Both laboratories were connected by another TWSTFT link using JCSAT-1B satellite at that time, which consisted also of a one-transponder configuration like Fig. 1 (b). Fig. 7 (a) shows the time-transfer results by the IS-4 and JCSAT-1B links established using one transponder. Both measurements were performed synchronously and the data were taken every second. The results were averaged for 5 minutes and convenient time offsets were inserted for better visibility in Fig. 7. The green line shows the double difference between them. Its standard deviation around the mean is 135 ps, which proves that the results obtained in both links agree well. Fig. 7 (b) shows their frequency stabilities. The short-term instabilities were lower than those in Fig. 2. Especially any bump cannot be seen in the double difference plot. If a site-dependent instability would exist, it is almost canceled by the subtraction of the results by the different links because of the synchronous measurement. On the other hand, since a site-independent instability is not canceled by the subtraction, the result indicates that such instability seems non-existent or at least negligible. The instability of 3×10^{-15} at half a day corresponds to a diurnal amplitude of 130 ps. This amplitude is much lower than that of about 900 ps shown in the links established by the two-transponder configuration. Thus the link performances showed clear difference between the one- and two-transponder configurations.

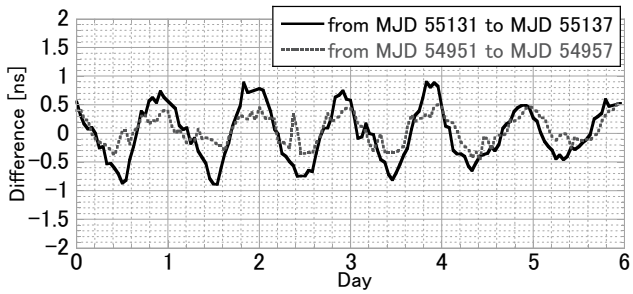


Fig. 6. Diurnal instabilities in the TL-PTB link.

B. via AM-2

The 1-hour averaged time differences and their frequency stabilities between TL and PTB by TWSTFT link via AM-2, GPS PPP and GPS P3AV are shown in Fig. 8. The results by the three methods look consistent. A bump around an averaging time of 40 hours exists in the TWSTFT instability whose reason is unresolved yet. The long-term agreement between TWSTFT and GPS data should be monitored carefully. Since any diurnal instability cannot be seen in GPS PPP and P3AV, the result by the AM-2 link was compared with them. Fig. 9 shows the double differences. Any clear functional variation was not observed. The instability obtained by TWSTFT at the averaging times longer than 2 hours reached the mid-10^{-15} level. It seems that such large diurnal instability as previously observed in the IS-4 link with the two-transponder configuration does not exist.

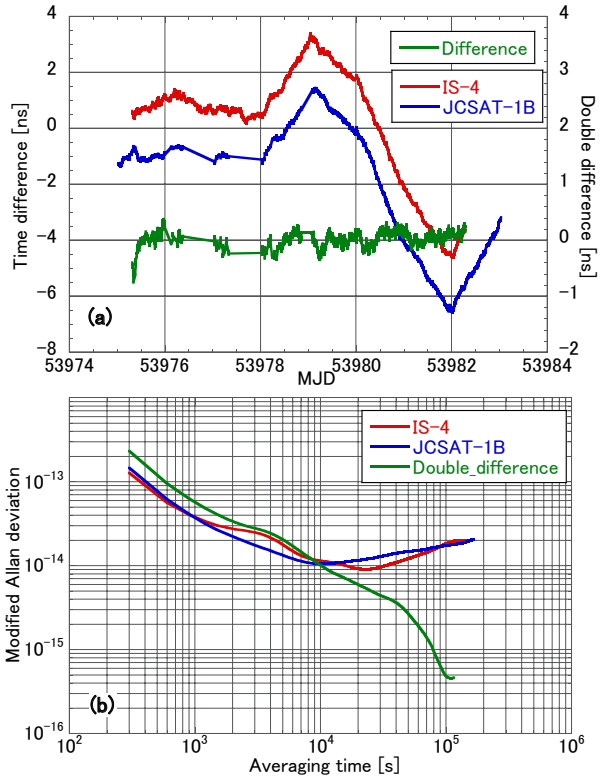


Fig. 7. (a) Time transfer results of NICT-KRISS by the IS-4 and JCSAT-1B links and (b) their frequency stabilities.

V. Conclusion

The Asia-Eu link was established using two transponders of IS-4 and restarted using one transponder of AM-2 after IS-4's malfunction. The diurnal phase variations have been observed since the beginning of the IS-4 link and possible error sources have been discussed so far. Thanks to the participation of the second European station, OP, it became clear that the diurnal instabilities looked site-independent. From the comparison of the one- and two-transponder configurations, such instability seemed existent only in the case when two transponders were involved. The current Asia-Eu link continues by one-transponder configuration on AM-2. Additionally any diurnal variation has not been seen there. It is thought that the on-board path asymmetry caused by two transponders induced such site-independent diurnal variation.

Acknowledgment

The authors would like to thank M. Rost, J. Becker and E. Staliuniene from PTB, and M. Aida from NICT, for their support to keep the regular measurements going.

References

[1] G. Panfilo and E. F. Arias, "Algorithms for International Atomic Time", *IEEE Trans. Ultrason. Ferrolectr. Freq. Control*, vol. 57, no. 1, pp. 140-150, 2010.

[2] D. Piester et al., "Time Transfer with nanosecond accuracy for the realization of International Atomic Time", *Metrologia*, vol. 45, pp. 185-198, 2008.

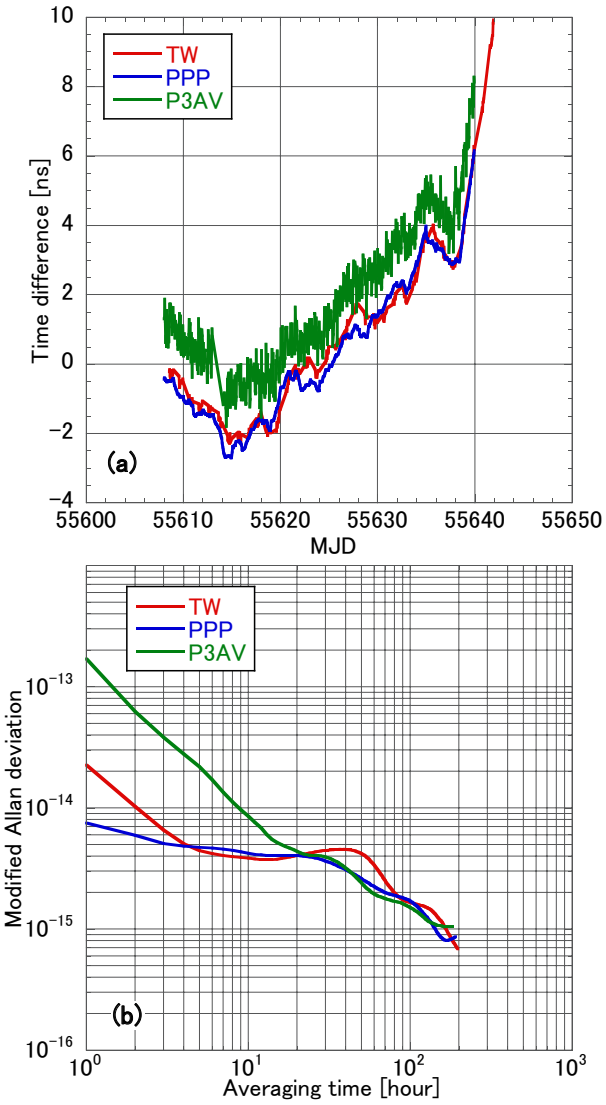


Fig. 8. Time differences of TL-PTB link by the AM-2 link, GPS PPP and GPS P3AV(a) and their stabilities(b).

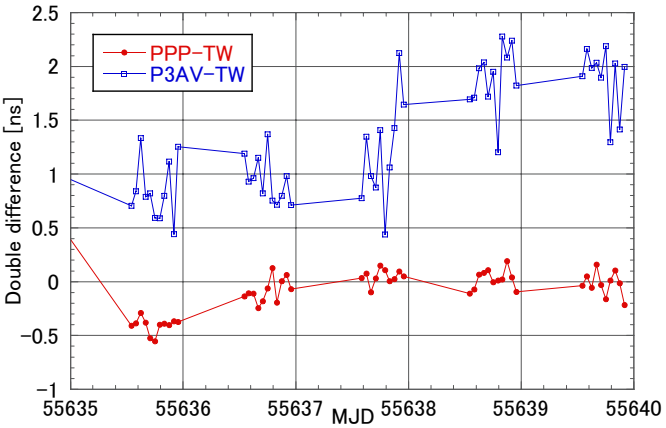


Fig. 9. Double differences between GPS and the AM-2 link.

[3] D. Kirchner, “Two-way satellite time and frequency transfer (TWSTFT): principle, implementation, and current performance”, in *Rev. of Radio Science*, Oxford University Press, 1999, pp. 27-44.
[4] A. Bauch et al., “Comparison between frequency standards in Europe and the USA at the 10^{-15} uncertainty level”, *Metrologia*, vol. 43, pp. 109-120, 2006.
[5] H. T. Lin et al., “The first two-way time transfer link between Asia and Europe”, in *Proc. of 35th Annual Precise Time and Time Interval (PTTI) Meeting*, 2003, pp. 71-79.
[6] H. Maeno et al., “Establishment of a TWSTFT link between Asia and Europe connecting NICT and PTB”, in *Proc. of 20th European Frequency and Time Forum (EFTF)*, 2006, pp. 575-579.
[7] M. Fujieda et al., “First comparison of primary frequency standards between Europe and Asia”, in *Proc. of 21th European Frequency and Time Forum (EFTF)*, 2007, pp. 937-941.
[8] T. E. Parker and V. Zhang, “Sources of Instabilities in Two-Way Satellite Time Transfer”, in *Proc. of 2005 Joint FCS/PTTI*, 2005, pp. 745-751.
[9] T. E. Parker et al., “Investigation of Instabilities in Two-Way Time Transfer”, in *Proc. of 34th Annual Precise Time and Time Interval (PTTI) Meeting*, 2002, pp. 381-390.
[10] M. Fujieda et al., “Long Baseline TWSTFT Between Asia and Europe”, in *Proc. of 38th Annual Precise Time and Time Interval (PTTI) Meeting*, 2006, pp. 499-511.
[11] D. Piester et al., “Studies on Instabilities in Long-baseline Two-way Satellite Time and Frequency Transfer (TWSTFT) Including a Troposphere Delay Model”, in *Proc. of 39th Annual Precise Time and Time Interval (PTTI) Meeting*, 2007, pp. 211-222.
[12] A. Bauch, D. Piester, M. Fujieda, and W. Lewandowski, “Directive for operational use and data handling in two-way satellite time and frequency transfer (TWSTFT)“, Rapport BIPM-2011/01, Bureau International des Poids et Mesures, Sevres, 2011.
[13] M. Fujieda et al., “Delay Difference Calibration of TWSTFT Earth Station Using Multichannel Modem”, *IEEE Trans. Instrum. Meas.*, vol. 56, no. 2, pp. 346-250, 2007.
[14] W. H. Tseng et al., “Analysis of Delay Fluctuations in Two-Way Time Transfer Earth Stations”, in *Proc. of 39th Annual Precise Time and Time Interval (PTTI) Meeting*, 2007, pp. 541-550.
[15] S. H. Yang et al., “Comparison of Time Transfer Using GPS Carrier Phase and Multichannel Two-Way Data in East Asia”, *IEEE Trans. Instrum. Meas.*, vol. 56, no. 2, pp. 664-668, 2007.

Slave to Master Clock Synchronization via TCP/IP for Time Distribution and Remote Calibration Applications

Laurent-Guy Bernier, André Stefanov

Time & Frequency Laboratory, METAS Federal Office of Metrology, Lindenweg 50, 3000 Bern-Wabern, Switzerland

***Abstract*—The goal is to distribute UTC(CH) in a remote location via TCP/IP. The MC (Master Clock) is locked to UTC(CH) via a combination of IRIG-B and 1-PPS signals. The SC (Slave Clock) is controlled by a PI control loop that tracks the MC. The SC-MC clock offset and the transmission delay are estimated simultaneously by means of a TW (Two-Way) comparison process similar to NTP but based on TCP/IP.**

I. Introduction

We seek to control a Slave Clock (SC) located in a remote location in order to track a Master Clock (MC) located at METAS. Clock comparison is performed via TCP/IP by means of the classical Two Way (TW) remote comparison technique. Every second, the TW processing yields a measurement of the TCP/IP transmission delay between the MC and the SC and a measurement of the SC-MC clock offset. A PI (Proportional Integral) control loop steers the slave clock by minimizing the low-pass filtered clock offset. The goal here is to use the network in order to synchronize a remote calibrated time reference. This is different from other works such as [2] that use an independent remote calibrated time reference in order to calibrate network controlled time sources such as NTP (Network Time Protocol) servers.

There is an optimum time constant which yields a minimal rms error when applied to the control loop. The better the stability of the oscillator used in the SC, the smaller the time synchronization error for a given amount of transmission delay noise. A typical NTP server steers the local system time generated from the PC microprocessor quartz oscillator, while the MC-SC system reported in this paper steers a hardware clock based on a high quality OCXO (Oven Controlled Quartz Oscillator).

The main limitation of such a MC-SC system is that the estimated clock offset is biased by the lack of symmetry of the send and receive TCP/IP transmission delays. Hence when the control loop steers the SC to minimize the low-pass filtered clock offset signal, the mean value of the resulting true clock offset is a constant different from zero that depends on the degree of asymmetry of the transmission delay.

The reason why we avoid using NTP is a need for flexibility regarding the network security rules. NTP is based on UDP/IP and other protocols that cannot be changed while we want to use TCP/IP and if required some encrypted point to point data link protocol in a taylor made software that complies with the specific network security requirements of the client.

This paper reports the laboratory experiments that we have performed in order to evaluate the performance of such a SC/MC system based on TCP/IP TW measurements. In a laboratory experiment the SC and MC are located in the same room, therefore it is possible to measure the true SC-MC clock offset by means of a time interval counter. In the real application, on the other hand, the performance of the SC control loop can only be estimated by analysis of the TW measurement results and of the the error signal.

II. Experimental Setup

The hardware is identical for the MC and for the SC. The hardware consists of a PC with a Spectracom TSYNC PCIe clock (PCI Express card). The clock is based on a 10-MHz OCXO.

The TSYNC PCIe clock is capable to synchronize automatically to a combination of time code and 1-PPS signal. The time code is used to determine the epoch and the 1-PPS signal is used to synchronize the OCXO as illustrated on Figure 1. This is what we do in the MC system in order to copy UTC(CH). The MC reproduces UTC(CH) with a 1 μs accuracy and generates an IRIG-B time code epoch, a 1-PPS clock signal and a 10-MHz frequency signal.

In the SC system, the OCXO frequency is steered digitally by software. The SC interrogates the MC via TCP/IP and controls itself by means of the PI error signal in order to track the MC as well as possible. The SC, like the MC, generates an IRIG-B time code epoch, a 1-PPS clock signal and a 10-MHz frequency signal.

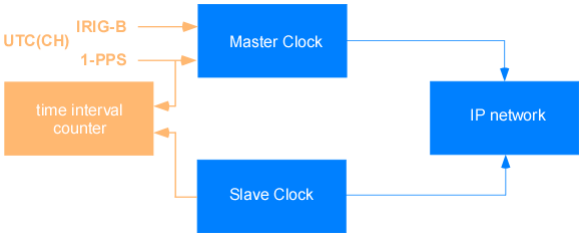


Figure 1. Experimental Setup

III. Two-Way Comparison Technique

The two-way comparison technique is similar to NTP [1]. The TCP/IP telegram exchanged between the SC and the MC

contains four binary time stamps.

The time stamps are defined in Table I. The TW measurements are generated as follows,

$$\theta = \frac{(T_2 - T_1) - (T_4 - T_3)}{2} \tag{1}$$

$$\delta = (T_4 - T_1) - (T_3 - T_2) \tag{2}$$

where θ is the estimated MC-SC clock offset and δ is the estimated TW transmission delay. Worst case clock offset error is half the TW delay (i.e. estimated one-way delay)

$$|\theta_{true} - \theta| < \frac{\delta}{2} \tag{3}$$

Note that, according to definition (1), the SC clock offset versus the MC, noted SC-MC, is given by $-\theta$. This is the sign convention used by Mills in [1].

Table I
DEFINITION OF TIME STAMPS T_1 TO T_4

Symbol	*Time Stamp Description*
T_1	Issued by SC when sending telegram to MC
T_2	Issued by MC when receiving telegram from SC
T_3	Issued by MC when sending telegram back to SC
T_4	Issued by SC when receiving telegram from MC

IV. THEORETICAL EXPECTATIONS

The Allan deviation of the OCXO is better than 1×10^{-10} and the Allan deviation of the TCP/IP delay is 1×10^{-9} for an averaging interval τ of $100ks$ (i.e. about one day) as shown on Figure 2. This means that with a control loop time constant as large as $T = 100ks$, the free OCXO is still more stable than the averaged noise of the TCP/IP half-TW delay. The time deviation of the measured TCP/IP delay is shown on Figure 3. The time deviation actually improves when the averaging interval is increased from $1s$ to $100ks$ but not that much because of the chaotic and non-stationary behaviour of the TCP/IP network. The theoretical expectation for the time stability of the SC-MC offset is about $20\mu s$ for a closed loop time constant of $T = 100ks$. Figure 4 shows a typical recording of the half-TW delay measurements. The chaotic nature of the TCP/IP transmission is visible. Figure 5 shows a typical histogram of the half-TW delay measurements. The histogram is multi-modal (in this example, mean $1.6ms$, standard deviation $74\mu s$). The average transmission delay of the TCP/IP packets across a switch depends on the number of clients connected to the switch and for a given network configuration, the mean value and the standard deviation of the delay between the MC and the SC are stationary. However the configuration of the network changes all the time as TCP/IP sessions are opened and closed by other clients. Hence the non-stationary behaviour of the delay.

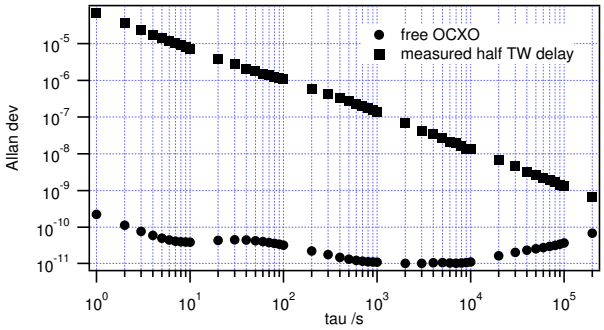


Figure 2. Allan Deviation of Free OCXO and of Half-TW Delay

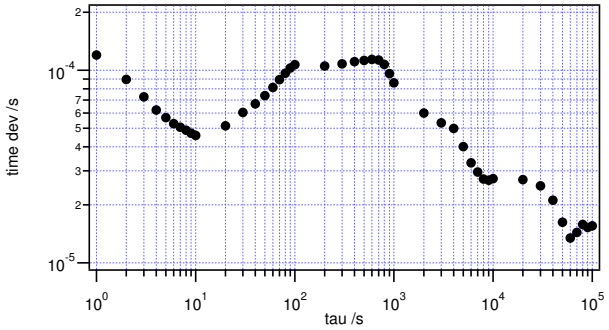


Figure 3. Time Deviation of Half-TW Delay

V. SECOND ORDER CONTROL LOOP

The control loop is similar to a PLL except that the control variable is time instead of phase. The control loop can be modelled as a linear feedback control circuit as illustrated on Figure 6. The DAC (Digital to Analog Converter) varactor frequency control of the SC is used to steer the OCXO. Since the control variable is time $x(t)$ and the error signal fed to the DAC is frequency $y(t)$, we have to model the DAC varactor control as an integrator. In order to have a zero time error in steady state, the loop filter $F(p)$ must include an integrator. Hence there are two integrators in series and the feedback control loop is potentially unstable. This is why we use a PI filter (i.e. the sum of a proportional gain and an integrator) as the loop filter $F(p)$. In closed loop the transfer function is of

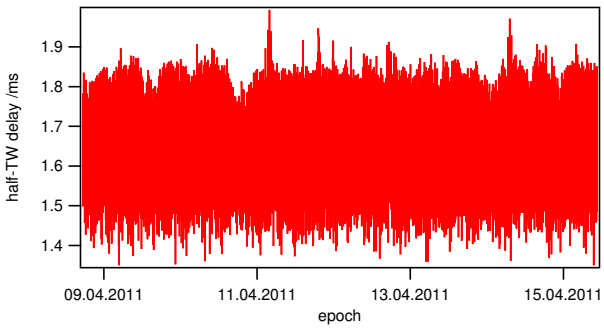


Figure 4. Typical Half-TW Delay Recording

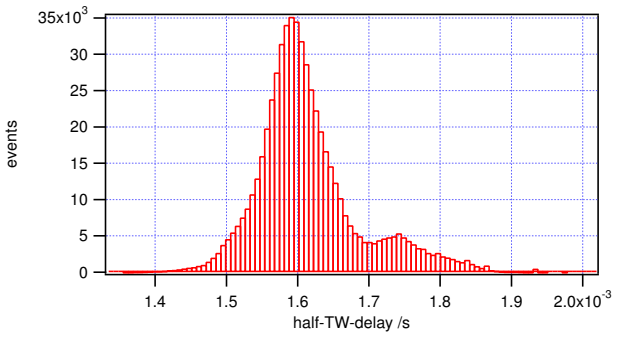


Figure 5. Histogram of Typical Half-TW Delay

the second order type and the two degrees of freedom given by the PI filter enable to set the time constant T as desired while keeping the damping at the critical value.

Calibration of the DAC frequency control of the OCXO yields a maximum rate of $\pm 1 \mu s/s$ (1 ppm) for a normalized maximum error signal of ±1 (DAC is actually16 bit). Figure 7 shows the initial transient when the control loop is started with a time constant of $1ks$. In this example the initial clock offset is $5000\mu s$. Initially the error signal is very large and the control loop works in non linear mode, i.e. the error signal is clipped to the 1 ppm maximum frequency offset of the OCXO which yields a slope of $1\mu s/s$. When the clock offset approaches zero, i.e. after 5000 s, the control loop starts to work in linear mode. Since the control loop parameters are adjusted for critical damping, there is a clean single overshoot, without any ringing visible in the transient.

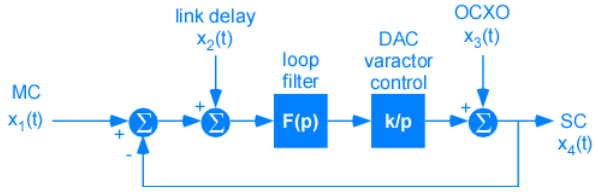


Figure 6. Schematics of PLL control loop

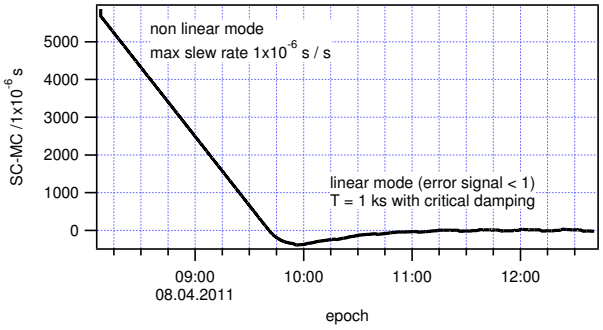


Figure 7. Initial Transient when Starting with Large Offset SC-MC

VI. Adaptative Control Algorithm

The advantage of a very long $T = 100ks$ control loop time constant is a very good noise filtering, but the disadvantage is the duration of transients: several days. An adaptative algorithm sets the SC epoch to the nearest second, when initialised, and then steers the SC frequency, switching between three discrete values of the time constant $T = 1ks$, $10ks$ or $100ks$ depending upon the value of the low-pass filtered SC-MC offset. Of course the goal is to minimize the duration of the transients. The loop is initialized with $T = 1ks$ and the time constant is progressively increased up to $T = 100ks$ in three steps as the low-pass filtered clock offset goes down. In the algorithm, a set of progammable criteria is used to decide when the time constant must be switched up or down.

VII. Estimation of the SC-MC Clock Offset

The measurement of the true SC-MC clock offset by means of a time interval counter is possible only in a lab experiment. In a real application the SC is operated in a remote location and only the TW data is available to the user. The best way to estimate the SC behaviour is by analysis of the error signal. Integrating the error signal using the calibrated DAC sensitivity gives a very good picture of the actual SC-MC time behaviour. Figure 8 shows an example of the true SC-MC clock offset versus the integrated error signal. We observe that the integrated error signal is a perfect image of the true SC-MC clock offset, except for the average value which cannot be determined by integration. By principle, again, it is not possible to deduce the true mean value of the SC-MC offset from the TW data because the asymmetry of the TCP/IP transmission delay is not an observable.

Figure 9 shows how the error signal changes when the control time constant is switched from $1ks$ to $10ks$ after the initial transient. When the time constant is increased by a factor of 10, the noise bandwidth is reduced proportionally. Therefore the rms value of the noise is suddenly reduced by a factor of 10.

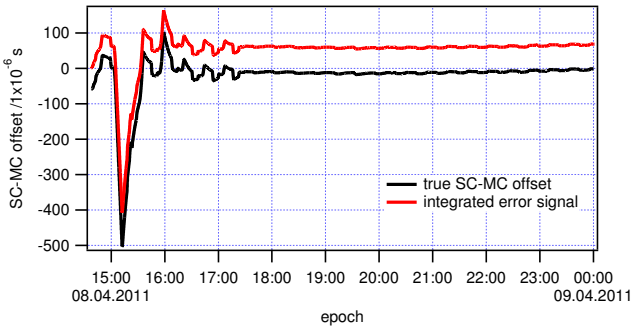


Figure 8. Example of True SC-MC Clock Offset vs Integrated Error Signal

VIII. Experimental Results

The results reported above in Sections IV-VII were taken from a simple single switch experiment, i.e. the MC and SC are located in the same room and connected to the same

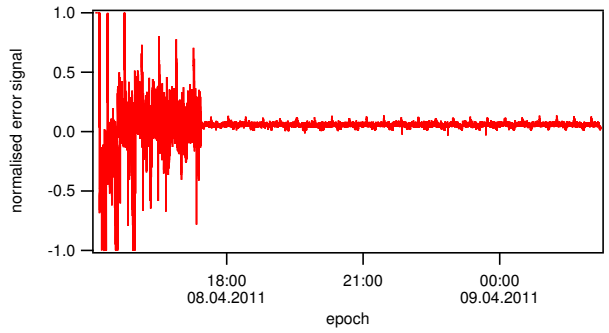


Figure 9. Example of Error Signal when Time Constant is Switched from 1 ks to 10 ks

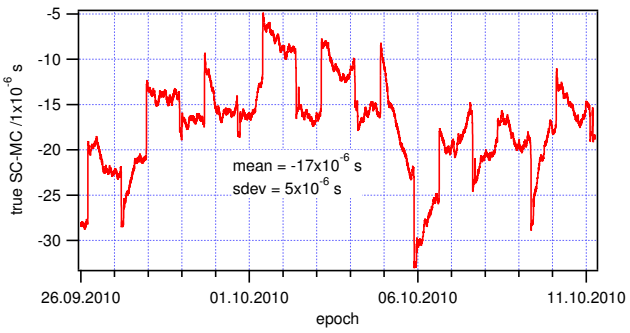


Figure 10. True SC-MC Clock Offset during Relay Experiment

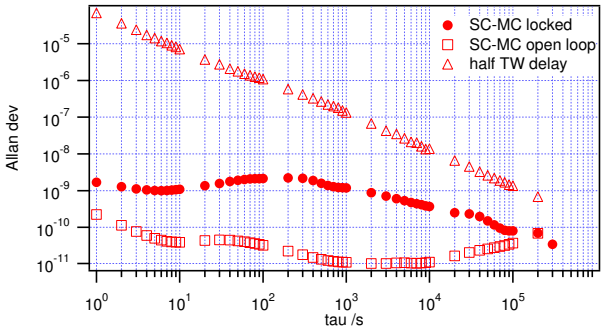


Figure 11. Allan Deviation of SC-MC Clock Offset: Locked vs Open Loop

Ethernet LAN switch. On the other hand the results of Section VIII are taken from a *relay* experiment which is more realistic regarding the everyday life TCP/IP network conditions. In the relay experiment the SC and MC computers are still connected to the same switch. However they do not communicate directly with each other. A third computer, located in another part of the METAS building, is used as a relay, receiving and forwarding all the TCP/IP packets exchanged between the SC and the MC. In the relay experiment several switches and routers are involved as well as the processing delays taking place inside the relay computer.

Figure 10 shows a 15 day recording of the true SC-MC offset as measured in steady state during the relay experiment. The peak-to-peak variation of the true offset over 15 days is only $25\mu s$. The SC-MC offset data shown yields a standard deviation of $5\mu s$ and a mean value of $-17\mu s$. Figure 10 shows three cases of the Allan deviation. Curve one is the SC-MC clock offset with the SC *locked* to the MC. Curve two is the SC-MC clock offset with the SC control *open loop*, i.e. with the OCXO free running. Curve three is the half-TW TCP/IP transmission delay. The linear feedback control loop actually filters out a large part of the TCP/IP delay noise but not all of it. If the delay noise were not filtered at all, then the *locked* SC-MC Allan deviation would be identical to the Allan deviation of the TCP/IP delay noise. On the other hand, if the delay noise were removed completely, then the *locked* SC-MC Allan deviation would be identical to the Allan deviation of the free running OCXO. As expected the *locked* SC-MC Allan deviation is to be found somewhere between these two limit cases. It can be observed that, in terms of Allan deviation, the noise suppression at 1s averaging time is 5 orders of magnitude.

IX. Conclusions

The performance of the prototype SC-MC system reported here, true SC-MC average offset $< 20\mu s$, is more than sufficient for the application we have in mind. Actually we intend to operate a similar SC-MC system with the MC located at METAS in the time & frequency lab and with the SC located at the BIT, the Swiss federal administration IT center. By means of the SC, UTC(CH) will become available to the BIT stratum 1 NTP servers together with other redundant sources of UTC.

References

[1] Mills D.L., "Internet Time Synchronization: the Network Time Protocol", *IEEE Trans. Communications*, COM-39, N°10, Oct. 1991, pp. 1482-1493.

[2] Smotlacha V., Cermak J., Palacio J., "On Calibration of Network Time Services", *Metrologia*, 45 (2008), pp. S51-S58.

[3] Vanier J., Têtu M., "Transfer of Frequency Stability from an Atomic Frequency Reference to a Quartz Crystal oscillator", IEEE Trans. on Instrum. and Meas., IM-28, N°3, Sept. 1979, pp. 188-193.

Timing and synchronization in mobile telecommunication networks

Rimantas Miškinis, Dmitrij Smirnov, Emilis Urba
Time and Frequency Standard Laboratory
Center for Physical Sciences and Technology
Vilnius, Lithuania
Email: miskinis@pfi.lt

Bronislovas Dzindzelėta
JSC "Omnitel
Vilnius", Lithuania

***Abstract*—Recent developments in mobile telecommunication networks offer a possibility for synchronization of mobile equipment. The possibility is explored here. To this end, a network operator is synchronized to the Lithuanian national time and frequency standard over internet via NTP (Network Time Protocol). A mobile testing system (MTS) created is synchronized to the operator's network via a radio modem operating in one of the several different modes. The time scale of the MTS synchronized over the mobile telecommunication network is compared to that of the national standard to evaluate the synchronization accuracy and to find out that the most suitable operating mode for the modem is CSD (Circuit Switched Data), which is likely due to its dedicated and symmetrical nature.**

I. INTRODUCTION

Implementation of new services and technologies in mobile telecommunication networks requires more precise timing and synchronization of the network systems as well as user's equipment. Correct timing may be very important for different M2M (Machine To Machine) applications, such as metering (measuring parameters of telecommunication channels), fleet management, mobile commerce, and public key infrastructure (PKI) applications, such as e-health, e-government, e-signature, etc. Some results on time synchronization over the second generation network GSM (Global Standard for Mobile Communications), via the GPRS (General Packet Radio Services) and EDGE (Enhanced Data rates for GSM) channels, are presented in the works [1, 2]. However, the two aspects are not considered – timing over dedicated low bit second generation networks via the CSD (Circuit Switched Data) channel and timing over the third generation network UMTS (Universal Mobile Telecommunications System) via the HSDPA (High-Speed Downlink Packet Access) channel. Networks of the third generation are very popular now. Nevertheless, channels of the third generation networks exhibit propagation characteristics very different from those of the second generation networks, such as GSM. The purpose of our work is a comparative testing of timing and synchronization characteristics of mobile GSM network via the CSD, GPRS, EDGE channels as well as over the UMTS (Release 99) network via the HSDPA channel. Propagation characteristics of different channels have been evaluated by using standard telecommunications testing equipment. Timing characteristics have been evaluated by using the data of NTP communications created during the synchronization of the mobile testing system (MTS) to the NTP servers of the Lithuanian National Time and Frequency Standard Laboratory.

II. TESTING TECHNIQUE

Timing characteristics of mobile telecommunication networks were evaluated by using NTP communications data created during the mobile testing system (MTS) synchronization with NTP servers of Lithuanian National Timing Laboratory (LT). The mobile testing system (MTS) developed consists of Radio modem, Computer (PC) running on LINUX UBUNTU server v.10.10 operating system and second pulses generator (Rb) (Fig. 1). The PC running the Linux Ubuntu Server 10.10 operating system, with operating LinuxPPS Kernel Module (PPS API (Pulse Per Second Application Programming Interface) implemented under RFC 2783) and ntpd (ntp_v4.2.6p3@1.2290 package) program installed, has been connected to the mobile communication network via the "Siemens MC35i" and "Nokia N95 8GB" modems.

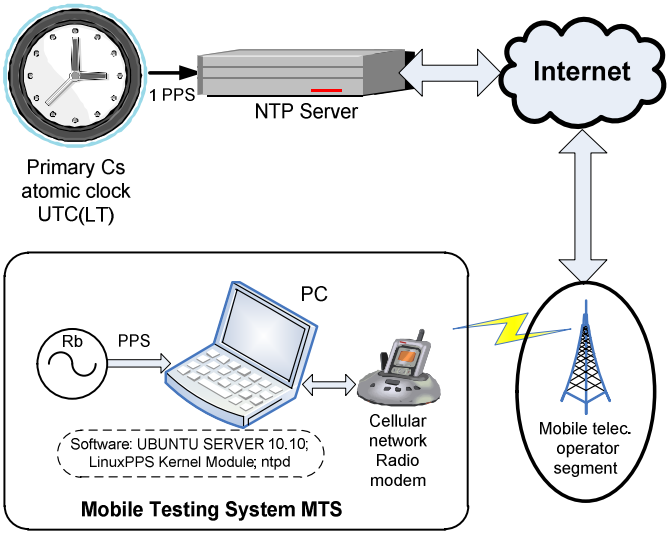


Figure 1. Device setup for investigation of time transfer via the mobile communication network

Such a setup is makes up a complete MTS version ready for measurements. The “Siemens MC35i” has been connected to the computer via the RS232 port and operated in the GSM-CSD and GSM-GPRS regimes. The “Nokia N95 8GB” has been connected to the PC via the USB interface and used to explore the time synchronization possibilities for the EDGE, UMTS, and HSDPA technologies.To estimate the possibilities for time synchronization in a mobile communication network, non-standardized measuring equipment has been used. Therefore, before making any measurements, it was necessary to explore its metrological characteristics. Fig. 2 reveals how the devices have been connected for MTS calibration. The calibration setup includes a primary Cs atomic clock HP5071A reproducing and keeping the UTC(LT) time scale and a “Symmetricom S250“ NTP server. The second pulses from the Rb generator, which during the calibration procedure was synchronized to UTC(LT) with a few nanosecond accuracy, are fed into the RS-232 port of PC. The LinuxPPS Kernel Module, having received a second pulse at the RS-232 port, forms a very accurate time mark, which is possible due to the fact that time marks are formed at the level of the system kernel. PC‘s clock has been synchronized every 16 seconds, using the ntpd daemon operating under the NTP protocol; it has received the time marks at the local (with the propagation time of less than 1 ms) network from the NTP server, which time scale is synchronized with UTC(LT) (Universal Time Coordinated realized by Lithuanian National Time Standard Laboratory) with the accuracy better than ±1 µs.

One can see from Fig. 3 that while synchronizing the computer clock in the local network yields a constant deviation of about 51 µs. It could be influenced by either the LinuxPPS module or the ntpd together with asymmetry of the local area network (LAN). Fig. 4 reveals the stability characteristics of the MTS system clock. From Fig. 4, one can see that at time intervals up to 40 s white phase noise of the LinuxPPS system prevails, while at longer intervals the noise character becomes different. The difference is caused by the fact that although the ntpd daemon is synchronized to the NTP server every 16 seconds, the corrections to the MTS‘ system clock frequency feedback loop are sent upon necessity only,

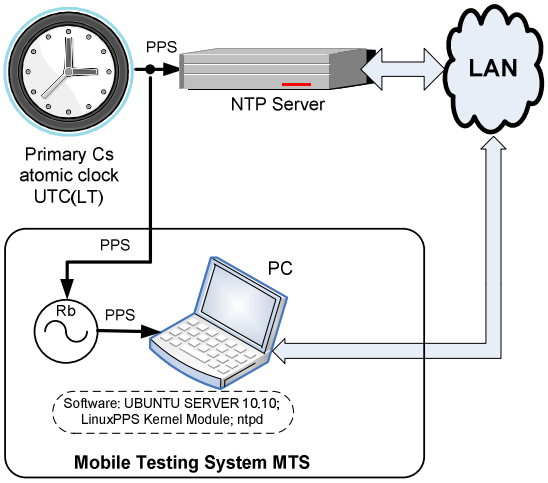


Figure 2. Schematic layout of MTC calibration

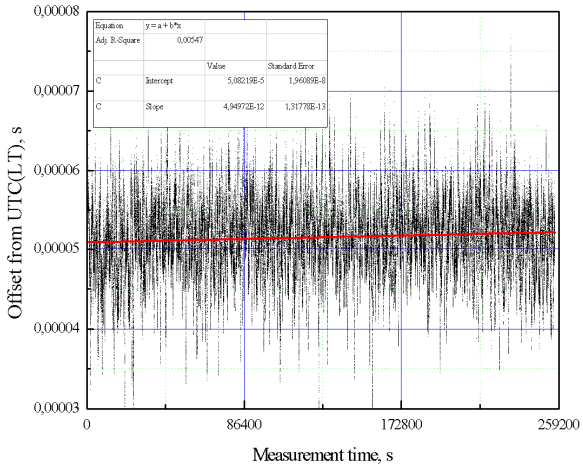


Figure 3. Sattering of the LinuxPPS – formed time marks, while synchronizing the computer clock by the NTP protocol to the server operating in the local network

and, in the most cases, it takes place every 100 s. Starting with intervals of about 1000 s, the white phase noise of the system becomes prevalent. Ntpd program has its algorithms, which automatically adjust the synchronization interval, which enables to improve the stability characteristics at short intervals (100...1000 seconds) and to decrease significantly the number of synchronization events, however, this doesn‘t improve the stability of the clock rate essentially (Fig. 5).

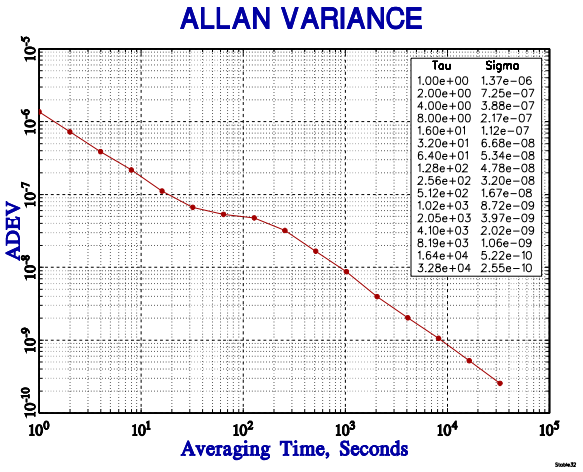


Figure 4. Stability characteristics of the MTS system clock, while synchronization takes place every 16 seconds by the NTP protocol in the local network

From the measurements carried out, one can conclude that MTS has a constant error not worse than 51 µs, while the uncertainty in phase measurement at the intervals starting with 1 s is better than ±1 µs.

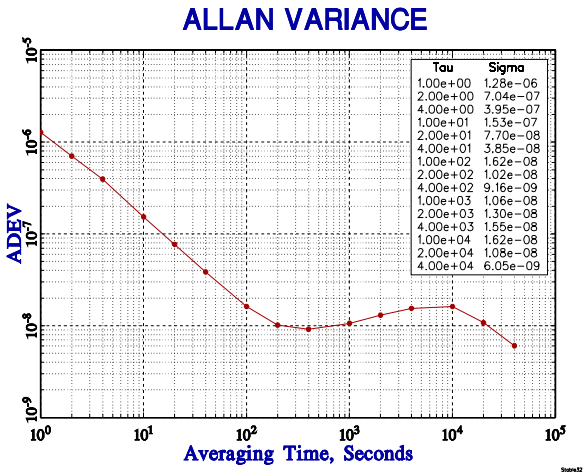


Figure 5. The stability characteristics of the system clock of MTS in the case when the decision concerning the necessity of synchronization is adopted by the ntpd program, and synchronization takes place in the local network by the NTP protocol

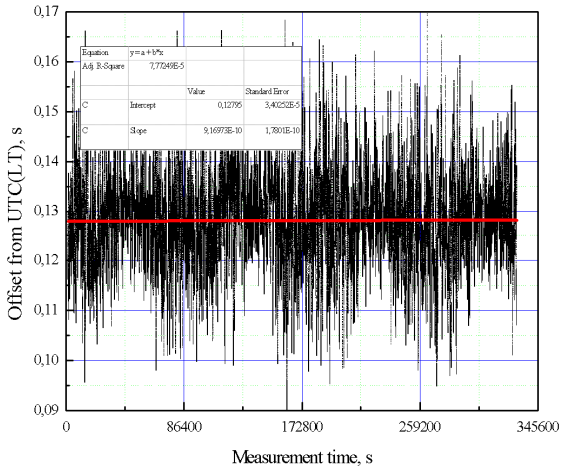


Figure 6. Scattering of the LinuxPPS–formed time marks, while synchronizing the computer clock to the time server by the NTP protocol, by means of the GPRS technology

III. RESULTS

Further, we will review only the results obtained while operating the MTS in the same base station (cell). Statistically, switching to another station is rare, however, such a switching usually causes the packet delay of about 3-5 s due to cell reselection procedure Sometimes, switching is even longer due to few cells reselection in raw. Yet, the investigation of such phenomena requires separate experimental measurements. We don't account for such phenomena here. In order to minimize cell reselection procedures mobile terminal (MTS) was placed close to the base station to have dominant cell all experiment time.

At the beginning, consider the results of the synchronization of the MTS clock while the time scale is transmitted by means of the GPRS technology.

Fig. 6 implies that, despite ntpd's effort to synchronize the system clock by minimizing its deviation from UTC(LT), average deviation of the system clock is about 130 ms. The worst thing is that we cannot see this deviation in the ntpd log files as it is caused by the network asymmetry, which cannot be detected and compensated by the ntpd. Network asymmetry in GPRS is caused by GPRS data scheduling differences in uplink direction (UL) and downlink direction (DL) due to both different algorithms as described in [4] and different cell loads in UL and DL.

Synchronization of the system clock by means of the EDGE technology yields results very similar to those obtained in the GPRS case; network asymmetry is about 150 ms.

The Allan deviation for the clock synchronized every 16 seconds via the GPRS tract has a maximum at 100 seconds because, as mentioned before, the corrections to the frequency feedback loop of the MTS system clock are sent upon necessity only, and, in the most cases, it takes place every 100 seconds (see Fig. 7). The Allan deviation graph implies that it is possible to decrease the overall noise level by means of decreasing the number of synchronization events. In this case, the corrections to the MTS system clock would be sent at longer time intervals because the short-term (100 s, 1000 s)

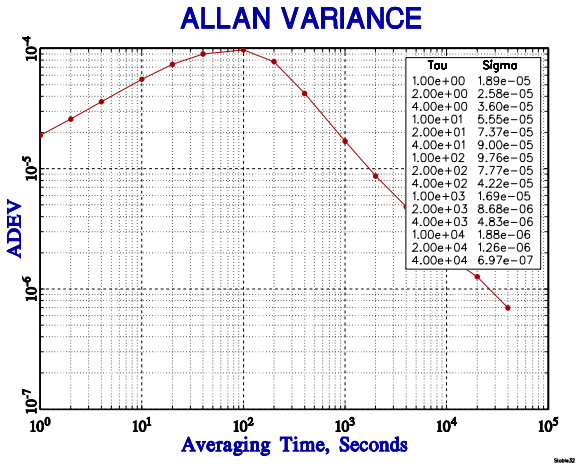


Figure 7. Stability characteristics of the MTS system clock, while synchronization takes place every 16 seconds by the NTP protocol, by means of the GPRS technology (ADEV – Allan Deviation [3])

stability of the clock is good enough. The experiments confirm the inference; Fig. 8 gives the system clock stability with synchronization every 1024 s as an example.

At the beginning of the graph of Fig. 8 (roughly for time intervals of up to 5 s), the measurement noise of the LinuxPPS part of the system is seen, whereas at longer intervals the PLL noise becomes dominant. At long intervals (10^2...10^4 s), its level decreases.

We can see that choosing the synchronization interval of about 1000 s allows obtaining rather good system clock stability. If the objective were to achieve the MTS system clock stability better than 4 µs for any averaging time, it would be reasonable to increase the synchronization interval even more, however, it is necessary to be careful at this point and to ensure that no other process can influence MTS system clock. It is possible to achieve in the laboratory conditions, however, it is more difficult to ensure when the computer works under heavy load.

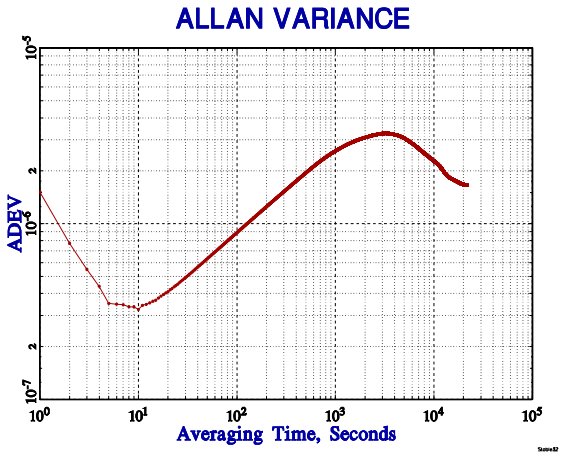


Figure 8. System clock stability with synchronization every 1024 s by the NTP protocol, by means of the GPRS technology

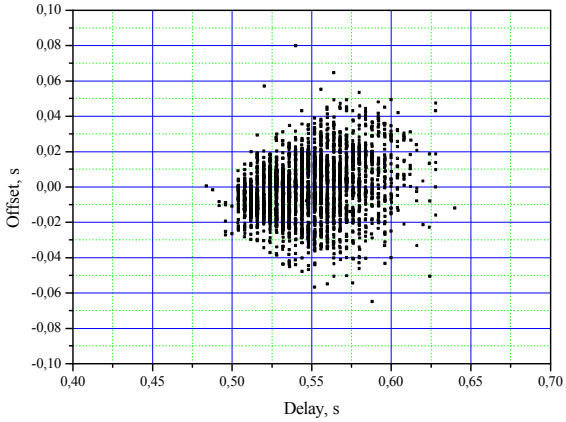


Figure 9. Delay vs.Offset characteristics of the MTS system clock, while synchronization takes place every 16 seconds by the NTP protocol, by means of the GPRS technology

Fig. 9 shows delay versus offset characteristics of the MTS system clock (generated from “peerstats” nptd log file), while synchronization takes place every 16 seconds by the NTP protocol, by means of the GPRS technology; it is seen that the discreteness in delay values is 4 ms. The EDGE technology yields similar results.

Experiments with the UMTS/HSDPA technologies reveal even larger asymmetry of about 230 ms. It is worth noting that this regime gives the worst results. Besides that, it isn‘t possible to evaluate the operation characteristics of UMTS and HSDPA separately. Depending on the network load, the mobile telecommunication operator automatically redirects the phones to use one or another technology. Statistically, in our case about 80% of measurements have been carried out under the UMTS, while the rest – under the HSDPA technology. HSDPA resources with different schedulers and different cell . Generally, the ratio can differ depending on the provider of the mobile communication. Reasons for asymmetry are mostly the same as those in the case of GPRS/EDGE – shared UMTS or loads in UL and DL directions are used for transmission.

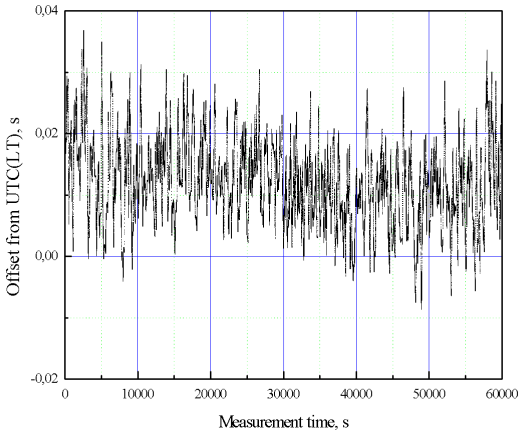


Figure 10. Scattering of the LinuxPPS–formed time marks, while synchronizing the computer clock to the time server by the NTP protocol, by means of the CSD technology

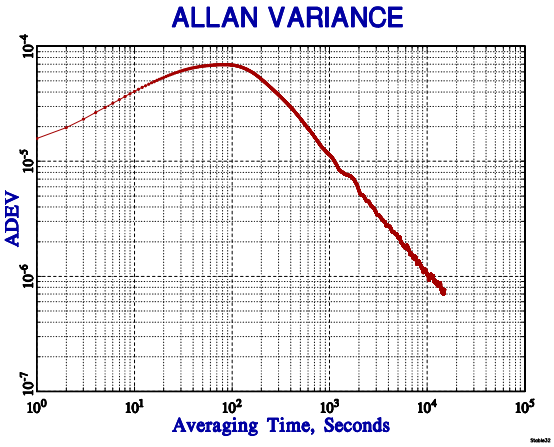


Figure 11. Stability characteristics of the MTS system clock, while synchronization takes place every 16 seconds by the NTP protocol, by means of the CSD technology

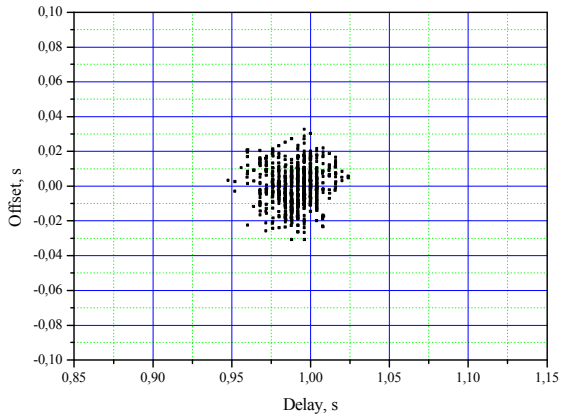


Figure 12. Delay vs.Offset characteristics of the MTS system clock, while synchronization takes place every 16 seconds by the NTP protocol, by means of the CSD technology

In addition, the GSM CSD channel with 9600 bps speed has been investigated. It works only within the GSM-type networks and has the largest possible delay of about 1 s due to bigger interleaving depth, however, asymmetry of the synchronization channel is only 10 ms. Measurement results are shown in Figs. 10 and 11, and can be explained taking into account that the CSD channel is a dedicated resource, therefore, the load doesn't influence channel delay as in the cases of GPRS/EDGE and UMTS R99/HSDPA. Also, the CSD channel processing in GSM is very similar in UL and DL directions, so propagation characteristics are almost the same in both directions.

Fig. 12 shows delay versus offset characteristics of the MTS system clock (generated from "peerstats" nptd log file), while synchronization takes place every 16 seconds by the NTP protocol, by means of the CSD technology.

IV. CONCLUSIONS

1. If the computer clock is synchronized via the mobile communication network by the NTP protocol, the network asymmetry is more important factor than the absolute delay time.
2. The delay time has no impact on the synchronization accuracy at all. While synchronizing by means of the CSD technology, the delay within the system is about 1000 ms, while within the HSDPA network it is about 80 ms, however, the asymmetry within the CSD network is almost 20 times less, whereas the stability of the clock rate synchronizing CSD and HSDPA differ insignificantly. The client doesn't see network asymmetry and can suppose that his clock behaves exactly as it shall do.
3. The stability of the clock rate while synchronizing the computer clock by the NTP protocol via a radio modem little depends on the technology of the signal transmission over the radio channel; it depends to the largest extent on the duration of the synchronization interval.
4. The CSD regime is best suited for time synchronization. It features the smallest asymmetry out of all mobile internet technologies. Besides that, stability characteristics of the MTS system clock in the case the synchronization is carried out by the NTP protocol, by means of the CSD technology, aren't worse than those obtained using other technologies.
5. The GPRS and EDGE technologies are very similar with respect to asymmetry and noise.
6. UMTS and HSDPA feature the worst parameters – the largest asymmetry and noise (although the delays are the smallest out of all the technologies).

REFERENCES

[1] Kovar, P., Molnar, K., "Precise Time Synchronization over GPRS and EDGE", *IFTP International Federation for Information Processing"*. Vol. 245. *Personal Wireless Communications*, B. Simak, R. Bestal, E. Kozowska, Eds. Boston: Springer, 2007, pp. 277-284. 1999.

[2] R. Miškinis, A. Burokas, J. Karuža, V. Simanavičius, D. Smirnov, E. Urba, A. Mašonis, "Timing of GSM audio connections by generating UTC(LT) traceable digital time stamps", *Proc. IEEE Int. Freq. Contr. Symp. jointly with the 21th Eur. Freq. Time Forum* [Online], 2007, pp. 836 – 839. Available: (www.ieee-uffc.org/main/publications/fcs/proceed/2007/178_7177.pdf)

[3] D. W. Allan, "Time and Frequency (Time-Domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators," *IEEE Trans. Ultrason. Ferroelectr. Freq. Contro*l, vol. UFFC-34, pp. 647-654, Nov. 1987.

[4] J. Eberspächer, Ch. Bettstetter, H.-J. Vögel, *GSM: architecture, protocols and services*, 3rd ed. London, UK: John Willey and Sons Ltd., 2009.

Examination of Time and Frequency Control Across Wide Area Networks Using IEEE-1588v2 Unicast Transmissions

A. Novick, M. Weiss, K. Lee, D. Sutton
of National Institute of Standards and Technology (NIST)
Boulder, CO, USA
Email: novick@nist.gov

***Abstract* - The IEEE 1588-2008 Precision Time Protocol (PTP) version 2 (IEEE 1588v2) can be used to synchronize a slave clock to a grandmaster clock over a wide area network (WAN). However, many of the algorithms the slaves use to steer to the master are optimized for a scenario where both devices are on the same subnet or local area network (LAN). This paper is a study of existing PTP hardware from a number of different manufacturers in unicast mode. We characterize the performance of the equipment, beginning with the timing outputs of the masters that are locked to their built-in Global Positioning System (GPS) receivers. Next, we compare the results of steering unicast clients to their masters through a LAN versus several wider-area network configurations such as virtual-LANs and the public Internet. Analysis of the results will show how clients of different manufacture handle the various network paths. It is our hope that these comparisons will instigate changes to clock steering and synchronization algorithms, which may help improve the overall capabilities of PTP for telecom and other networking environments. As network synchronization techniques improve, the quality of the PTP masters will become more significant. Therefore, the performance and calibration of PTP masters with respect to UTC(NIST) is also discussed.**

I. Introduction

Network timing has become of great interest over the last decade, especially since the implementation of IEEE 1588-2008, where Precision Time Protocol (PTP) was defined to enable precise synchronization of clocks in measurement and control systems through packed based networks [1]. PTP uses a two-way timing method [2] to characterize the delays between two different points across a network. By repeatedly exchanging time-stamped messages between two devices (a master and a slave), the slave clock can be synchronized to, and will 'follow', the master as long as data are being exchanged. The client attempts to determine the delay of the packets through the network and adjusts the timing output to compensate. The use of 1588v2 in a multicast environment is limited primarily to local area networks (LANs) where the client devices (slaves) search the network for masters and begin to exchange time stamps. Network providers may or may not allow multicast on public networks, so the unicast method should be used outside of a LAN [3]. A unicast client can find a master outside the LAN if it has the IP address of the master to look for. We employ the unicast method because we are using the public network in some cases, and we also want to guarantee that we are using a designated PTP master.

The purpose of this paper is to compare the timing synchronization abilities of PTP hardware from several manufacturers in different real-world network settings to show the capabilities and weaknesses of these devices. We were not merely trying to determine which one is the 'best' or to advocate particular hardware. No brand names are used. Also, no boundary clocks or transparent clocks are implemented in the path, only *ordinary clocks* (master and slave) are used, and there is no use of simulations or network traffic generators.

In order to assess the output of a remotely steered client device, the master must first have a reference time source. Typically, the masters synchronize to signals from an internal Global Positioning System (GPS) receiver, a Network Time Protocol (NTP) server or a 1 pulse-per-second (pps) input. Ideally, we would synchronize the master to the NIST timescale, UTC(NIST), and measure the output of the slave compared to UTC(NIST). However, some of the master devices do not allow a 1 pps input as the reference and many facilities may not have an on-time 1 pps available. Therefore, we looked at comparisons using the GPS receiver built in to the master as the reference. Also, the internal receivers were allowed to survey their own positions, as opposed to entering known coordinates. Later in this paper, we will discuss the time offset and stability of the GPS reference inside each master unit and how we calibrated them with UTC(NIST).

II. Experiments

We set up a test bed for multiple PTP masters and clients, including GPS antennas on the roof, time interval counters and multiple network connections, including one outside the NIST firewall on the public network. Figure 1 shows the first measurement setup to compare the 1 pps outputs of the PTP masters to UTC(NIST) by use of a commercial time interval counter (TIC) with an external time base frequency input from UTC(NIST). The results of the measurements were recorded directly into files every second on a desktop computer with a serial connection to the TIC. We calibrated the masters for twenty-four hours and entered the average delays into the units to compensate for their offset from UTC(NIST) due to position errors, cable delays and hardware delays.

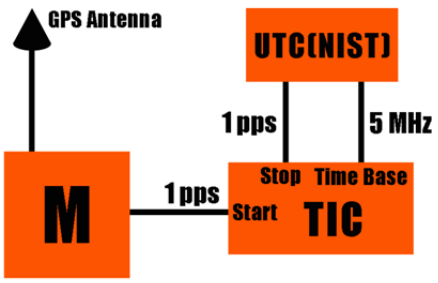


Figure 1. Comparing the PTP master (M) to UTC(NIST)

The PTP masters have internal GPS disciplined oscillators (GPSDOs) that can vary significantly from device to device [4]. Figure 2 shows a comparison of the 1 pps outputs of the four PTP masters, with peak-to-peak variations that range from 30 ns to greater than 110 ns.

Our goal was to compare multiple PTP clients over several different network scenarios, so we began with the simplest setup: a master and slave on the same subnet in the same room (Figure 3). Each client was configured to follow the master of the same manufacture, although subsequent experiments showed cross-operability. The devices were not connected together with a crossover cable but, instead, connected to the LAN via separate jacks in the wall.

The clients were configured to be in unicast mode, with a synchronization rate of 64 packets per second. Also, we allowed the clients ample time to steer to the PTP data before recording any measurements. Our interest was in how well they can do in a steady-state, as opposed to how quickly they can lock to the master. We compared the 1 pps output of each client to UTC(NIST). On the same subnet, the synchronization varied between the different PTP clients, as shown in Figure 4.

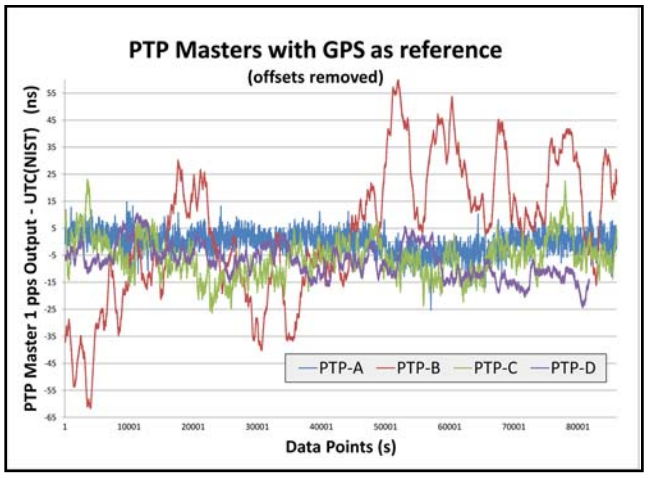


Figure 2. Comparisons of PTP masters to UTC(NIST)

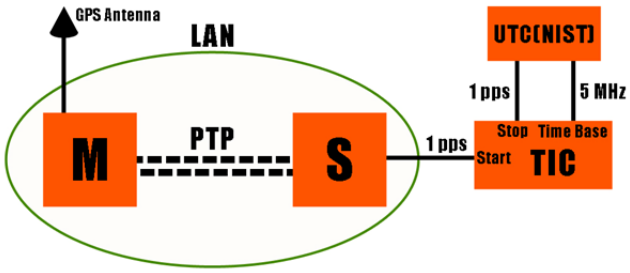


Figure 3. Measuring a slave (S) on the same subnet as the master (M)

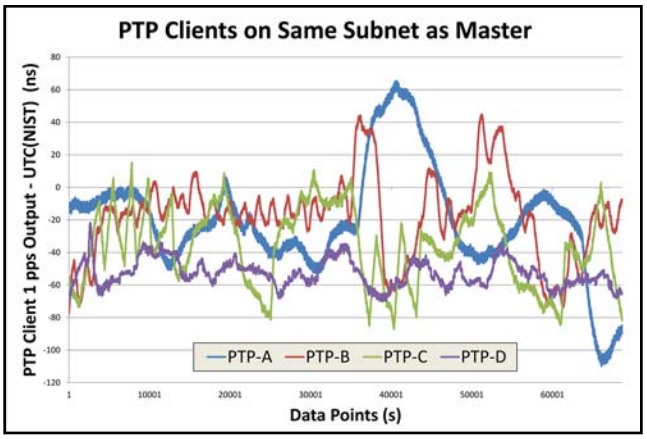


Figure 4. Comparisons of PTP clients on the same subnet

The average time offset for each client was less than 60 ns compared to its master, but the peak-to-peak variations ranged from 45 ns to over 170 ns. It is interesting that the PTP-A master had the tightest range, but the PTP-A client had the largest range. Figure 5 shows how closely each client followed its master. Different steering methods by manufacturers cover a wide range of possibilities for how closely a PTP client follows the master. Following too closely (PTP-B) causes worse stability in the network scenarios we measured, especially when, in this case, the master is the least stable. The PTP-C slave seems to over-react to changes in the master with high gain adjustments.

Next, we put the PTP masters on the public network outside the firewall at the NIST location in Boulder, Colorado, while the clients were still on the internal LAN (see Figure 6). The masters and clients were in the same building, but there were several network elements in between. These elements (referred to as "hops") can cause noise and asymmetries in the timing of the data exchange. In this case there were five hops between the devices.

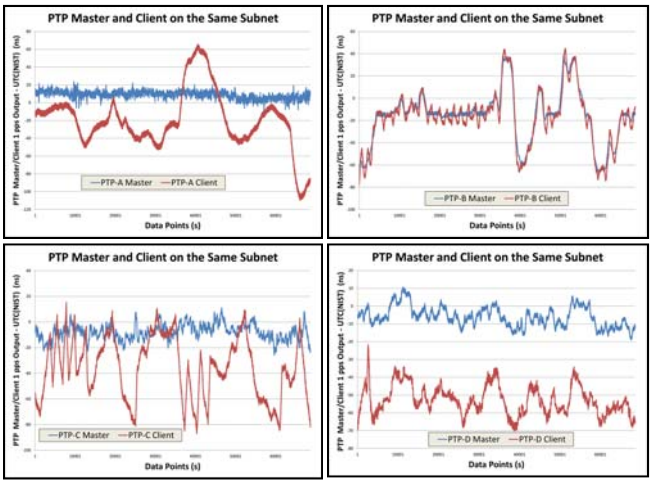


Figure 5. Graphs showing how well each client followed its master

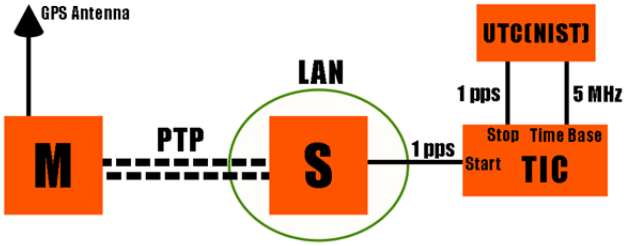


Figure 6. The master is outside the NIST local area network

The differences between the client devices with masters outside the local network are shown in Figure 7. The average time offset of PTP-D was 2.3 µs and the peak-to-peak variation was 4.2 µs. PTP-B had a similar average, but the stability was worse, varying by 9.4 µs. PTP-C had a large average time offset of 22.9 µs, but a good stability, with a peak-to-peak range of 2.0 µs. The timing stabilities are shown in the time deviation (TDEV) plot in Figure 8. The timing requirement for some telecommunication networks is one microsecond [5,6]. None of the devices meet this criterion of accuracy in this case. However, some of the devices remain below one microsecond in stability, so if the network path were calibrated and the delay amount were entered into the client device so its output was advanced by the delay amount, meeting the requirement may be possible. However, much longer data runs would be necessary to see how the network delays change over time. Also, a reference at the client site (like GPS) would be necessary in order to calibrate the PTP client output. We measured an NTP client with a server outside the local network as well. It had an average time offset of 55.4 µs, and $\sigma_x(\tau)$ is shown to be worse than one microsecond for τ >10,000 s (Figure 8). It should be noted that NTP has a considerably slower synchronization rate and a much different intended use than PTP.

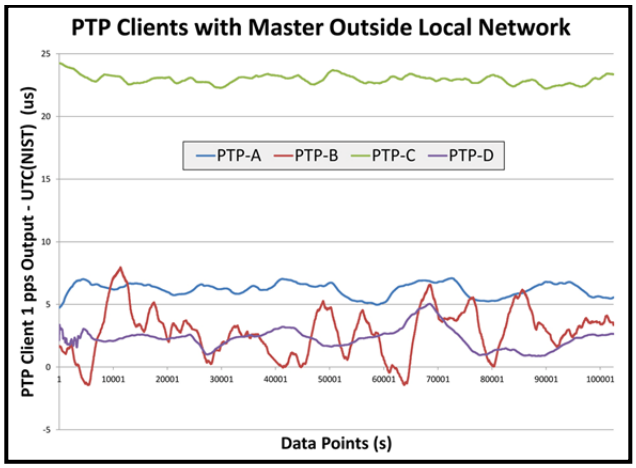


Figure 7. Performance of the clients with the master outside the LAN

The International Telecommunication Union (ITU) Option 2 (less stringent) requirement mask for a locked slave clock in telecommunication networks [7] is shown in Figure 8. The slaves did not stay under this mask for τ > 100 s (best case).

We were also able to compare some of the units at remote NIST sites. Radio stations WWV and WWVB broadcast time signals from a NIST site near Ft. Collins, Colorado, which is about 80 km (direct path) from the NIST location in Boulder. There is a dedicated T1 (1.544 Mbit/s) leased line between the two sites that is considered to be virtual LAN (VLAN) inside the NIST network. We set up the client devices at the remote site (Figure 9) and compared them to 1 pps signals from a backup UTC(NIST) generated at that site, which is typically within 30 ns of the official UTC(NIST) in Boulder. There were four network hops between the master and client in this setup.

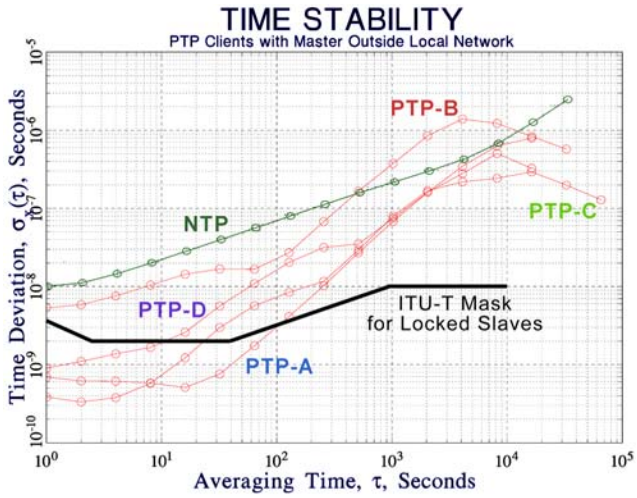


Figure 8. TDEV plot of the clients with the master outside the LAN

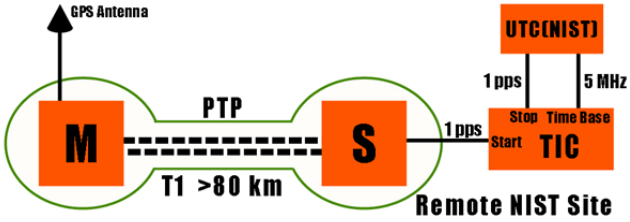


Figure 9. Remote PTP synchronization setup

PTP-A and PTP-D performed the best in this case (Figure 10), but each had an average time offset > 30 µs. PTP-D had the best peak-to-peak variation (< 1.4 µs). As before, if the network path were calibrated, and the outputs of the clients were adjusted to be centered around zero, this still may meet the needs of some applications, *assuming the characteristics of the network do not change*. The TDEV plot in Figure 11 shows the time stability of PTP-D (and maybe PTP-A) remaining below a microsecond, but none of the clients were within the ITU requirements for a locked slave [7] for τ > 250 s.

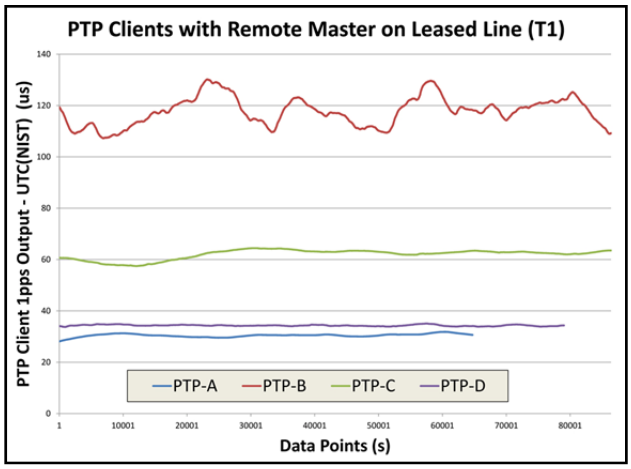


Figure 10. Performance of clients with remote master

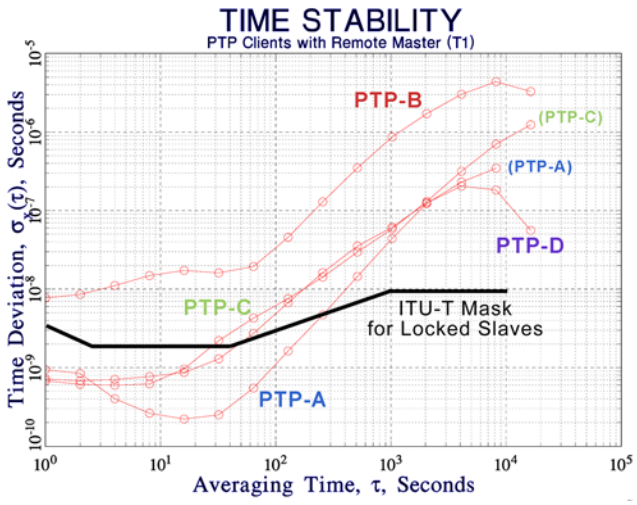


Figure 11. TDEV of clients with remote master

Next, we compared the clients when they were synchronized to the masters by use of the public Internet. The masters were at the NIST Boulder site on the public network, and the clients were at the NIST radio broadcast site, also on the public network (Figure 12). There were nine network hops between the sites.

PTP-D was the "best" with an average time offset of just under 10 ms (Figure 13) and a peak-to-peak variation of 56 µs. The NTP measurement of a remote server using the public Internet was actually better, remaining below 3 ms but varying by more than 2 ms peak to peak. The TDEV plots (not shown) for PTP and NTP clients with a master on the public Internet do not approach the quality of the telecommunication requirements. There may be client settings on some of these PTP devices that would work better in this environment, but from our observations, PTP clients do not work well on the public Internet and should be steered after much more averaging, at least to match the average time offset of NTP. However, not only are the network paths very asymmetric to begin with, but the packets can change course from one synchronization to the next, so calibration of the client output may not improve the timing performance.

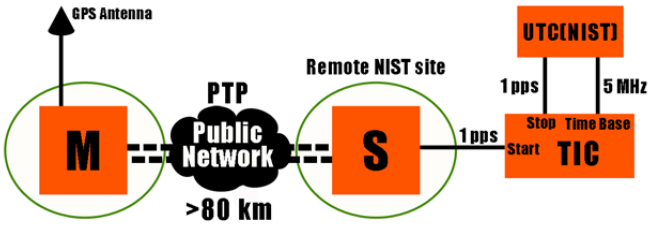


Figure 12. PTP master and slave across the public internet

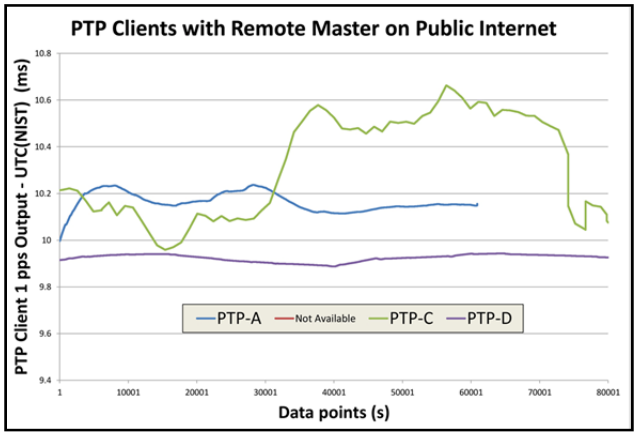


Figure 13. Clients with remote masters the public Internet

The last remote master experiment was conducted between NIST in Boulder, Colorado and NIST in Gaithersburg, Maryland, which is over 2400 km away. There are dual T3 (44.7 Mbit/s) leased lines composing the NIST VLAN, with a significant amount of traffic between the two sites. In this case there were five network hops. Because there is no equivalent to UTC(NIST) generated at the Gaithersburg site, the master there used GPS as the reference. The client was in Boulder, and its output was compared to UTC(NIST). Figure 14 shows that the average time offset was 473.0 µs, with peak-to-peak variations of 128.7 µs. PTP over a leased line may not have large timing steps due to packets taking different paths, but the timing variation is large, possibly caused by variable (and abundant) network traffic.

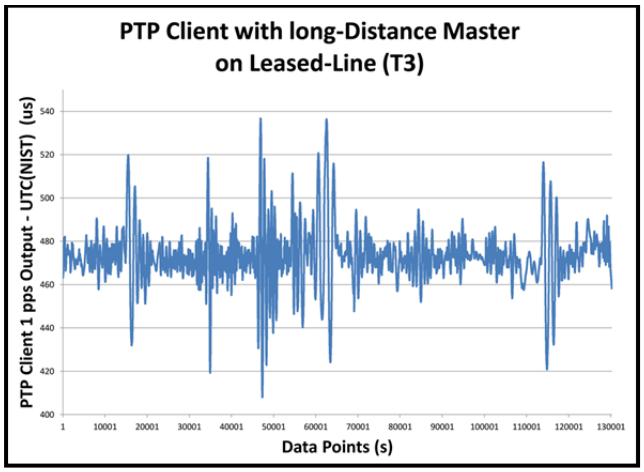


Figure 14. Results with client at a site 2400 km away

III. Calibration of the Masters

Each master's GPS receiver was allowed to survey its own antenna position, but the exact locations of the antennas were well known due to geodetic surveys performed recently. Also, the antenna cable delays were measured. Table I shows the errors in latitude/longitude and height of the surveyed coordinates compared to the geodetic solution. Also, the table shows the measured antenna cable delays and the delays entered into the PTP masters to adjust the output to match UTC(NIST). The difference between these two delays is the resulting master delay, which is affected by the height survey error and the delay through the master itself. It is not known why PTP-D has a negative receiver delay, but it is possible that the manufacturer advances the output of the device to compensate for *presumed* delays. Most important, without calibration of the master, the output can be in error, which will result in client error on the other end.

TABLE I. GPS Errors in the PTP Masters

	Lat/Lon Survey Error	*Height Survey Error*	*Antenna Cable Delay*	*Output Adjustment (based on calibration)*	*Resulting Master Delay*
PTP-A Master	0.4 m	18.2 m	122.1 ns	163 ns	40.9 ns
PTP-B Master	9.0 m	13.3 m	123.3 ns	147 ns	23.7 ns
PTP-C Master	2.7 m	- 0.9 m	108.9 ns	112 ns	3.1 ns
PTP-D Master	1.9 m	19.1 m	101.6	56.0 ns	-45.6 ns

A GPS disciplined oscillator (GPSDO) is sometimes considered a self-calibrating device. However, we have shown that even when the antenna cable delay is entered, position error and receiver delays have caused average time offsets over 40 ns in some cases. Also, even though the mid-term frequency stabilities of the masters (Figure 15) are very good because they are steered to GPS, we noted earlier (Figure 2) that there were short-term peak-to-peak fluctuations of greater than 120 ns in the worst case. Figure 16 shows all of the masters within the ITU specification for time stability of primary reference clocks [8]. At present, there is no ITU requirement for time accuracy. In order to know how well a particular PTP master works with a GPSDO as its reference, it is advisable to have it calibrated by NIST or another National Metrology Institute (NMI) or a calibration laboratory [4]. Ideally, a GPSDO calibration should be performed with the antenna and cable that will be used in the field, leaving only the coordinate error. In order to achieve the best performance with a calibrated GPSDO, precisely surveyed coordinates should be entered into the receiver.

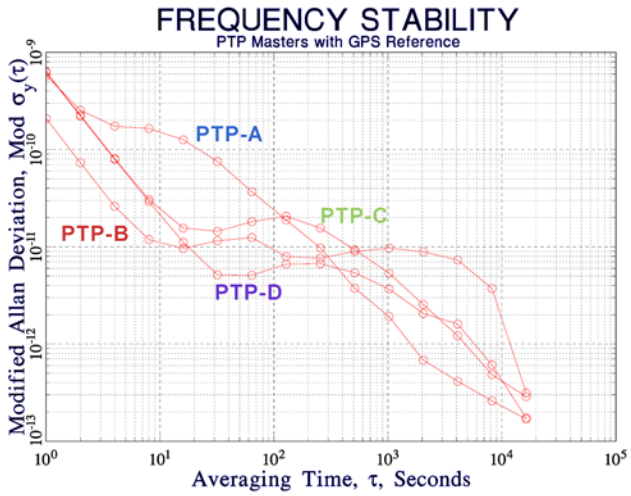


Figure 15. Frequency stability of the PTP masters

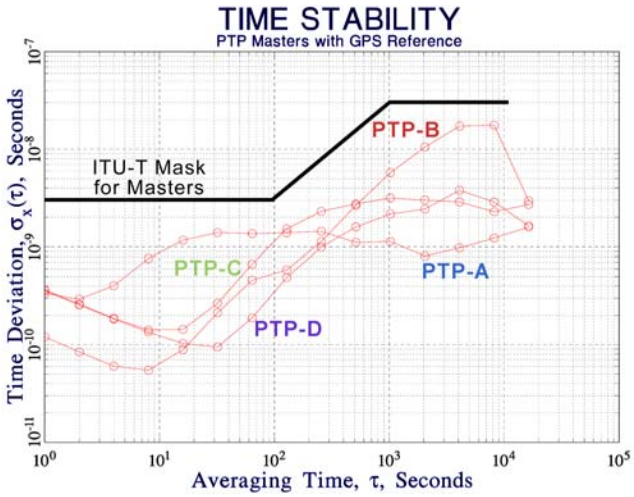


Figure 16. Time stability of the PTP masters

IV. CONCLUSION

These results show several real-world network scenarios and how different commercial PTP hardware reacts in each situation. We see that there is usually one that is worse than the other three, but it is not always the same one. As more distance and network elements are added to the network, the asymmetry of the paths can increase, causing larger time offsets. As PTP continues to grow as a tool for synchronization across networks, and the types of network situations continues to vary, we hope that providers of PTP hardware continue to develop their devices and steering methods to account for this. Also, increasing network infrastructure and opportunities for different levels of service (symmetric paths) from network providers will improve long-distance PTP synchronization. Another possibility of operation is GPS-assisted PTP, where the client would also have an embedded GPS receiver, and the PTP timing would be calibrated continually, so it could be used in a holdover mode if GPS becomes unavailable.

ACKNOWLEDGMENT

We would like to thank Robert Sorensen and Dr. Eugene Song of NIST for their help with these measurements.

REFERENCES

[1] IEEE-1588, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," July 2008.

[2] J. Edison, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems, A Tutorial," 2005 Conference on IEEE-1588, 10-12 October, 2005, Winterhur, Switzerland.

[3] L. Cosart, "Precision Packet Delay Measurements Using IEEE 1588v2," on Proceedings of the 2007 International IEEE Symposium on Precision Clock Synchronization (ISPCS) for Measurement, Control and Communication, 1-3, October 2007, Vienna, Austria, pp. 85-91.

[4] M. A. Lombardi, A. N. Novick and V. S. Zhang, "Characterizing the performance of GPS Disciplined Oscillators with Respect to UTC(NIST)," in *Proceedings of the 2005 Joint Meeting of the IEEE International Frequency Control Symposium and Precise Time and Time Interval (PTTI),* 29-31 August 2005, Vancouver, Canada, pp. 677-684.

[5] J. Ferrant, M. Gilson, S. Jorbert, M. Mayer, L. Montini, M. Ouellette, S. Rodrigues, S. Ruffini, "Development of the First IEEE 1588 Telecom Profile to Address Mobile Backhaul Needs," IEEE Communications Magazine, vol. 48, no. 10, October 2010, pp. 118-126.

[6] 3rd Generation Partnership Project 2 (3GPP2), "Recommended Minimum Performance Standards for cdma2000 Spread Spectrum Base Stations: Release B," *C.S0010-B, Version 2.0*, February 2004.

[7] ITU-T G.813, "Timing Characteristics of SDH Equipment Slave Clocks," March 2003.

[8] ITU-T G.811, "Timing Characteristics of Primary Reference Clocks," September 1997.

Time and frequency transfer combining GLONASS and GPS data

Pascale Defraigne
Royal Observatory of Belgium
Brussels, Belgium
p.defraigne@oma.be

Aurélie Harmegnies, Gérard Petit
Bureau International des Poids et Mesures:
92312 Sèvres, France
aharmeg@bipm.org, gpetit@bipm.org

***Abstract*—In order to allow the use of GLONASS data from geodetic type receivers providing only raw data in RINEX format, we updated the R2CGGTTS software developed initially at the Royal Observatory of Belgium (ROB) for GPS data. The GLONASS navigation files are used for the determination of satellite clocks and positions, and the computation procedure to get the CGGTTS data from the pseudorange measurements is applied similarly as for the GPS satellites. Using the satellite orbits and clock products provided by the IGS analysis center ESA, we combine GPS and GLONASS in All in view. Adding GLONASS decreases the noise of the clock solution with respect to the GPS-only solution by about 20% thanks to the increased number of satellite observations. We also quantify the added value of using GLONASS data in addition to GPS data for time and frequency transfer based on Precise Point Positioning (PPP). A particular attention is given to the determination of the frequency-dependent hardware delays in both approaches.**

I. INTRODUCTION

Measurements from Global Navigation Satellite Systems (GNSS) are used since the eighties [1] to perform precise and accurate Time and Frequency Transfer (TFT). In its classical version, the GPS time transfer is performed using clock offsets collected in a fixed format, called CGGTTS (Common GPS GLONASS Time Transfer Standard), as described in [2,3]. These clock offsets are obtained from the pseudorange measurements, corrected for the signal travel time (satellite-station), for the troposphere and ionosphere delays, and for the relativistic effects. A smoothing is then performed over 13 minutes observation tracks. Starting with C/A code receivers, the method was then upgraded to take benefice from the dual-frequency receivers measuring codes on both frequencies, which allows to remove the ionosphere delays at the first order (i.e. 99.9 percent of the effect), thanks to the ionosphere-free dual-frequency combination. This led to a factor 2 improvement in the precision of the intercontinental time links (e.g. [4]). Some time receivers directly provides the CGGTTS results, but in order to be able to use classical geodetic receivers driven by an external clock for the time transfer applications with the same standards, a dedicated software named R2CGGTTS was produced to compute the CGGTTS files from the raw observation data provided by the receivers in the RINEX format [5]. This software is presently used by a large portion of the time laboratories for their participation to TAI, through the technique TAIP3 [6]. Thanks to the growing constellation of GLONASS, some time laboratories now upgraded their equipment with receivers capable to observe both GPS and GLONASS. We therefore adapted the R2CGGTTS software to the use of GLONASS observations as will be explained in the first part of the paper.

A more precise version of GNSS TFT consists in using both code and carrier phase observations. The use of PPP, i.e. Precise Point Positioning [7], for time and frequency transfer has been extensively studied and developed in the last years [e.g. 8-11]. Presently several time links in TAI are based on PPP solutions [12]. Only GPS measurements are currently used, while the method should find a significant improvement from the combined use of additional constellations, GLONASS, Galileo, COMPASS, thanks to the increased number of satellites and signals. We propose a first step in this evolution with the combination of GPS and GLONASS data using the Atomium PPP software [9]. This is now possible thanks to the effort of the IGS during the last decade to use the GLONASS data for positioning and providing therefore precise orbits for the GLONASS satellites consistent with the GPS satellite orbits.

A major difference between the GLONASS and GPS constellations is that all the GLONASS satellites do not transmit the same frequency, so that the signal delay in the receiver is different for each satellite group emitting a given frequency. This affects the code measurements, with differential biases up to 25 ns, as already shown by several authors [13-14]. These differential biases must be either determined by calibration or estimated in addition to the clock solution.

The paper will study the added value of GLONASS in CGGTTS and PPP. The PPP results were fully presented in [15], we just here provide a summary.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

II. CGGTTS COMBINING GPS AND GLONASS DATA

As explained in the introduction, the CGGTTS files contain clock solutions corresponding to the pseudorange measurements, corrected for the signal travel time (satellite-station) based on broadcast satellite orbits, for the troposphere and ionosphere delays, and for the relativistic effects. The final results provided are the midpoint of a linear fit performed over 13 minutes observation tracks for each visible satellite. Using dual-frequency observations allows to remove the first order ionosphere delay, which will be done here with GLONASS observations.

As the two time scales GPS and GLONASS are differing from an integer number of leap seconds (GPS Time = GLONASS Time + leap seconds) and as the observation files are dated in GPS time, while the broadcast navigation message uses GLONASS time as reference, a first correction for system time must be applied. Secondly, while GPS navigation messages provide parameters of keplerian orbits in WGS84 (very close to the ITRF) each two hour, the broadcast navigation message of the GLONASS constellation is given as a set of positions and velocities in the PZ90 Earth-fixed system at a sampling rate of 30 minutes. As we are computing the CGGTTS from RINEX data in a post-processing, we have the choice between computing the GLONASS satellite positions either by numerical integration, or by polynomial interpolation. A comparison between the CGGTTS results obtained for one day with either interpolation or numerical integration Runge Kutta 4th order was performed [16] using the GLONASS pseudoranges from a RINEX observation file provided by the TTS4 receiver located at the BIPM, and connected to a H-maser. As the noise of the results was larger with the polynomial interpolation, we decided to implement the Runge Kutta 4th order integration in the R2CGGTTS software. Finally, as done by the BIPM for the computation of TAI, we corrected the CGGTTS results computed with broadcast navigation messages by introducing more precise satellite positions and clocks as computed by the ESOC analysis center of the International GNSS Service (IGS). The results so obtained are presented in Fig. 1, with a comparison with similar results for the GPS satellites. We can directly observe a larger dispersion of the GLONASS results with respect to the GPS results, due to the satellite-dependent hardware delays associated with the satellite-dependent carrier frequencies of the GLONASS constellation. This is illustrated in Fig. 2 which presents the differences between the clock solutions obtained with the GLONASS satellites and the GPS All in View (AV) clock solution for the BIPM receiver BP1B. The satellite of the same GLONASS frequency channel have been plotted with a same color; each color represents therefore the biases of one frequency channel. The frequency-dependent biases are clearly visible on the picture. It must be noted that these biases correspond to biases of the P3 combination, i.e. a combination of the frequency-dependent biases of the P1 and the frequency-dependent biases of P2. The variations of the ensemble from day to day are due to the day-to-day changes of reference of the ESA satellite clock solutions. Changing the reference can indeed also change the bias between the GPS pseudoranges and the GLONASS pseudoranges.

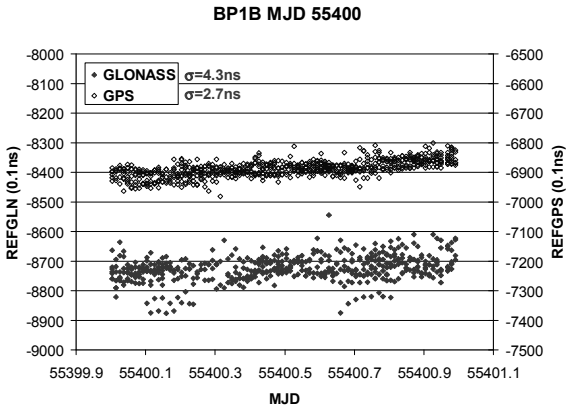


Figure 1. Comparison between the CGGTTS results obtained with either GLONASS satellites or GPS satellites after correction for precise satellite positions and clocks. The ESOC post-processed satellite orbits and clocks are used in both case.

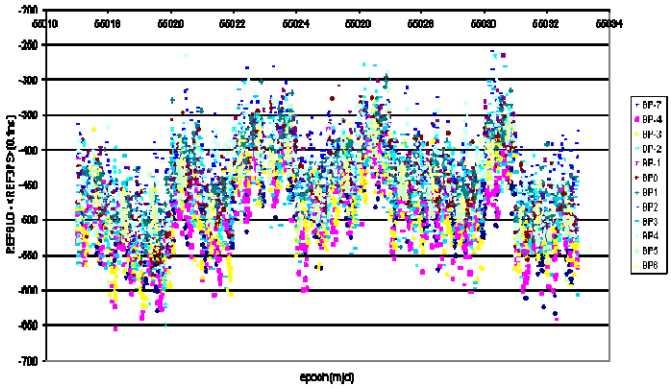


Figure 2. Differences between the clock solutions obtained with the GLONASS satellites and the GPS clock solution for the BIPM receiver BP1B. Each color corresponds to one GLONASS frequency channel.

In order to compute an All in View solution with GLONASS and GPS data, it is therefore necessary to correct the CGGTTS results for the inter-frequency and inter-system biases. As presently no GPS-GLONASS receivers have been calibrated for the GLONASS part, the procedure we propose it to determine for each month (regular basis for the TAI computation) and for each GLONASS frequency channel, one bias with respect to the GPS solution (hopefully calibrated). In order to account for the day-to-day differences due to the changes of reference, the procedure proposed is as follows: first determine for each day j a daily bias *D(j)* between GPS and GLONASS, as the average of the differences between GLONASS results and the GPS AV solution during that day. Then corrects all the frequency-dependent biases of Fig. 2 for *D(j)*, so that for each day they are centered on zero. The monthly bias *B(f)* for each frequency channel *f* is then computed as the average of the biases reported to a zero average. Finally, for each day, the correction for frequency-dependent biases is obtained as the sum of the monthly bias *B(f)* and the daily offset between GPS and GLONASS results *D(j)*.

Figure 3 presents an example of AV solution obtained with GLONASS data, before and after the correction for frequency-dependent biases applied with the procedure just described. The improvement of the stability of the solution is clearly visible.

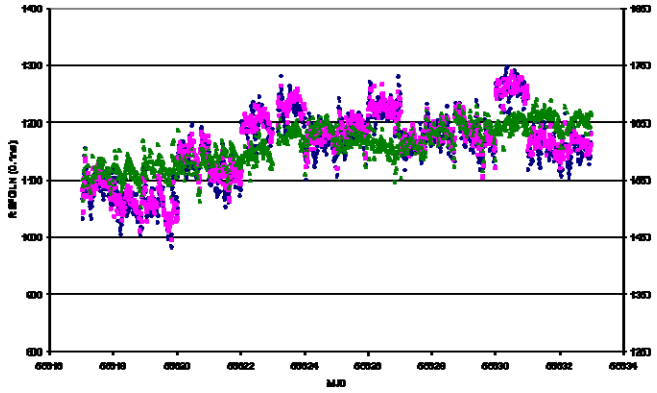


Figure 3. AV clock solution obtained with the GLONASS satellites BIPM receiver BP1B. blue: without correction of inter-frequency biases; pink: using a monthly mean of the biases; green: using a daily mean of the biases.

In order to quantify the added value of GLONASS observations in time transfer links based on CGGTTS results, we computed the link BIPM-IPQ, using the RINEX files from the TTS4 receivers installed in both labs. The time links obtained in All in View with either only GPS or GPS+GLONASS are plotted in Fig. 4, and the stability of both solutions is depicted in Fig. 5. Thanks to the increased number of observations, the noise of the final solution is reduced by about 20% when adding GLONASS data.

III. PPP COMBINING GPS AND GLONASS DATA

The combination of GPS and GLONASS observations for time and frequency transfer was done using the Atomium software [9]. This software is based on a least square analysis of code and carrier phase GPS measurements, using satellite orbits in the format sp3 and clocks in RINEX clock format, to determine the receiver clock synchronization error at each observation epoch, a daily position and the tropospheric zenith path delays each 15 minutes. The precise satellite positions and clocks were also taken from the ESOC products.

The clock solution, obtained from the ensemble of GPS and GLONASS observations, corresponds therefore to the synchronization error between the ground clock and the ESOC reference.

The inter-frequency hardware delays are estimated as additional parameters in the least square inversion, as biases with respect to the GPS pseudoranges. As the data are analysed in daily data batches, one set of frequency-dependent hardware delays is determined for each day analysed.

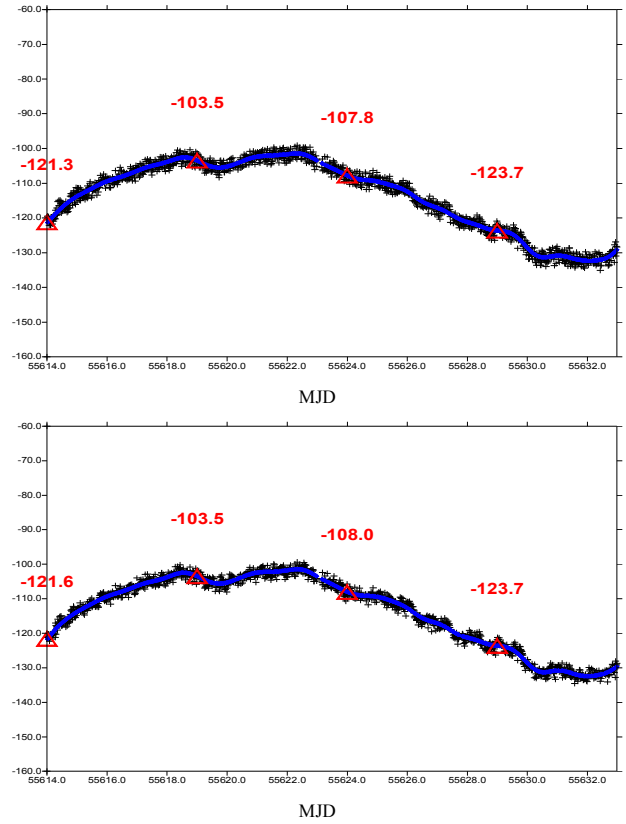


Figure 4. Time link BIPM-IPQ during 18 days, based on All-in-View, up: using only GPS data, down: using GPS+GLONASS data. Units are nanosecond.

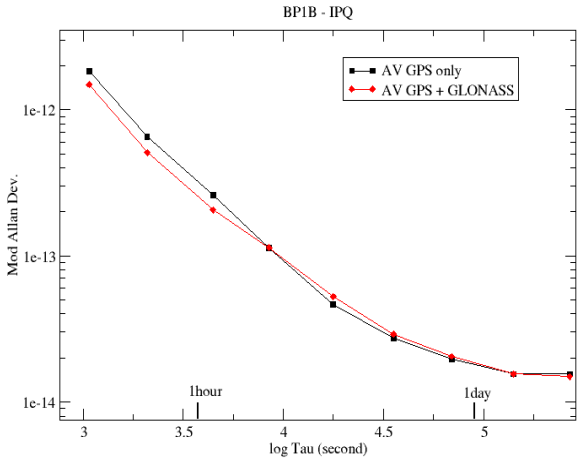


Figure 5. Mod. Allan Deviation of the two curves presented in Fig. 4.

Fig. 6 presents the differences between the PPP clock solutions obtained with GPS data only or using both GPS and GLONASS data for three IGS stations equipped with H-masers: MAR6, ONSA and WTZR. These differences are lower than 200 ps peak to peak. The differences were also computed over a 6 week period for the station WTZR and the same conclusion was drawn concerning the magnitude of the differences.

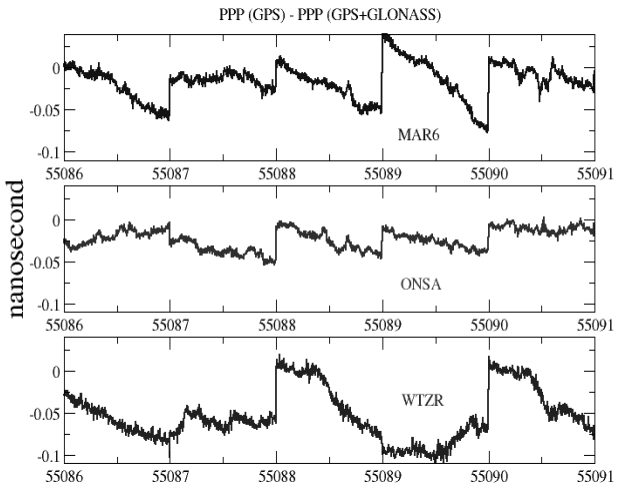


Figure 6. Differences between the PPP clock solutions obtained with GPS data only or using both GPS and GLONASS data for the IGS stations MAR6, ONSA and WTZR.

Figure 7 shows one situation where the addition of GLONASS observations brings an improvement to the global solution. There was a tracking interruption, with a very short data gap in the WTZR observations. This forces the Atomium software to estimate different ambiguities before and after the gap, using approximately two times less pseudorange data in each part. This jump of 200 ps disappears when adding GLONASS code and phase measurements to GPS data within the analysis, thanks to the increased number of data which favors a better determination of the ambiguities in both parts, before and after the data gap in WTZR.

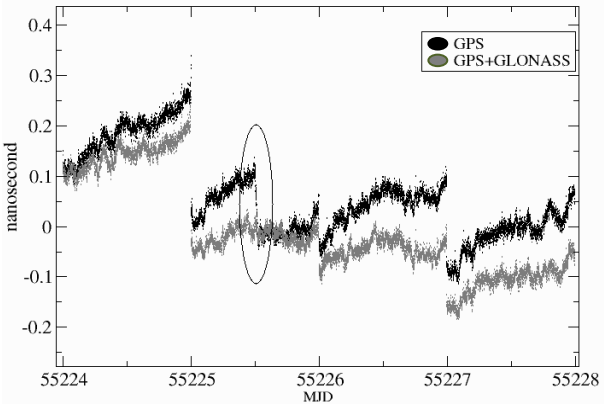


Figure 7. Differences between the Atomium PPP solution and the IGS solution for the time link ONSA-WTZR, with a tracking interruption in WTZR during the MJD 55225.

IV. CONCLUSIONS

The **R2CGGTTS** software has been upgraded and is now able to be applied on GLONASS data. Thanks to the ESA products, it is possible to combine GPS and GLONASS data in AV. The inter-frequency biases of GLONASS data should however firstly been corrected for. These biases are computed on a monthly basis with respect to the GPS solution, further removing the daily average of the differences between GPS and GLONASS solutions due to the daily change of reference of the ESA clock products. For the time link investigated here, i.e. BIPM-IPQ, the addition of GLONASS data to the GPS data in AV reduces the noise level by about 20% thanks to the increased Clock solutions from PPP can now be obtained from an ensemble of GPS + GLONASS observations, thanks to the ESOC products for satellite clocks. Concerning time transfer, i.e. the absolute value of the synchronization errors, the addition of GLONASS observations brings a significant improvement in the case of short data batches, because the higher number of observation allows a better determination of the carrier phase ambiguities This can be important in case of tracking interruption as was the case in the example presented here.

Concerning the frequency transfer with PPP, it was shown that adding GLONASS data modifies the shape of the curve, with a maximum difference of 150 ps peak to peak between the results obtained with and without GLONASS over the 6 week analysis performed.

ACKNOWLEDGMENT

We acknowledge the IGS and ESA/ESOC for making their solutions available.

REFERENCES

[1] D.W. Allan, M. Weiss, "Accurate time and frequency transfer during common-view of a GPS satellite", Proc. 1980 IEEE Freq. Contr. Symp., Philadelphia, PA, pp. 334-356, 1980.

[2] D.W Allan., C. Thomas, "Technical directives for standardization of GPS time receiver software", Metrologia 31, PP. 69-79, 1994.

[3] J. Azoubib, W. Lewandowski, "CGGTTS GPS/GLONASS data format Version 02", 7th CGGTTS meeting, 1998.

[4] P. Defraigne, G. Petit, "Time Transfer to TAI using geodetic receivers", Metrologia 40, 184-188, 2003.

[5] P. Defraigne, G. Petit, "Proposal to use geodetic-type receivers for time transfer using the CGGTTS format", BIPM Time section Technical Memorandum TM.110, 2001.

[6] G. Petit, Z. Jiang, "Stability and accuracy of GPS P3 time links", Proc. 18th EFTF, 2004, CD-Rom (file 066.pdf)

[7] J.Kouba, P. Heroux, "Precise Point Positioning using IGS orbits and clock products". GPS Solutions 5(2):12–28, 2001.

[8] N. Guyennon, G. Cerretto, P. Tavella, F. Lahaye, "Further characterization of the time transfer capabilities of precise point positioning (PPP): the Sliding Batch Procedure". IEEE Trans Ultrason Ferroelectr Freq Control, 56(8), pp. 1634-41, 2009.

[9] P. Defraigne, N. Guyennon, C. Bruyninx, "PPP and Phase-Only GPS Frequency transfer", EFTF 2007 proceedings, pp. 904-909.

[10] F. Lahaye, P. Collins, "Advances in time and frequency transfer from dual-frequency GPS pseudorange and carrier-phase observations", Proc. 40th PTTI, 2008.

[11] J. Delporte, F. Mercier, D. Laurichesse, O. Galy, “Fixing integer ambiguities for GPS carrier phase time transfer,” in Proceedings of the IEEE International Frequency Control Symposium (IFCS ’07), pp. 927–932, Geneva, Switzerland,May 2007.

[12] G. Petit, Z. Jiang, "Precise point positioning for TAI computation", Int. J. Navig. Obs., ID 562878, 8 pages, 2008.

[13] F. Roosbeek, P. Defraigne, C. Bruyninx, "Time transfer experiments using Glonass P-code measurements from RINEX files", GPS Solutions, 5(2), 51-62, 2001.

[14] J. Nawrocki, W. Lewandowski, P. Nogas, et al., "An Experiment of GPS+GLONASS Common-View Time Transfer Using new Multi-System Receivers", Proc. Europ. Freq. and Time Forum, 2006, 904-909, 2006.

[15] P. Defraigne, Q. Baire, "Combining GPS and GLONASS for Time and Frequency Transfer", J. Adv. Space Res., 2011 (in press)

[16] P. Defraigne, A. Harmegnies, Q. Baire, "Time and frequency transfer combining glonass and gps data", proc. PTTI 2010. (CD-rom)

[17] K. Senior, P. Koppang, J. Ray, "Developing an IGS time scale", IEEE Trans. Ultrason. Ferroelectr. Freq. Control, 50, 585-593, 2003.

[18] T.A. Springer, "NAPEOS Mathematical Models and Algorithms", DOPS-SYS-TN-0100-OPS-GN, 1.0, Nov 5, 2009.

[19] Baire, Q., Defraigne, P., Pottiaux, E., Influence of Troposphere in PPP Time Transfer, Proc. Europ. Freq. Time Forum 2009, 1065 - 1068, 2009.

Time-Code Assisted Low-Power GNSS Single-Shot Receiver for Mobile Devices

Christoph Kandziora and Robert Weigel

Institute for Electronics Engineering, University of Erlangen-Nuremberg,
Cauerstr. 9, 91058 Erlangen, Germany, {kandziora,weigel}@lte.eei.uni-erlangen.de

***Abstract*— The scope of this paper is a new approach to assisted navigation using both signals from a global navigation satellite system (GNSS) and a long-range time code transmitter. The information gathered from the time code signal gives a priori knowledge of signal reception time in the sub-microsecond range. Together with the approximate position of the device and assistance navigation data, both derived from a cellular network, a pre-estimation of code-phases is possible. This leads to a significant reduction in baseband processing effort for the computation of the real code phases for the navigation solution.**

The structure of the time code signal gives the opportunity to reuse the core of the baseband signal processing of the GNSS single-shot receiver with only minor modifications to determine the reception time. Therefore the number of cross correlation channels in the baseband processing hardware can be reduced by at least a factor of two.

I. Introduction

The market for mobile devices is increasingly dominated by smartphones, which incorporate a lot of different functions such as camera, FM-radio receiver or GNSS receivers. Therefore each of these functions has to be optimized regarding power consumption and speed. Looking at GNSS receivers, the introduction of the assisted GNSS principle provides a major advantage in the performance of the GNSS receivers. These concepts strongly rely on a cellular network connection which feed the GNSS receiver with the appropriate information about the satellite constellation and errors in the navigation solution due to atmospheric effects, hence the navigation message. The baseband signal processing can then be arranged in a massive parallel structure to scan the whole code-phase-Doppler search space for all satellites in view in parallel. This leads to a very small time-to-first-fix (TTFF). The main disadvantage of this concept is that the massive parallel structure consumes a lot of hardware.

The proposed concept overcomes the problem of extensive hardware usage by adding information about the reception time of the satellite signal to the navigation solution. The time information is gathered by a time code receiver for the German DCF77 transmitter. Compared to a common Single-Shot receiver concept the architecture has to be expanded by only an analog RF front-end for the reception of the DCF77 signal. Since the DCF77 transmitter emits a 512 bit long PRN sequence modulated on a 77.5 kHz sine wave carrier, the baseband signal processing hardware of the common single-shot receiver can be reused with minor changes to the concept but major changes to the amount of applied cross correlation channels.

II. DCF77 Signal in Space

The DCF77 transmitter is located near Frankfurt, Germany, and emits a 77.5 kHz sine wave which is amplitude modulated in the way that at the beginning of each second the amplitude is reduced by 75% for 100 ms or 200 ms whether a '0' or '1' is transmitted. To set the minute mark the amplitude at the beginning of the 59th second is not reduced but held at 100% level [1]. Additionally the phase of the sine wave is modulated with a 512-bit long PRN-sequence in the segment of the second when the amplitude is at 100% level, hence from 200 ms after the second mark until the next second mark [2]. The PRN sequence is generated from a 9 stage linear feedback shift register which is reset every second. The generation of the DCF77 PRN sequence follows equation(1):

$$y = x^9 + x^5 \tag{1}$$

The PRN sequence also carriers the information encoded in the amplitude modulation [2]. Dependent on whether the PRN sequence is transmitted directly or with inverted phase a binary '0' or '1' is transmitted. Fig.1 shows the behavior of the DCF77 signal.

III. RF Front-end

The demands to the GNSS RF front-end are similar to those of common GNSS receivers. Therefore its characteristics will not be described here. For further information the reader is refered to the literature [3] [4] [5].

The demands to the RF front-end for the presented architecture of the time-code receiver are on the one hand an extremly narrow band receiver design (around 10-20 Hz) for the generation of a reference frequency and a design

978-1-61284-110-6/11/$26.00 ©2011 IEEE

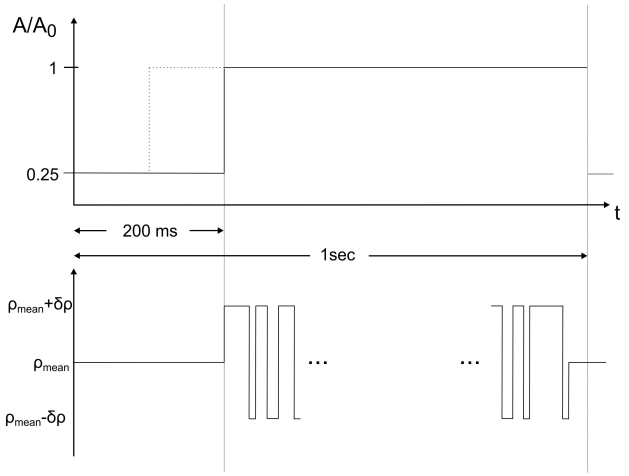


Fig. 1. Behavior of DCF77 signal in amplitude and phase. [2]

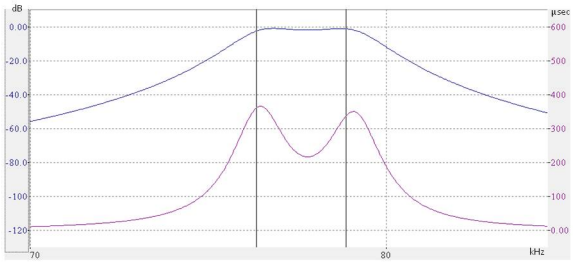


Fig. 3. Attenuation and group delay characteristic of band pass filter

with a wider bandwidth of around 2.6 kHz in order to allow demodulation of the information.

These two, at first glance, contrary demands can be overcome by the architecture shown in Fig. 2.

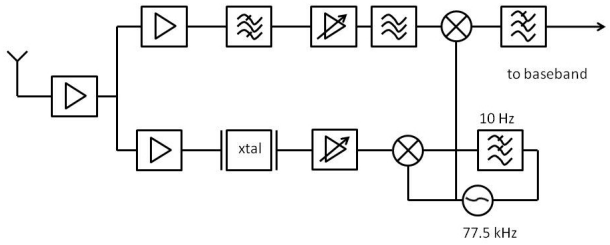


Fig. 2. Block diagram of the RF front-end architecture

Similar to the well know receiver chipsets, the proposed architecture uses a ferrite antenna to receive the DCF77 signal. After a first amplification stage, which again also decouples the antenna from the rest of the circuit, the signal is fed into two parallel branches of the circuit. The lower branch in Fig.2 filters the signal, after further amplification, by a high-Q crystal filter. The bandwidth of this filter is around 10-20 Hz. This eliminates all kind of distortions and even affects the rising edge of the amplitude modulation. After filtering the signal is again amplified and then fed as reference frequency into a PLL. The VCO output signal is by that tightly coupled to the received highly precise 77.5 kHz sine wave and therfore serves as the demodulating reference frequency.

The second branch of the RF front-end amplifies the signal, too, and filters it with a bandpass filter with a 3 dB-bandwidth of around 2.6 kHz. The filter is realized by a four-pole multiple-feedback (MFB) configuration. A screenshot of the simulated filter characteristic regarding attenuation and group delay is shown in Fig 3. After filtering the signal is again amplified by an automatic gain control (AGC) amplifier to compensate losses from the filter stage and prevent the signal from clipping.

The next block in the demodulating branch of the RF front-end represents an allpass filter which is used to compensate differences in the signal delays of the two branches. The phase rotation introduced by this allpass filter is necessary to generate a DC-offset free output of the following phase detector (PD). This is important for proper operation of the digital baseband processing engine. The phase rotation introduced by the allpass filter is compensated in the digital baseband.

The following phase detector is fed by the output of the allpass and the frequency generated in the other branch of the circuit. The phase difference at the inputs of the phase detector is constant and can easily be measured to calulate the characteristic of the precedent allpass.

The demodulated signal is then digitized by a 2 bit flash ADC for further signal processing.

IV. Digital Baseband

The digital baseband processing engine is based on that of the common single-shot receiver. It is used for the computation of the cross correlation function between the received time-code or the GNSS satellite signals and their local replica. As mentioned before both are a pseudo-random-noise (PRN) codes of different length. All possible code-phases for one signal are processed in a massive parallel approach. The structure of one cross correlation channel is shown in Fig. 4. The correlation function is performed in two steps. The first step is a coherent correlation which must be performed for at least one complete code epoch, namely 1 s for the time code signal or 1 ms for GPS signals.

Depending on the received power level and distortions it can be necessary to correlate over more than one epoch to determine the actual code phase of the received signal. Since the actual data bit of the time or navigation message are coded into the PRN sequence by transmitting either non-inverted

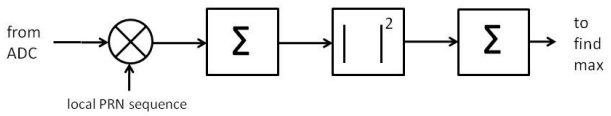


Fig. 4. Block diagram of baseband processing channel.

or inverted PRN-sequences [6] coherent integration lead to false results if correlated over data bit boundaries [7]. That is why it is essential to perform the subsequent non-coherent integration.

The proposed architecture, incorporating 1024 channels shown in Fig. 4, is shown in Fig. 5. The channels are arranged in parallel each fed with the same digitized demodulated signal (either time-code or GNSS signals) and the local generated replica of the appropriate PRN sequence. The PRN replica is half chip-wise shifted regarding adjacent channels. The position of the maximum result of the parallel cross correlation functions correspond to either the actual error of the local RTC with respect to UTC or the actual code phase for the actual space vehicle (SV).

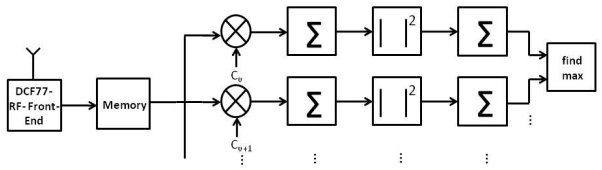


Fig. 5. Block diagram of baseband processing engine.

Regarding the result for corellating the time code signal, the resulting clock offset has to be corrected from two main errors. First of all the offset introduced by the allpass filter in the RF front-end. This error is an offset which is constant for a particular receiver and varies from receiver to receiver due to component tolerances. As already mentioned it can be measured once for each receiver and then be hard-coded in the baseband processing engine or the software of the receiver, e.g. by an initial shift of the local PRN replica. Second the distance between the transmitter and the receiver has to be taken into account.

V. Positioning Algorithm

The algorithm to determine the actual user position is controlled by a host platform based on an OMAP3530 embbeded processor running Google Android.

The first step is to collect the GNSS navigation message over a cellular network connection. As already mentioned this connection is essential for the single-shot operation mode of the proposed architecture, due to the fact that the navigation message does not have to be extracted from the satellite signal. Together with an approximate position and time information which again can be derived from the cellular network two essential pieces of information can be extracted. First of all the time of flight of the time-code signal is now known. The error because of the error in the inital position approximation is negligible. The host then initiates the reception of both the time-code and the GNSS signals. The received samples are stored in memory for offline processing and both front-ends can be shut down for power saving.

The host then instructs the baseband processing engine to cross correlate the DCF77 signal. The code-phase computed directly relates to the offset between the local RTC and UTC. With the known reception time the actual satellite constellation can be computed and the SVs with the best GDOP can be selected.

Together with their position at the calculated time and the estimated position of the user an estimation of the code phases for the selected SVs can be calculated by the baseband processing engine either. Since the error of the inital user position approximation and the error in the reception time calculated are together smaller than 500 μs the search window for the real code phase for the appropriate SV can be cut down to 512 chips. Therefore again the 1024 cross correlation channels available in the baseband processing engine are sufficent to calculate the real code phase.

After four code phases are computed the well known algorithms [8] for calculating the position from code-phases can be applied.

VI. Conclusion

A new low-power approach to assisted navigation is proposed. Incorporating a long-range time-code receiver gives a priori knowledge of the reception time of the satellite signal. With the already know constellation a estimation of the code phases for the appropriate SVs at the estimated position of the user can be calculated. Since the resulting search window is much smaller compared to the search window for common receiver architectures the 1024 parallel cross-correlation channels for the determination of the reception time are sufficient even for the determination of the code phases of the GNSS satellite signals. This directly correlates to a cut in hardware effort by a factor of two regarding GPS satellite signals and a factor of 8 for Galileo satellite signals. Since active hardware strongly correlates to power consumption the cut in active hardware prooves the low-power approach mentioned at the beginning.

REFERENCES

[1] D. Piester, P. Hetzel, and A. Bauch, "Zeit- und Normalfrequenzverbreitung mit DCF77," Physikalisch Technische Bundesanstalt, Tech. Rep., 2004.

[2] P. Hetzel, "Zeitübertragung auf Langwelle durch amplitudenmodulierte Zeitsignale und pseudozufällige Umtastung der Trägerphase," Ph.D. dissertation, Institut für Zeitmeßtechnik, Fein- und Mikrotechnik der Universität Stuttgart, 1987.

[3] *MAX2769 Datasheet*, MAXIM Inc.

[4] A. Miskiewicz, "Integration of Satellite Navigation RF Receiver Paths into Mobile Radio Platforms," Ph.D. dissertation, Lehrstuhl für Technische Elektronik der Universität Erlangen-Nürnberg, 2009.

[5] H. Ehm, "Single Shot Radio Receiver Architectures for Mobile Stations," Ph.D. dissertation, Lehrstuhl für Technische Elektronik der Universität Erlangen-Nürnberg, 2008.

[6] *NAVSTAR GPS User Equipment Introduction*, U.S. Department of Homeland Security, September 1996.

[7] A. Schmidt, A. Neubauer, H. Ehm, R. Weigel, N. Lemke, G. Heinrichs, J. Winkel, J. A. Avila-Rodriguez, R. Kaniuth, T. Pany, B. Eisfeller, G. Rohmer, B. Niemann, and M. Overbeck, "Enabling location based servies with a combined Galileo/GPS receiver architecture," in *Proceedings of ION GNSS*, 2004.

[8] E. D. Kaplan and C. J. Hegarty, Eds., *Understanding GPS :principles and applications*. Artech House, Norwood, MA, 2006.

Towards an All-digital Time Scale

Stefania Römisch, Steven R. Jefferts, Thomas E. Parker
Time and Frequency Division
NIST
Boulder, CO USA
romisch@boulder.nist.gov

***Abstract* — The implementation of an all-digital time scale is under way at NIST, by use of a novel 8-channel, all-digital phase measurement system based on subsampling techniques. The phase measurement system is used to compare output signals from several commercial atomic frequency standards; the phase differences between these signals, at different measurement times, will then be input to the algorithm used to generate the digital time scale. The subsampling technique allows the elimination of analog mixers in the system's front end, with their noise contributions, and yields performance that is comparable to or better than the present state of the art. Long-term (up to 80 days) comparisons of both common-clock performance and fractional frequency measurements between the latest generation of the digital measurement system and the commercially available system used in the implementation of UTC(NIST) are presented.**

I. Introduction

The implementation of a time scale requires the ability to compare the time evolution of several clock signals. The instantaneous absolute phase of each clock is expressed in seconds to account for the accumulation of more than one cycle (2π radians) of phase difference between two signals. The quantity used in the computation of time scales is still commonly called *phase* but is in fact indicated with x and defined according to (1) below:

$$x(t)=\frac{\varphi(t)}{2\pi\nu_0}, \tag{1}$$

where $\varphi(t)$ is the instantaneous phase of the signal in radians and ν_0 is the nominal frequency of the same signal.

The differences between the instantaneous phase of n different clocks with respect to a clock r, elected to be the reference for the time scale, are indicated as $x_i(t)$-$x_r(t)$ and are then used by a weighting algorithm that computes the instantaneous phase of a virtual clock, the *ensemble clock*.

The ensemble clock, calculated by use of the algorithm *at1e* at NIST, is the representation of a local time scale, $TA(t)$. At NIST there are two real-time time scales, called AT1 and TSC, computed separately (but using the same algorithm) from distinct measurements of the same set of commercial atomic standards comprising both hydrogen masers and cesium thermal-beam clocks.

A third time scale, called DIG, will be computed, using the same algorithm, from the phase measurements on the same set of commercial atomic standards as executed by the novel, all-digital measurement system developed at NIST in the past few years.

The addition of a third time scale will provide increased redundancy for the entire time-scale operation, and the fundamentally different technology at the core of the digital measurement system (with respect to the ones presently used to provide the data for AT1 and TSC) will serve as a discriminating tool for the occasional aberrant behavior of the measurement systems. In fact, different technologies usually have different shortcomings that are often triggered by different occurrences.

II. The Measurement of Phase Differences

The high stability of the clocks involved in time-scale operations requires high-resolution measurement systems, making it impossible to use currently available time-interval or frequency counters without the use of sensitivity-enhancing techniques. The traditional and our all-digital sensitivity-enhancing techniques are briefly described below.

A. The dual-mixer technique

The technique traditionally used to compare the phase of two (or more) atomic frequency standards is referred to as dual-mixer measurement technique [1], [2] and is illustrated in its basic conception in Fig. 1.

The mixers in Fig. 1(a) simply down-convert each of the clocks' signals (generally at 5 MHz) to a much lower frequency (generally at 10 Hz) to increase the effective resolution of the time-interval counter. In fact, the phase difference in radians between the two 5 MHz waveforms amounts to the same number of radians for the 10 Hz waveforms, resulting in a time difference which is $5\cdot10^5$ times larger. A time-interval counter with a resolution of 10 ps used in this manner will therefore yield an effective resolution of about $2\pi\cdot10^{-9}$ rad for the phase measurement of the 5 MHz signal.

To first order the quality of the synthesized intermediate frequency (4.999990 MHz in this case) is irrelevant, because its noise contributions are common to both down-converting

processes and therefore cancel out, resulting in easier implementation of this measurement scheme.

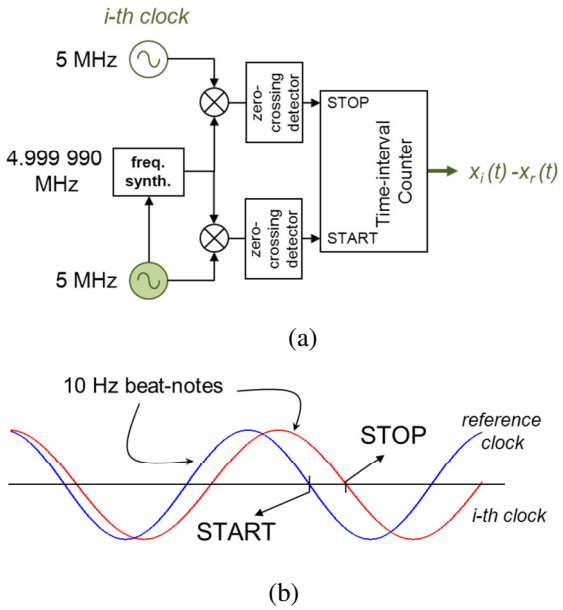


Figure 1. The dual-mixer technique for measuring phase differences between two signals. (a) The down-converting mixers produce two beat-note signals used to measure the time (phase) difference. (b) The two down-converted signals and their relation with the start and stop of the time-interval measurement.

The basic concept depicted in Fig. 1 can be expanded to simultaneously measure phase differences between *n* clocks and a common reference.

B. The All-digital technique

The all-digital technique for measuring phase differences between clocks can be summarized in one basic concept: the down-conversion performed by the analog mixers in Fig. 1 is replaced by digital sub-sampling of the clock signals [3]. The result of this operation is a digitized signal that is fitted by a sinusoidal waveform at a much lower frequency than that of the original clock signal.

The sub-sampling operation, like frequency mixing, preserves the signal's phase at the beginning of each measurement instance, allowing the measurement of the phase of a 5 MHz clock signal with a sensitivity gain similar to the one described for the dual-mixer technique.

In Fig. 2 is shown a graphical representation (not to scale) of the sub-sampling process used in the all-digital phase measurement system. The nominal 5 MHz clock signal is sampled at the rate set by the sampling time base square-wave for the duration of the measurement window, providing a set of samples that will be fitted by a sinusoidal waveform that is at a much lower frequency than both the sampling rate and the clock's frequency.

The frequency of the "beat-note" is determined by the residual of the ratio between the clock's frequency (5 MHz) and the sampling rate, which in fact must not be an integer fraction of 5 MHz.

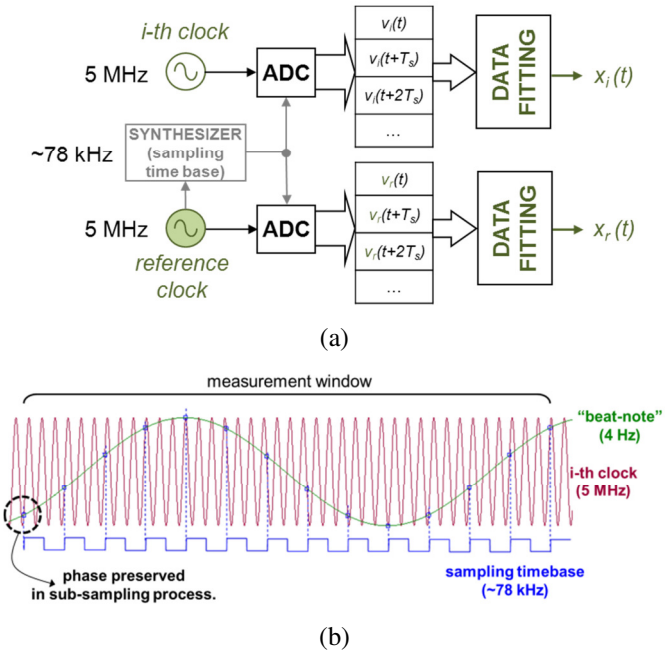


Figure 2. (a) Phase measurements using the all-digital technique. A synthesized sampling time base drives the subsampling process and the resulting samples are then fitted by a sinuoisdal waveform. The initial phase of this waveform is the final result of the measurement. (b) Not-to-scale representation of the sub-sampling process. The 5 MHz signal (in red) is sampled at the rate set by the sampling timebase square-wave (which must not be an integer fraction of the nominal 5 MHz) producing a set of samples indicated by the blue squares. These samples are then fitted by a sinusoidal waveform whose frequency is much lower that 5 MHz. The phase of the 5 MHz signal at the beginning of the measurement window is preserved by the sub-sampling process.

The length of the measurement window is a free parameter that can be chosen together with the sampling rate to determine the number of samples acquired for each measurement instance. The sampling rate and the length of the measurement window are parameters that should be chosen with some care. In fact, if a higher number of samples available for the "beat-note" fitting generally means a higher accuracy in the waveform's initial phase, a longer measurement window may result in a set of samples whose statistical properties are no longer stationary, introducing a bias in the determination of the waveform initial phase.

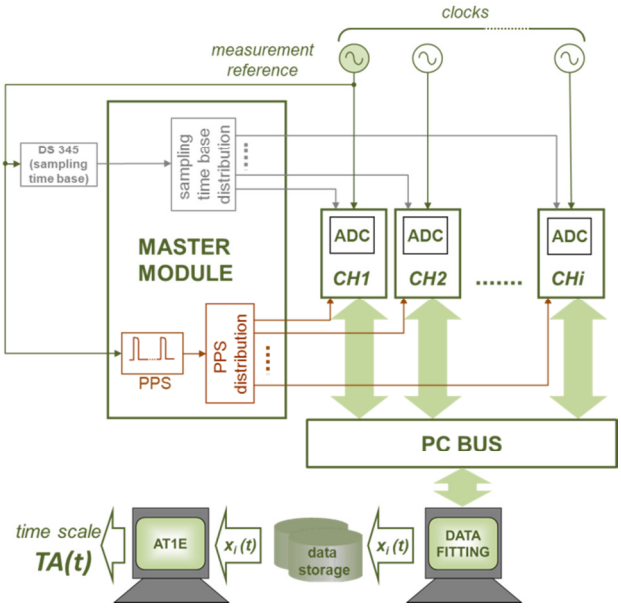
III. THE ALL-DIGITAL MEASUREMENS SYSTEM

The all-digital measurement system presently implemented at NIST is a re-engineered version (V1.5) of the prototype (V1) described in [3] and it can be seen in the photograph shown in Fig. 3(a).

The three main system components are a Master Module (top enclosure), a PC Bus (second enclosure from the top) and the crate housing the six active acquisition modules. The same three components are also indicated in the functional block diagram in Fig. 3(b). One of the clocks under measurement is used to synthesize both the sampling time-base and the pulse-per-second (PPS) signal used to trigger the measurement instances of the system, making that clock the reference for the measurement system. The sampling time-base is not an integer fraction of the nominal 5 MHz frequency of all clocks'

signals, so as an intermediate step towards full Master Module implementation with a Digital Direct Synthesizer (DDS) a commercial synthesizer was used. Both the sampling time-base and the PPS signals are distributed by the Master Module to all acquisition modules indicated as CH1, CH2 and CHi in the diagram.

(a)



(b)

Figure 3. (a) Photograph showing the re-engineered measurement system V1.5. The top enclosure is the Master Module, the second enclosure from the top is the PC Bus, and the crate below them houses the six active acquisition modules. (b) Functional block diagram of the same sytem. The sampling time base and the PPS signal are synthesized from one of the available clocks, which is then the reference clock for the measurement system.

The PC Bus allows the PC dedicated to data fitting to poll in turn each acquisition module, reading the acquired samples from the flash memory chip where they were stored by the microcontroller managing the acquisition process. At the end of the fitting process the phase of each 5 MHz signal at the beginning of each measurement instance, $x_i(t)$, is stored to be processed at a later time by the time-scale algorithm of choice.

Because any time-scale algorithm is substantially a prediction algorithm for clock *differences*, one of the ensemble's clocks is chosen to be the reference clock for the time-scale calculations. This reference clock must not be confused with the clock that is used as the *measurement reference* as indicated in fig. 3(b): while the two may be chosen to be the same physical clock, their functions are separate and different. Although some measurement systems may not allow a complete freedom of choice in this regard, and despite the setup chosen for the measurements shown in this paper, the conceptually optimal choice should use UTC(NIST) as the *measurement* reference clock. The effect of using, instead, one of the clocks as the *measurement* reference has the effect of introducing a skew in the measured data, due to the drift of that clock with respect to the ensemble clock representing the local UTC. As it turns out, the effect of such drift is of no consequence, as discussed in the next section.

IV. System Characterization

The names of the frequency sources (clocks) used in the measurements described in this paper are summarized in Table I, where the channel numbers are indicated for both all-digital (DIG) and dual-mixer-based (AT1, TSC) measurement systems.

TABLE I. Summary of Measured Clocks

DIG Channel number	AT1 Channel number	TSC Channel number	Clock name and description
1	30	30	ST004 – Hydrogen maser
2	5	5, 25, 26	ST003 – Hydrogen maser
3	7	7	AOG03 – Auxiliary Offset Generator (ST006 steered by UTC(NIST))
4, 5	13	13	ST0022 – Hydrogen maser (reference for sampling time-base, PPS)
7	18	18	ST006 – Hydrogen maser
-	17, 24	17, 24	HP1074 –Cs thermal-beam clock

There are several ways to characterize this measurement system, each one of them providing different and complementary information about its overall performance in the domain of time-scale applications.

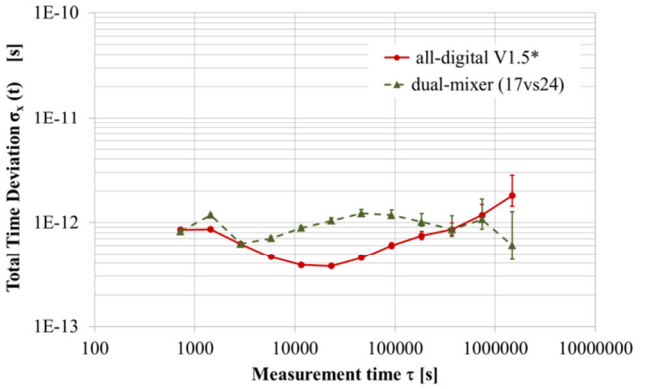
A. Common-clock Measurements

In a common-clock measurement, two distinct channels (acquisition modules) measure the same clock, thereby offering as a result the residual noise of each channel, summed in quadrature. Presuming the two measuring channels are virtually identical, the resulting "noise floor" for each of them is assumed to have a total time deviation (or a total fractional frequency deviation) that is 0.7 times better than the one computed from the measured noise levels.

Several results for this configuration are shown and compared for different measurement systems in Fig. 4, where the quantity depicted is the total time deviation (total TDEV)

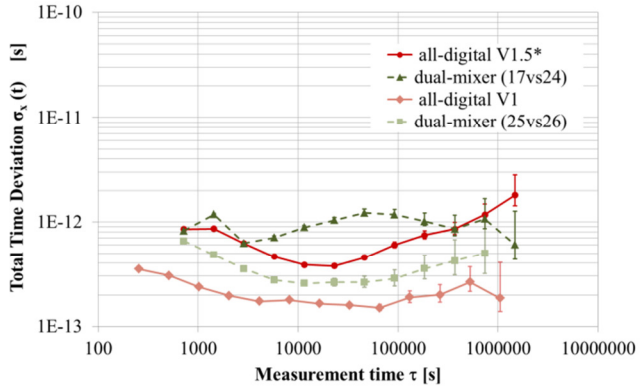
computed from the measured noise, without the 0.7 factor reduction.

Fig. 4(a) shows the time instability of channels 4 and 5 for the all-digital measurement system (V1.5), as compared with the results of similar measurement for channels 17 and 24 of the dual-mixer-based system that is the core of UTC(NIST). The all-digital system exibits a time instability below 1 ps for measurement periods up to approximately 10 days, and it is comparable with, or better than, the dual-mixer system. The measurements were conducted between MJD 55560 and 55680.



* Glitches removed are > 5 s and amount to less than 3·10-3 of total data points

(a)



* Glitches removed are > 5 s and amount to less than 3•10-3 of total data points

(b)

Figure 4. (a) Total time deviation calculated for channels 4 and 5 for the all-digital measurement system (V1.5), as compared with the results of similar measurement for channels 17 and 24 of the dual-mixer-based system that is the core of UTC(NIST). (b) Total time deviation calculated for common clock measurements using the prototype all-digital system prior to its re-engineering and a different pair of channels of the dual-mixer system, compared with the data in (a).

In Fig. 4(b) the same data are shown together with the results from common clock measurements using the prototype all-digital system (V1) prior to its re-engineering. Data from a better-performing pair of channels for the dual-mixer measurement system (for the same period spanning approximately MJD 55350-55400) are also shown.

The discrepancy between the two sets of measurements illustrates the dual-mixer system channel-to-channel performance variability and the unwanted effect of the re-engineering process of the all-digital system, which significantly raised its noise floor, albeit without invalidating its viability for time-scale applications. In fact, even the worse performance is still within the spread of performance of the dual-mixer system generating UTC(NIST).

Finally, a comparison of the projected effect of using a specific clock as measurement reference clock, instead of UTC(NIST), is compared with the noise floor of the system and determined to be inconsequential, at least for the measurement periods described in this paper. In fact, if the measured drift of ST006 (used as the measurement reference clock) is approximately $-1.27 \cdot 10^{-21}$, after 100 days of measurements the 720 s measurement interval will have an error of approximately 10 fs. This would be the timing error of the measurement system if it were otherwise noiseless: in reality it is well below even the best noise floor of the all-digital measurement system. The measurement reference clock drift with respect to UTC(NIST) is a negligible contribution to the all-digital system for measurement periods up to 2000 days.

B. *Frequency Offset and Double-difference Measurements*

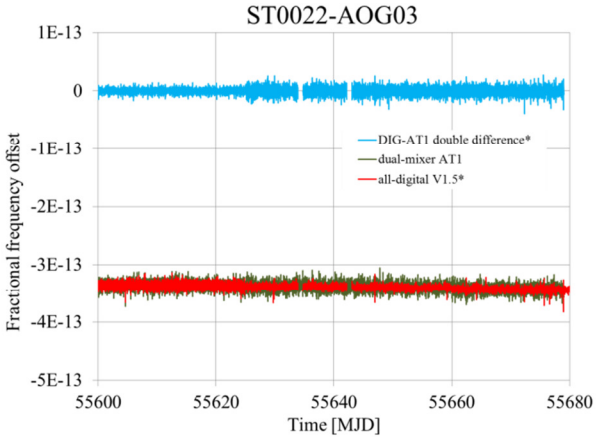
Although the common-clock measurements are very important in determining the capabilities of the measurement system, its typical operating conditions require the simultaneous measurement of *n* different clocks' phases.

Apart from channels 4 and 5, all other active channels in the all-digital measurement system are measuring different clocks (see Table I). The measurement of the absolute frequency of a clock happens only through its comparison with a primary frequency standard. The clocks that constitute a time scale are all secondary standards, so any time scale algorithm is essentially a predictor for clock *differences*. In Fig. 5, therefore, is shown one example of the frequency offset (difference) between maser ST0022 and the output of the AOG03, as measured by the all-digital measurement system and the dual-mixer-based one, AT1 in this case. The blue curve (around zero) is the difference between the same frequency offset, as measured by the two systems: it represents the combined residual noise of the two measurement systems and is obtained by calculating the difference between the red and the green curve of Fig. 5:

$$\text{double diff.} = (\text{ST0022-AOG03})_{\text{DIG}} - (\text{ST0022-AOG03})_{\text{AT1}}$$

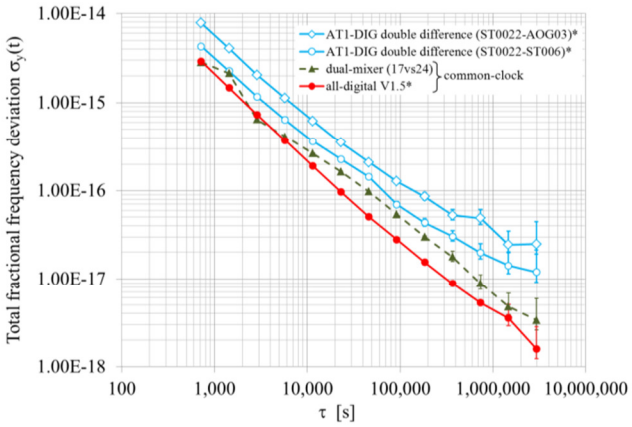
The data in Fig. 5 are representative of the frequency offsets measured between all other clocks listed in Table I: they confirm the ability of the measurement system to accurately measure phase differences evolving with time.

An evaluation of this accuracy is shown by the total fractional frequency deviations shown in Fig. 6, where the combined residual noise of both measurement systems is compared with the common-clock measurements for each one of the measurement systems. The total fractional frequency deviation for the common-clock measurements is computed from the same sets of data used to compute the TDEV shown in Fig. 4.



* Glitches removed are > 5 s and amount to less than 3•10-3 of total data points

Figure 5. Measurement results for the frequency offset (difference) between maser ST0022 and AOG03, as measured by the all-digital system (red curve) and the dual-mixer-based one (green curve). The blue curve is the difference between the red and the green curves, and it represents the combined residual noise of the two measurement systems.



* Glitches removed are > 5 s and amount to less than 3•10-3 of total data points

Figure 6. Total fractional frequency deviation computed for the common-clock measurements (same data sets used for the TDEV shown in Fig. 4, red and green curves) and for the combined residual noise of the two measuremetn systems, for two different frequency offsets: ST0022-AOG03 and ST0022-ST006. The combined residual noise is obtained by subtracting the frequency offset of each clock pair as measured by each measurement system.

In particular, two double differences are shown in Fig. 6, for two different pairs of clocks: (ST0022-AOG03) and (ST0022-ST006). The residual noise calculated through double differences is expected to be $\sqrt{2}$ above the noise of a common-clock measurement, because the noise of each pair of channels is summed in quadrature. The differences between the two double difference curves in Fig. 6 (empty circles and empty diamonds) is due to the variability between the each channel's noise floor, as shown by the difference between the results of common-clock measurements between different pairs of channels of the same dual-mixer-based measurement system.

The interesting feature of Fig. 6 is in the flicker-type behavior shown by both the double difference curves, but not by the curves based on common-clock measurements. At this time it is not clear what the cause of it is, and it is currently under investigation.

V. Conclusions and Future Work

The performance of the re-engineered, multi-channel, all-digital measurement system presently active at NIST is presented, and compared, for both common-clock performance and operational (difference clocks) performance with the dual-mixer-based system that is at the core of UTC(NIST).

Apart from validating the all-digital system as fit to be used as part of an operational time scale, the comparison between the measurements obtained from both systems offers the possibility to investigate long-term behaviors not seen previously, as shown in Fig. 6.

Future work includes reducing the residual noise for the all-digital system in order to reach the levels warranted by the prototype V1, implementing ancillary systems allowing for automated processing of the data into a time-scale algorithm, and investigating the long-term trends of both measurement systems.

Aknowledgemnts

The authors thank Judah Levine for extremely helpful discussions, regarding both the system's design and the subsequent data analysis.

References

[1] D. Allan, H. Daams, "Picosecond time difference measurement system," Proceedings of the 1975 International Frequency Control Symposium, pp. 404-411.

[2] S. Stein, et al., "Automated high accuracy phase measurement system," IEEE Transactions on Instrumentation and Measurement, Vol. 32, pp. 227-231, 1983.

[3] S. Römisch, T.E. Parker, S.R. Jefferts, "Novel, all-digital phase measurement system for time scales," Proc. of the 41st PTTI Meeting, pp. 397-408, 2009.

The Method of Time Synchronization Based on the Combination of COMPASS GEO Pseudo-range and Two-way data

Xuhai Yang[1], Langming Ma[1, 2], Baoqi Sun[1], Hui Lei[1], Le Sun[1, 2,] Zhigang Li[1], Wei Zhou[3]

[1]National Time Service Center, Chinese Academy of Sciences, Xi'an, 710600, China

2Graduate University, Chinese Academy of Sciences Beijing, 100049, China

3Xidian University, Xi'an, 710071, China

Email: yangxh@ntsc.ac.cn malangming@163.com

***Abstract*—This paper presents a time synchronization method based on the COMPASS code measurements in combination with two-way transfer ranging, which results the clock offset between COMPASS GEO satellite and the ground station. The clock offset of COMPASS G3 satellite on board is computed by using the data observed on October 2010. In order to verify the accuracy of this method, the results are compared with the results, which are determined only by pseudo-range observation. By comparison and analysis, the following facts can be drawn. ①The precision of navigation satellite clock bias is better than 10 ns, when the clock bias is calculated by the new method .There is no systematic error between our results and clock bias deduced from precise satellite orbit. ②The new method can separate satellite orbit and clock bias, it is beneficial to researching satellite clock bias.**

I. INTRODUCTION

Time synchronization is a basic key technology and performance of a satellite navigation system, directly affecting the navigation, positioning and timing accuracy. GPS system in order to improve accuracy for time synchronization, the first request the satellite clocks are atomic clocks with high stability, also require the satellite and ground system clock on a regular basis compared with the standard clock to give the clock parameters of the satellites clocks than the relative standard clock .In GPS control and operation, or IGS system, the satellite clock correction determined using the method "orbital inversion", this method is applied to GPS, it works well. Characteristics of this method is dependent on the satellite orbit, there will be a problem in COMPASS geostationary satellites. The high-level precise orbit determination for geostationary satellites is difficult to achieve, because the tracking geometry for GEO is almost unchanged due to the geostationary feature of the GEO satellite, resulting in weak dynamical constraints.GEO satellites are often in the orbit maneuver, this time it also will be difficulties in determining the orbit of the satellite. So clock correction determined based on the satellite orbit will be a problem. This paper presents a time synchronization method with the combination pseudo-range and two-way transfer ranging data, free from influence of the orbital maneuvering to determine the satellite clock correction.

Two conventional GEO tracking techniques are the pseudo-ranging and the transfer ranging. The L-band downlink navigation signal is received by the receiver and the pseudo-range is observed with clock corrections of both the satellite and the receiver. The transfer ranging is a two-way ranging mode. The C-band ranging signal from a ground station is transmitted by a transfer onboard GEO satellites designed for communication, and is also received by the station, without clock corrections in the ranging data.

This paper proposes a new method which directly combines the two-way transfer ranging data and code pseudo-range data to determine clock correction of the satellite and the ground station. This combined solution haves the advantages, i.e. a high accuracy clock correction free from satellite orbit. Section 2 of this paper will describe the algorithm and methodology used for the combination. In section 3, this direct method will be applied to COMPASS G3 in order to quantify its efficiency, advantages and limitations.

II. COMBINATION STRATEGY

The C-band transfer ranging is a two-way ranging mode. Ranging signals are emitted by a transfer tracking station, transmitted by the transponders onboard a GEO satellite and the transmitted signals are received by the original tracking station to realize the distance measurement between GEO satellite and tracking station. Ignoring the delay correction, can be written as

$$\rho = \frac{\tau}{2} c$$

where c is the speed of light.

For the L-band pseudo-range, the signal propagation time between a satellite and a receiver is determined by correlation processing between the local receiver code and the pseudo-range code of the satellite. To accurately determine the distance from the satellite to station, the satellite clock and the receiver clock must maintain a strict synchronization, which is difficult to achieve in practice. As the satellite clock and receiver clock are not synchronized, so the pseudo-range is observed with clock corrections of both the satellite and the receiver, which is equal to the measured distance from the satellite to the receiver plus two clocks correction, omitted delay correction, can be written as

$$\rho^{'} = \rho + c\left(\Delta t_r - \Delta t_S\right)$$

Where $\rho^{'}$ is the pseudo-range, Δt_r the receiver clock offset, Δt_s satellite clock offset.

Since the two-way transfer range directly determine the distance between the satellite and the station, and without clock error information. If the two-way transfer ranging tracking station and the COMPASS tracking receiver using the same clock, at the same time and the same site to tracking the same satellite to attain distances between satellite and stations. So

$$c\left(\Delta t_r - \Delta t_s\right) = \rho^{'} - \rho$$

It can be seen from the above equation as long as there are transfer ranging and pseudo-range at the same ground station, it can directly be access the clock correction of satellite and ground station.

Combine two-way transfer ranging with pseudo-range to achieve time synchronization must be processing on transfer ranging and pseudo-range data, including synchronized record timestamp of two observational data, error correction, and the antenna of the L-band receiver and C-band transfer ranging system convert to the same center corrected.

A. Error correction

The pseudo-range measurement model is

$$\rho^{'} = \rho + c\left(\Delta t_r - \Delta t_s\right) + \Delta^{ion} + \Delta^{trop} + \Delta^{sagnac} + \varepsilon$$

where Δt_r is the clock correction of a station and Δt_s is the clock correction of a satellite, Δ^{trop} is tropospheric delay, which has nothing to do with the signal frequency , and the values of troposheric delay of two-way ranging are canceled out each other. Δ^{ion} is ionospheric delay, and here the dual-frequency ionosphere-free combination of pseudo-range observations is used. Δ^{sagnac} is the sagnac effect correction，which is directly corrected by the formula from the reference[1]. ε is the observing noise.

The two-way transfer range measurement model is

$$\tau c = 2\rho + \Delta_{up}^{ion} + \Delta_{down}^{ion} + \Delta_{up}^{trop} + \Delta_{down}^{trop} + \left(I^e + I^r\right)c + \tau_s c + \varepsilon$$

where R_{up} is the uplink geometric distance between the satellite and the station, R_{down} is the downlink geometric distance between the satellite and the station, we suppose $R_{up} = R_{down} = \rho$, τ_s is the instrumental delay of the satellite transponder , I^e and I^r are the signal instrumental delays for transmitter and receiver, which can be determined by a special survey. Δ_{up}^{trop} , Δ_{down}^{trop} , Δ_{up}^{ion} and Δ_{down}^{ion} are the uplink and downlink tropospheric delay and ionospheric delay respectively, the ionospheric errors are corrected by using global precise ionosphere delay model provided by CODE. The tropospheric delay has nothing to do with the signal frequency, then uplink and downlink can be approximately considered equal， when combination with pseudo-range, the tropospheric delay can be cancelled out， and ε is the observing noise. The upward and downward sagnac effects for two-way transfer range are the same with opposite signs, so they can be canceled out each other.

B. Timestamp synchronization of observation records

The recorded timestamp of two-way transfer range is the time for UTC (converted into GPST) of ranging signal emitted from the MODEM, marked as t_{tw}^{0}, while the recorded timestamp of pseudo-range is the received time for BDT (also converted into GPST) of the signal by COMPASS receiver, marked as t_{sw} ,and such recorded timestamps of two observations must be translated into the same basis. Two-way ranging signal transmission time is converted to the signal reception time $t_{tw} = t_{tw} + \tau$, where τ includes instrumental delays, and the errors are negligible if the accuracy of time synchronization is 1 nanosecond. The time scale only effects the observing time. Typically, t_{tw} and t_{sw} are not synchronized. In order to solve the integration of two-way transfer range and pseudo-range, synchronization correction must be done. As a basis of t_{sw}, the two-way range is interpolated into t_{sw} to form virtual two-way range as a reference timestamp using t_{sw}, and two observing data in the same timestamp are attained, where a 6-order Lagrange interpolation formula is used.

C. Correction for centering between transfer ranging tracking station and COMPASS receiver

Although the geometry distance of the satellite and the ground transfer station ρ and pseudo-range only including clock corrections of both satellite and the receiver ρ' above has been calculated, transfer station and the receiver are not at the same absolute position on the ground. Correction for centering between transfer ranging tracking station and COMPASS receiver must be carried out. The formula to correct for centering is

$$\Delta\rho = \frac{\rho_x \Delta E + \rho_y \Delta N + \rho_z \Delta H}{\sqrt{\rho_x^2 + \rho_y^2 + \rho_z^2}}$$

Where (ρ_x, ρ_y, ρ_z) is the satellite position vector in the station's coordinate system, $\Delta E, \Delta N, \Delta H$ are ,respectively, relative to the center of the receiver east, north, and up deviation of the transfer station center.

In the calculation, the satellite coordinates calculated using broadcast ephemeris; $\Delta E, \Delta N, \Delta H$ can be calculated

$$\begin{bmatrix} \Delta E \\ \Delta N \\ \Delta H \end{bmatrix} = \begin{bmatrix} X_{compass} - X_{tw} \\ Y_{compass} - Y_{tw} \\ Z_{compass} - Z_{tw} \end{bmatrix}$$

where $(X_{compass}, Y_{compass}, Z_{compass}), (X_{tw}, Y_{tw}, Z_{tw})$ are the coordinates of COMPASS receiver and coordinates of transfer tracking station in the ITRF 2000 coordinate system, respectively. M is transformation matrix from ECEF to station coordinate system, the specific expression see reference [2].

Due to the geostationary feature of the GEO satellite, the change of correction for centering is up to 2 ns, broadcast ephemeris is not used to correct for centering, only elevation and the distance between transfer station and pseudo-range receiver are used to correct for centering. This helps separation of the satellite orbit and clock errors. From this perspective, the integration solution does not rely on the satellite orbit, and the error of the orbit of the satellite will not be introduced into the satellite clock error.

III. COMPUTATION AND ANALYSIS

A. Experimental device configuration

Relying on transfer ranging tracking network in National Time Service Center, Chinese Academy of Science, the domestic tracking network of the COMPASS satellite navigation system are established. In addition to each tracking station equipped with geodesic receiver, high-performance atomic clocks and a two-way satellite time transfer (TWSTFT) system are equipped. Test system configuration as shown in Figure 1.

In tracking stations, the output of cesium atomic clock, a 10MHz signal of high stability performance, is send to the receiver.

TAB.1 DIFFERENCES OF CLOCKS BETWEEN TWO SOLUTIONS

Date	Max/ns	Min/ns	Mean/ns	sigma/ns
278	9.9495	2.9926	-6.1905	3.7016
279	11.1708	-4.5301	4.0059	3.3288
280	12.1989	-4.5533	4.2118	3.1839

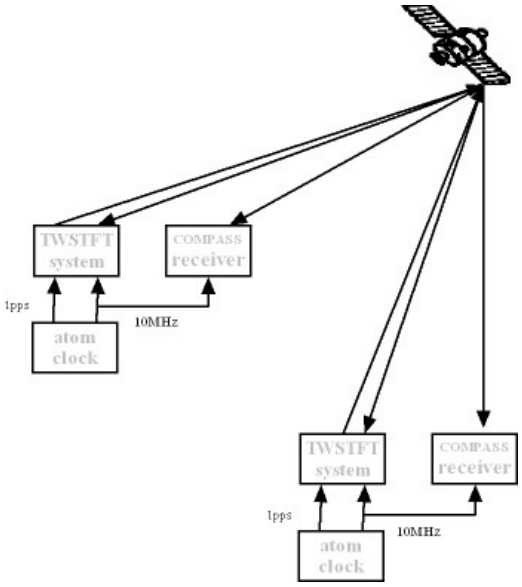


figure1. Test system configuration

B. *Experimental results and analysis*

The experimental data were from the Lintong tracking station from 2nd to 7th October, 2010. The clock error of satellite G3 and the Lintong ground station is calculated with the following two strategies. Solution ①: our new combined method; Solution②: deduced from the precise satellite orbit, which is determined by pseudo-range observations. At this time interval to evaluate the orbit accuracy based on the pseudo-range data, we used the SLR data of Changchun station to evaluate the orbit radial error. The maximum was less than 1m. So the error of clock bias between satellite and ground computed by solution ② is relatively small as a standard of comparison reference. Finally, the clock errors calculated by two solutions were directly compared.

During the test calculations, first the clock error was computed using solution①, then the clock error corresponding to timestamp was calculated by solution②, and the last two results were compared directly. Table 1 shows the comparison of the two results. Fig2-4 shows the results of October 5th -7th, the upper are the results using two solutions, the lower are the two solutions comparison residuals.

The compared results show that: from October 2nd to 4th, there are big systematic errors, taking into account that the compared results contain not only clock offset error of solution ①, but also the clock bias error of solution ②. We need further analyze whether the GEO satellite maneuvers occur these days or the change of the instrument delays and so on.

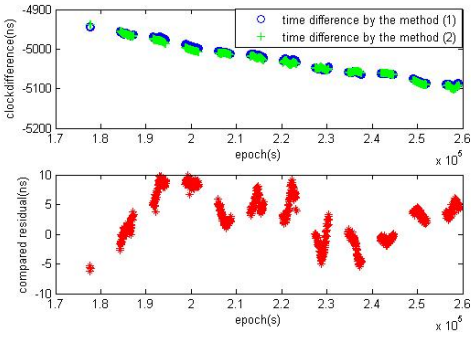


Fig2. the results on October 5th

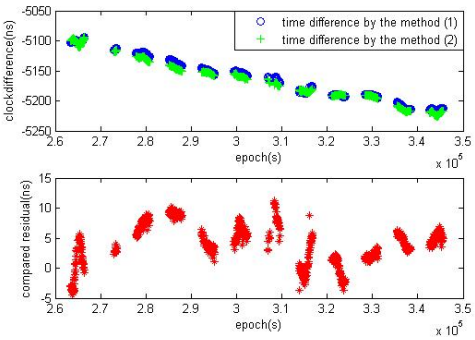


Fig3. the results on October 6th

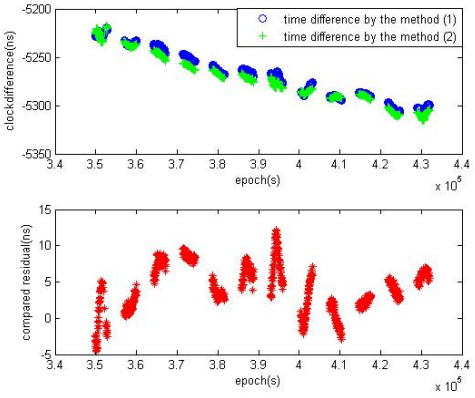


Fig4. the results on October 7th

From October 5th to 7th, the comparison shows that there are some systematic errors, which have not been corrected completely, such as the ionospheric delay. we need further analysis.

IV. CONCLUSIONS

From the analysis and comparison above the following conclusions can be drawn:

1. Combination of two-way transfer range and pseudo-range data can be directly measured the clock errors of satellite and ground station;

2. Time comparison between the navigation satellite-borne clock and the ground system standard clock can be achieved by integration of

two-way transfer range data and pseudo-range. As the GEO satellite needs maintain its position relative to the equator in some time intervals, it must be adjusted regularly. Solution ② is not reliable during and after an orbital maneuver, because the orbital errors determined affect the satellite-borne clock bias measured, it could harm the separation of the satellite coordinates and on-board clock errors. The use of integrated solver ④ to determine the satellite clock error reflects the actual satellite clock error. This can be based on navigation satellite clock error prediction to improve the real-time navigation and positioning accuracy.

This paper presents the measurement method of clock errors of satellites and ground stations based on combination two-way transfer range with pseudo-range observations. The preliminary test results show that the method can weaken the interaction of GEO navigation satellite coordinates and clock offsets, determining the satellite clock errors using a single tracking station. Taking into account the carrier phase measurement accuracy is far higher than the pseudo-range observation accuracy, in future work, we will further develop integration of two-way ranging and carrier phase observations to measure clock errors of the GEO navigation satellite.

V. REFERENCES

[1] Li Xiaohui,Yang Xuhai, et al. Precision measurement of time-frequency signal[M], Beijing: Science Press,2010. (in Chinese) pp197-215

[2] LI Jisheng. Satellite Precision Orbit Determination [M]. Beijing: Army Publishing House, 1995. (in Chinese)

[3] GAN Guoqiang. Navigation and Positioning [M]. Beijing: National Defense Industry Press , 2000. (in Chinese)

[4] Li Z G, Yang X H, Ai G X, et al. A new method for determination of satellite orbits by transfer. Science China Ser G-Phys Mech Astron, 2009,52: 384–392

[5] QIN Xian-ping, YANG Yuan-xi, JIAO Wen-hai, WANG Gang, Determination of Navigation Satellite Clock Bias Using SLR and Pseudo-range Data [J]. ACTA GEODAETICA et CAR TO GRAOHICA SINICA, 2004, 33 (3):205-209(in Chinese)

Active optics for the interrogation of a single ion

Gaetan Hagel, Olivier Morizot, Emeline Bizri, Didier Guyomarc'h,
Caroline Champenois, Marie Houssin, Jofre Pedregosa, Martina Knoop

PIIM, CNRS/Aix-Marseille Université, Centre de St Jerome, Case C21, F-13397 Marseille cedex 20, France
Email : Martina.Knoop@univ-provence.fr

***Abstract*— The signal-to-noise ratio in the detection of a single ion is one of the determining factors in the speed of the detection of the ion's internal states. Its value defines the necessary interrogation time to distinguish between an ion in the "on" and an ion in the "off" state. The improvement of the detection signal by means of active optics is a new approach in single-ion experiments. They accompany further progress in the stabilization of the clock laser for the single ion interrogation.**

I. Introduction

The frequency stability of an optical frequency standard can be quantified by the evaluation of the Allan deviation [1], where the stability value depends on the interrogation time, the cycle time of the measurement as well as the signal-to-noise ratio (SNR) which can be realized throughout the probing of the clock transition.

Single ion detection is always carried out on the cooling laser's wavelength as it scatters a maximum of photons. Today, trap structures are getting smaller and smaller and even with tightly focused laser beams, scattered light from the metallic electrodes contributes to the background light level. A high quality detection objective [2] as well as spatial filtering in the imaging process are among the stringent requirements for the detection of a single trapped ion.

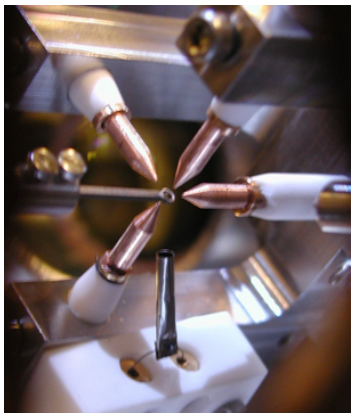


Fig. 1. Miniature ion trap

A perfect Gaussian beam of 100 μW focused to a waist diameter of 20 μm in the centre of a trap (even if it is a miniature trap) should give a scattered light level of less than one photon per second collected by the optical detection solid angle. However, it can be observed in essentially all single ion trap experiments ([3], and references therein) that the background level is typically larger than the photomultiplier's dark count rate (10 photons per second). Even with high-quality UV optics (a large majority of single-ion experiments has a cooling laser in the blue or near UV) in many experiments a deformation of the laser-beam can be observed, as well as an additional halo due to scattering in the optics.

Aberrations in a beam have been corrected with active (or adaptive) optics (AO) elements at multiple occasions since seminal works in high-intensity lasers [4] and of course in astronomy [5]. Moreover, it has been shown that strong scattering in optics can be corrected by an AO element [6].

We propose the use of an AO element to optimize the SNR in the detection of a single trapped ion. In this manuscript we will develop the necessary requirements and envisaged solutions. We suppose that aberrations in our system do not show temporal fluctuations, we will thus concentrate on means of active optics rather than adaptive.

II. Cooling and detection of a single ion

The interrogation of a single ion as a frequency standard is made with the help of the electron-shelving technique [7], the interrogation statistics is built by sequential probing of the fluorescence state of the ion which indicates whether the ion is in the "on" (scattering photons) $4S_{1/2}$-$4P_{1/2}$-$3D_{3/2}$ - cycle or shelved in the dark ("off") $3D_{5/2}$-state (cf. energy level scheme in Figure 2).

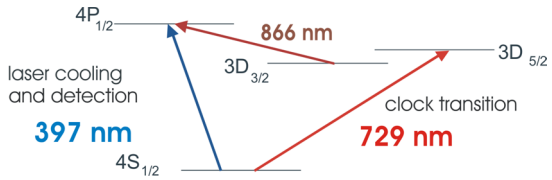


Fig. 2. Scheme of the first energy levels of $^{40}Ca^{+}$

Quantum jump statistics are presented in Figure 3, the right part of this figure shows the distinction between "on" and "off" state. The higher the SNR the faster the interrogation procedure can be. This value conditions the cycle time in the interrogation of the single ion for all types of high-resolution experiments.

In our experiment, a frequency-doubled titanium-sapphire laser at 397 nm is used for laser-cooling the ion on the $4S_{1/2}$-$4P_{1/2}$ electric dipole transition. This laser is transported onto

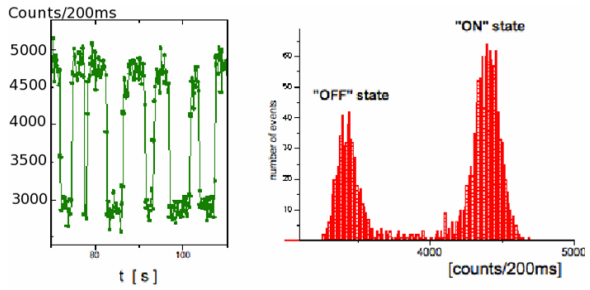


Fig. 3. Quantum jump statistics in the interrogation of a single trapped ion

the ion experiment by a single-mode optical fiber, which also spatially cleans the laser beam. The laser is then split in three beams in order to dispose of different probing axis. This separation is done by half-wave plates in conjunction with polarization beam splitter cubes. Each beam is then focused by a 50 mm-diameter achromat lens into the vacuum vessel to a waist of about 20 μm radius.

A. Preliminary characterization of the cooling laser

In order to obtain a quasi-perfect beam in the plane of the single ion, i.e. a Gaussian beam focused on a waist of 20 μm, it is necessary to measure the deformations of this beam as well as to estimate the scattering of this beam by the optical elements of the system, before applying eventual corrections. Scattering in optics has been suppressed by spatially filtering with a pinhole in the image plane of the laser waist and is considered to be negligible in the final background detected.

In order to identify the origin of the unusually high background of scattered light in our detection system, a realistic description of the beam profile is required, which considers the beam not only as a perfect Gaussian envelope but as a Gaussian field amplitude with an aberrated wavefront. This approach has to be made experimentally as well as theoretically. The corner stone of the description of an aberrated beam is the decomposition of the wavefront surface on the orthonormal basis of the Zernike polynomials.

Hence, a beam characterization was carried out on our set-up by different means. Analysis were made with two types of wave-front sensors based on different technologies (Shack-Hartmann and phase-shearing), and imaging with a camera is under way. The presence of aberrations in the laser beam can be quantitatively described with the Strehl ratio, or in terms of encircled energy, which is the intensity deposited in a given area. Indeed, the fiber output shows an almost perfect Gaussian

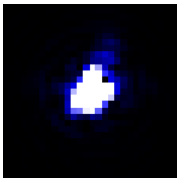


Fig. 4. Deformed laser beam as measured with all optical elements inserted in the beam-line

beam, with a Strehl ratio higher than 95%. Travelling through the optical system, the beam is then deformed, and the Strehl ratio drops below 80% at the site of its focus on the ion, as can be seen in Figure 4. Monitoring the beam by a CCD camera is a complementary method of waist characterization. For reasons of acquisition dynamics, the center of the beam has to be masked in order to reveal residuals in the wings.

Correction of the observed distortions in the beam profile can be carried out by different means, actually, several types of AO corrective elements are available : deformable mirrors (DM), spatial light modulators (SLM) and optical valves (OV). While DM have a number of actuators typically limited to a few hundreds, and therefore present a limited spatial resolution, they show a modulation depth of typically a few micrometer. On the other hand, spatial light modulators are LCD devices with an important number of pixels (on the order of 10^5), their main drawback is the important diffraction patterns they produce. Finally, OV probably present the best spatial resolution, but the amplitude dynamics of the correction is rather weak.

Only the precise characterization of the laser-beam with the identification of the induced perturbations allows to determine the spatial resolution and amplitude dynamics necessary for the choice of the AO element adapted to his application.

The influence of aberrations in the beam on the noise level in the single-ion detection has been evaluated with the use of a deformable mirror. This 97-actuator device with a mirror diameter of 13.5 mm was set-up in normal incidence at the entrance of the vacuum vessel. Preliminary measurements have demonstrated the link between the SNR and wavefront modifications. The measured noise level is represented on the graph in Figure 5 as a function of induced aberrations in the beam.

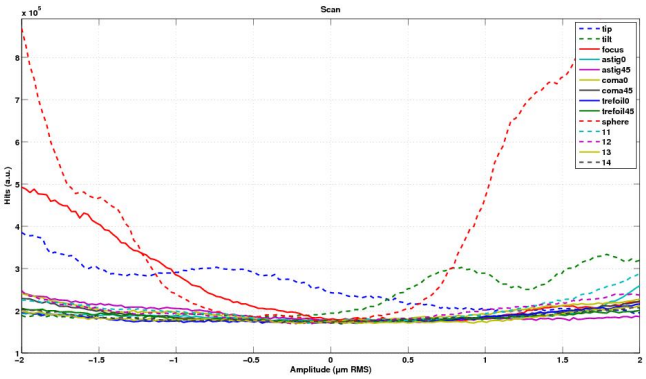


Fig. 5. Noise level in the optical detection line as a function of the amplitude of aberration, decomposed for the different Zernike polynomials

B. Simulation of the cooling laser beam

In parallel to the experimental characterization of the optical set-up, a quantitative analysis of its defects based on simulations of the full optical system up to the trap is carried out. Using an optical design software (Zemax), we can precisely reproduce each optical element (plane mirrors of finite size,

achromatic lenses, pinholes and the entrance viewport of the vacuum chamber) and predict the amount of aberrations that affects the beam in the end, considering the consequences of chromaticity, tilting and off-centering of the optics. The results of this study consist of a qualification and a quantification of the optical defects of the system decomposed on the basis of Zernike polynomials. For each polynomial of this basis, the depth value of the deformation in units of λ is obtained and this computer assisted study, validating the experimental observations, provides us with a versatile tool to precisely understand and even predict the effects of the modifications introduced in the set-up (especially modifications of the angle of incidence).

A second type of simulation is based on the measurement of the Zernike coefficients in the pupil plane of the system. With the input of the measured coefficients in a numerical program and the propagation of the aberrated field using the Fresnel formalism, it is possible to deduce the intensity profile (point-spread function) of the beam, at the site of its focus inside the trap. Figure 6 shows the simulated profile of the focused aberrated Gaussian beam with 0.4 μm of coma in x and y directions and -0.2 μm of 3rd order spherical aberration. This work directly allows to quantify the percentage of light that can be intercepted and diffused by the trap structure and to understand the consequence of each aberration on the detection noise we observe.

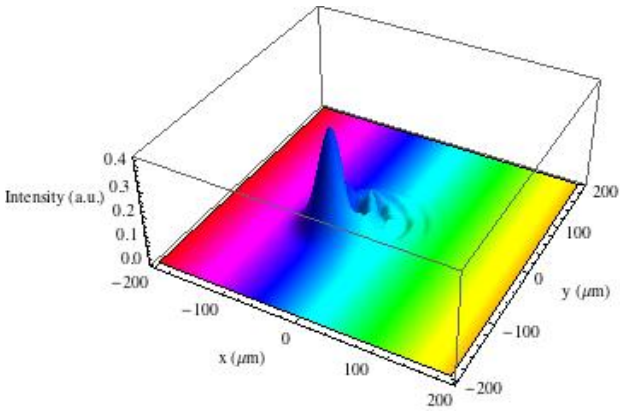


Fig. 6. Simulated point-spread function of the aberrated Gaussian beam with 0.4 μm of coma and -0.2 μm of 3rd order spherical aberration. Note that the beam waist is approx. 20 μm and the trap's inner diameter is 190 μm.

C. Feedback algorithm on the AO element

In astronomy the feedback on the active optics element is made by a two-dimensional signal either from a wavefront analyser or from a CCD camera, located at a reference position. The presented laser-ion interaction experiment could allow to test correction algorithms based on the only signal and noise values. Using these two 1D signals for correction falls into the category of the novel approach of wavefront sensorless adaptive optics [8].

An intermediate approach would be to use a classical type of correction, for instance a pre-registered correction map, for a first rough adjustment, and to apply corrections based on 1D signals only around this pre-optimized solution. The main interest of such a method is that it could widely increase the correction bandwidth. Moreover, in case of success, this would also reduce the requirement of a permanently installed wavefront analyser on the experiment.

III. Clock laser stabilisation

To probe the narrow, electric quadrupole transition $4S_{1/2}$ - $3D_{5/2}$ of the Ca^+-ion at 729 nm, the spectral linewidth and the frequency stability on a 1-second time-scale of the clock laser have to be of the same order of magnitude as the clock transition's natural linewidth (below 1 Hz, [9]).

Our home-built titanium-sapphire laser at 729 nm is pre-stabilized onto a low-finesse (≈ 1000) Invar-cavity in a 10^{-6}-vacuum by the Pound-Drever-Hall technique resulting in a linewidth of the order of a few kHz. Further reduction of the linewidth and in particular frequency stabilisation is made by locking to a vertically mounted, high-finesse ULE cavity of 150 mm length. This cavity has been numerically and mechanically optimized [10], to render it vibration-insensitive to a maximum extent. Fused silica mirrors are optically contacted to the ULE-spacer, the finesse of this cavity has been measured to be 140 000.

The cavity is surrounded by three thermal screens of polished aluminum, mounted in a vacuum-vessel. This set-up is enclosed in a box, with an active thermal control and mounted on a passive vibration isolation platform (see Figure 7). Locking of the clock laser on this cavity is made by a commercial locking electronics (Toptica).

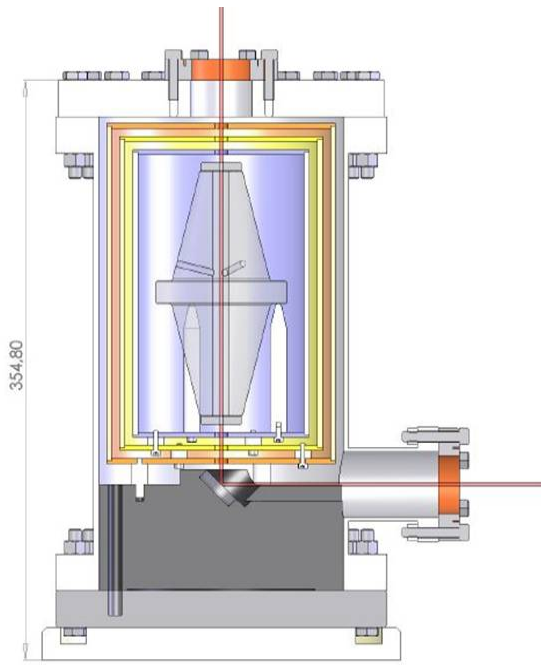


Fig. 7. The high-finesse cavity is thermally and vibrationally stabilized by different means, see detailed description in the text

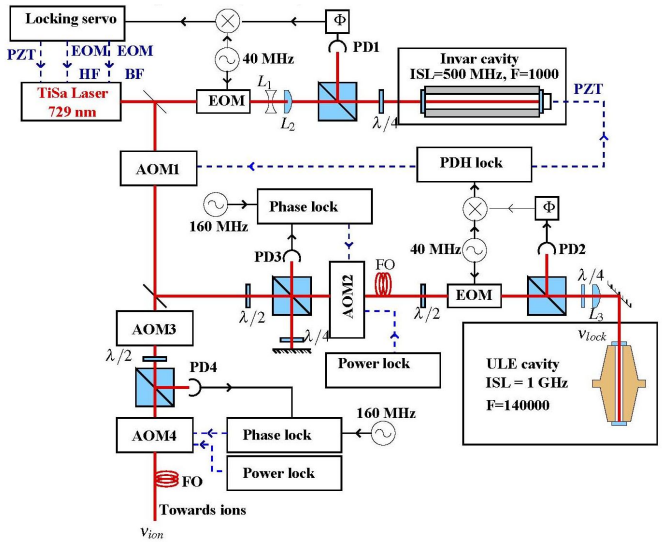


Fig. 8. Clock laser stabilisation scheme

IV. Conclusion

The use of an active optics device to control the wavefront and intensity profile of a cw laser interrogating single atoms in a trap opens the way to a tailored interaction of the electromagnetic wave with the atom. This approach will help to increase the SNR and thus optimize the frequency stability of an atomic clock.

Additional advantages are expected for the implementation of a similar device for the probing of single ions in micro- and surface traps, which have even smaller dimensions, and which are extremely sensible to the illumination by ultraviolet light. Addressing of individual ions in ion chains [11] is a supplementary application of the use of AO devices, allowing a 'real-time' correction of laser beams probing ions at different sites.

Acknowledgment

The authors acknowledge Thierry Fusco, ONERA and Brice Leroux, LAM for fruitful discussions, and ALPAO for the gracious loan of the deformable mirror.

References

[1] D. W. Allan, "Statistics of Atomic Frequency Standards," *Proc. of the IEEE*, vol. 54, pp. 221–230, 1966.

[2] W. Alt, "An objective lens for efficient fluorescence detection of single atoms," *Optik - International Journal for Light and Electron Optics*, vol. 113, no. 3, pp. 142 – 144, 2002. [Online]. Available : http ://www.sciencedirect.com/science/article/B7GVT-4DR0VJR-1B/2/f36294f532126ad448d6f981861b1f18

[3] M. Knoop, L. Hilico, and J. Eschner, "Modern applications of trapped ions," *Journal of Physics B : Atomic, Molecular and Optical Physics*, vol. 42, no. 15, p. 150201 (1pp), 2009. [Online]. Available : http ://stacks.iop.org/0953-4075/42/150201

[4] F. Druon, G. Chériaux, J. Faure, J. Nees, M. Nantel, A. Maksimchuk, G. Mourou, J. C. Chanteloup, and G. Vdovin, "Wave-front correction of femtosecond terawatt lasers by deformable mirrors," *Opt. Lett.*, vol. 23, no. 13, pp. 1043–1045, Jul 1998. [Online]. Available : http ://ol.osa.org/abstract.cfm ?URI=ol-23-13-1043

[5] P. Léna, "Adaptive optics : a breakthrough in astronomy," *Experimental Astronomy*, vol. 26, pp. 35–48, 2009, 10.1007/s10686-009-9155-8. [Online]. Available : http ://dx.doi.org/10.1007/s10686-009-9155-8

[6] I. M. Vellekoop and A. P. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.*, vol. 32, pp. 2309–2311, 2007.

[7] H. Dehmelt, "Radiofrequency spectroscopy of stored ions I : storage," *Advances in Atomic and Molecular Physics*, vol. 3, pp. 53–72, 1967.

[8] H. Linhai and C. Rao, "Wavefront sensorless adaptive optics : a general model-based approach," *Opt. Express*, vol. 19, no. 1, pp. 371–379, Jan 2011. [Online]. Available : http ://www.opticsexpress.org/abstract.cfm ?URI=oe-19-1-371

[9] M. Knoop, C. Champenois, G. Hagel, M. Houssin, C. Lisowski, M. Vedel, and F. Vedel, "Metastable level lifetimes from electron-shelving measurements with ion clouds and single ions," *Eur. Phys. J. D*, vol. 29, pp. 163–171, 2004.

[10] D. Guyomarc'h, G. Hagel, C. Zumsteg, and M. Knoop, "Some aspects of simulation and realization of an optical reference cavity," *Phys. Rev. A*, vol. 80, no. 6, p. 063802, Dec 2009.

[11] M. Marciante, C. Champenois, J. Pedregosa-Gutierrez, A. Calisti, and M. Knoop, "Parallel ion strings in linear multipole traps," *Phys. Rev. A*, vol. 83, no. 2, p. 021404, Feb 2011.

Development of a Cryogenic Sub-Hz Laser System for Optical Clocks

Thomas Kessler, Christian Hagemann,
Thomas Legero, Uwe Sterr and Fritz Riehle
PTB and QUEST
Bundesallee 100, 38116 Braunschweig
GERMANY
Email: christian.hagemann@ptb.de

Michael Martin and Jun Ye
JILA, NIST and
University of Colorado, 440 UCB, Boulder, CO 80309-0440
USA

***Abstract*— We present the setup of a cryogenic single-crystal silicon optical cavity with a potential thermal noise floor below 10^{-16}. The system comprises a vibration-free cryostat based on gaseous nitrogen as coolant.**

Thomas Kessler
May 30, 2011

I. Introduction

The stability of optical clocks is currently limited by the performance of the interrogating laser system. The reference cavities used to stabilize the laser suffer from thermal noise on a level of fractional instability of $10^{-16} - 10^{-15}$ at one second [1]–[3]. To overcome these limits a sub-Hz laser system based on a monocrystalline silicon resonator is being developed with expected instability below 10^{-16}. To avoid a large sensitivity to temperature fluctuations, state-of-the art cavities are typically operated close to the zero-crossing of the coefficient of thermal expansion of the cavity material. For silicon, the zero-crossing lies at approximately 124 K [4], therefore requiring construction of a dedicated cryostat. As an additional requirement the maximum accelerations induced by the cryostat should not exceed 1 μg.

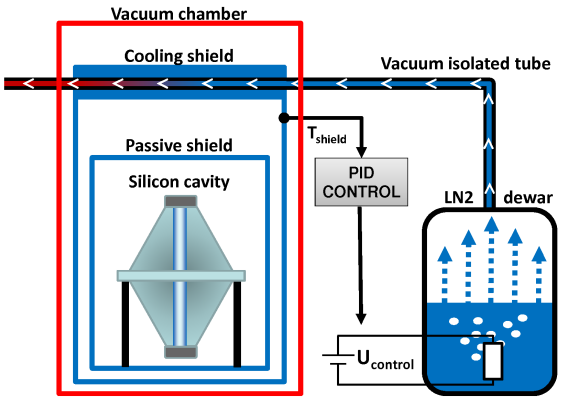


Fig. 1. Schematic sketch of the cryostat. The outer shield is cooled by nitrogen gas evaporated from a dewar.

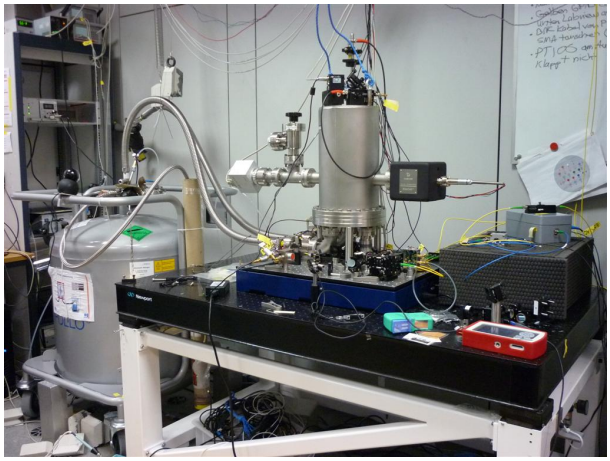


Fig. 2. Photograph of the cryostat.

II. Experimental Setup

The cryostat has been developed in a collaboration between PTB and JILA. A sketch of the setup is shown in figure 1. The cavity is held by three rods which are made of Teflon to suppress heat exchange with the base. To avoid refractive index fluctuations and for thermal insulation, the cavity is placed in a vacuum chamber. The vacuum chamber is set up on a vibration isolation table. The outer shield of the cryostat is actively cooled by a flow of nitrogen gas evaporated from a dewar. The temperature of the shield is obtained by a Platinum sensor and fed into computer-controlled feed-back system. Additional thermal insulation of the cavity from its surroundings is reached by a gold-plated passive copper shield placed between the cavity and the outer shield.

Figure 2 shows a photograph of the complete setup. The gas is evaporated from a 200 liter liquid nitrogen dewar and guided by superinsulated vacuum tubing to the cryostat. The tubes are suspended by additional weights for improved vibration isolation. A fiber laser operated at a wavelength of 1.5 μm is locked onto the cavity. All free-space optics are placed on the vibration-isolation platform to minimize optical paths. In thermal equilibrium a consumption of 1 liter of liquid nitrogen per hour is achieved with residual vibrations well below the

978-1-61284-110-6/11/$26.00 ©2011 IEEE

required specification. The thermal shield suppresses residual Temperature fluctuations of a few mK at the cryostat down to μK level for the cavity.

III. Conclusion and Outlook

A vibration-free cryostat has been built for operation of a single-crystal silicon cavity at a temperature of 124 K. First tests indicate that the performance of the cavity will not be limited by the cryogenic enviroment and therefore a stability below 10^{-16} seems feasible. After the completion of the characterization of the cryostat we will focus on a direct frequency comparison with state-of-the-art room-temperature stable cavities to explore the potential of the silicon cavity.

References

[1] Y. Y. Jiang, A. D. Ludlow, N. D. Lemke, R. W. Fox, J. A. Sherman, L.-S. Ma, and C. W. Oates, "Making optical atomic clocks more stable with 10^{-16} level laser stabilization," *Nature Photonics*, vol. 5, pp. 158–161, 2011.

[2] J. Millo, D. V. Magalhães, C. Mandache, Y. Le Coq, E. M. L. English, P. G. Westergaard, J. Lodewyck, S. Bize, P. Lemonde, and G. Santarelli, "Ultrastable lasers based on vibration insensitive cavities," *Phys. Rev. A*, vol. 79, p. 053829, 2009.

[3] S. A. Webster, M. Oxborrow, S. Pugla, J. Millo, and P. Gill, "Thermal-noise-limited optical cavity," *Phys. Rev. A*, vol. 77, pp. 033 847–1–6, 2008.

[4] C. Swenson, "Recommended Values for the Thermal expansitivity of Silicon from 0 to 1000 K," *J. Phys. Chem. Ref. Data*, vol. 12, pp. 179–182, 1983.

Doppler-free spectroscopy using a mid-infrared optical parametric oscillator linked to a frequency comb generator

Markku Vainio[1,2], Mikko Merimaa[1], and Lauri Halonen[2]
[1]Centre for Metrology and Accreditation (MIKES)
Espoo, Finland
[2]Laboratory of Physical Chemistry, Department of Chemistry
University of Helsinki
Helsinki, Finland
Email: markku.vainio@helsinki.fi

***Abstract*—A method to measure the absolute frequencies of molecular transitions in the mid-infrared region is presented. The method is based on a mid-infrared cw optical parametric oscillator, which can be used for sub-Doppler molecular spectroscopy at ~ 3.4 μm while being phase-coherently linked to an optical frequency comb in the visible/near-infrared.**

I. Introduction

Absolute optical frequency measurements can be done relatively easily in the visible and near-infrared using self-referenced, octave-spanning optical frequency combs (OFCs) based on mode-locked femtosecond lasers [1, 2]. The task is significantly more difficult in the mid-infrared. Despite the recent remarkable progress in the development of mid-infrared frequency comb generators based on femtosecond optical parametric oscillators [3, 4], most optical absolute frequency measurements reported so far at wavelengths longer than 2.5 μm have been done using continuous-wave techniques. The advantage of the cw techniques is that they allow spectroscopic measurements to be done with sub-Doppler resolution without spectral filtering, and in some cases also without an optical power enhancement cavity. As an example, Takahata et al. measured the absolute frequencies of several tetrahedral fine structure components of the P(6) and P(7) lines of the ν_3 band of methane using a narrow-linewidth mid-infrared beam that was produced by difference frequency generation (DFG) from two near-infrared laser beams [5]. The absolute frequencies of the mid-infrared transitions were obtained by referencing the two near-infrared beams to an OFC based on an Er-doped fiber laser. The measurement can alternatively be referenced to a mid-infrared OFC that has been generated by DFG from a near-infrared OFC [6, 7]. Also in this case, a cw mid-infrared source is required to probe the molecular transition, owing to the low power of the mid-infrared OFC obtained by DFG. Another technique reported earlier is to use a mid-infrared quantum cascade laser for spectroscopy and to link it to a visible/near-infrared OFC by sum-frequency mixing with a near-infrared laser [8].

Here we present a new technique for the measurement of absolute frequencies of molecular transitions in the mid-infrared. The technique is based on a cw optical parametric oscillator (OPO), which can be used for sub-Doppler mid-infrared spectroscopy. A phase-coherent link between the OPO and a visible/near-infrared OFC is established through the pump and signal beams of the OPO. This method resembles another one previously reported by Kovalchuk et al., in which a frequency comparison between a 3.39-μm CH_4-stabilized HeNe laser and a 532-nm I_2/Nd:YAG laser was demonstrated using a combination of pump-enhanced cw OPO and Ti:sapphire OFC [9]. Unlike any of the previously reported methods, our solution is based on a singly-resonant cw OPO, which produces high output power and easy wavelength tuning in the mid-infrared. This makes the experimental setup for frequency-comb referenced sub-Doppler spectroscopy simple. A particular advantage of our method is that the carrier-envelope offset frequency of the OFC does not need to be known or stabilized. This, and other characteristics of the method are described in more detail in the following section. The experimental setup is presented in Section III, and the results in Section IV.

II. Method

The basic principle of the method is to use a continuous-wave optical parametric oscillator (cw OPO) to phase-coherently link the mid-infrared and the visible/near-infrared regions of the optical spectrum. The OPO is

978-1-61284-110-6/11/$26.00 ©2011 IEEE

pumped by a near-infrared laser at frequency ν_p. Through a parametric nonlinear frequency conversion process it produces two output beams: a near-infrared beam called signal (ν_s), and a mid-infrared beam called idler (ν_i). The frequencies of the three optical beams involved in the parametric process are linked to each other by the law of energy conservation,

$$\nu_i = \nu_p - \nu_s. \quad (1)$$

According to (1), the absolute frequency ν_i of the mid-infrared beam can be determined by measuring frequencies ν_p and ν_s. This can be done using a femtosecond frequency comb generator. A typical mid-infrared cw OPO emitting at $\lambda_i \sim 3$ μm is pumped with a 1-μm laser, which sets the signal wavelength to $\lambda_s \sim 1.5$ μm [10]. Both λ_p and λ_s are therefore within the range of a typical near-infrared OFC based on an Er-fiber laser (0.95-2.1 μm) and can easily be measured. Another approach, which is the one we have used in this work, is to frequency double the pump and signal frequencies so that they can be measured with a Ti:sapphire-laser OFC that typically extends from 0.5 to 1.1 μm. The measurements can be done with a very high accuracy if the OFC is phase-locked to a microwave frequency standard. If that is the case, then also the mid-infrared frequency becomes referenced to the same standard, and a phase-coherent link between the mid-infrared and microwave regions is established (in addition to that between the mid-infrared and visible/near-infrared).

To have a closer look at the method we first assume an OFC generator (such as the Er-fiber comb) that can be used to directly measure the pump and signal frequencies of the OPO. The OFC consists of equidistant peaks, whose frequencies are

$$f_m = f_{ceo} + mf_{rep}, \quad (2)$$

where m is an integer, f_{rep} is the laser repetition rate, and f_{ceo} is the (carrier-envelope) offset frequency. The beat frequency f_p between the OPO pump frequency and the nearest comb peak m_p is then

$$f_p = f_{ceo} + m_p f_{rep} - \nu_p. \quad (3)$$

For the sake of clarity, we assume that the pump frequency is measured against the comb peak that is on its high-frequency side (so that $f_p > 0$). In a similar manner, beat frequency f_s between the OPO signal frequency and the nearest comb peak m_s (again on the high-frequency side) is

$$f_s = f_{ceo} + m_s f_{rep} - \nu_s. \quad (4)$$

Using (1)-(4), we get the absolute frequency of the mid-infrared idler beam as

$$\nu_i = (f_s - f_p) + (m_p - m_s) f_{rep}. \quad (5a)$$

In practice, the two beat frequencies can be counted with frequency counters referenced to a stable frequency standard. The difference $m_p - m_s$ of the comb mode numbers can be determined by measuring the pump and signal frequencies with an accurate wavelength meter. The repetition rate f_{rep} is phase-locked to a microwave frequency standard and hence known with high accuracy. It is important to note that the measurement is independent of the offset frequency f_{ceo}. This simplifies the experimental setup, since f_{ceo} does not have to be stabilized or measured. (In a typical octave-spanning OFC, in which the spectrum is broadened using a microstructured fiber, slow degradation and misalignment of the fiber makes long-term stabilization of f_{ceo} more difficult than stabilization of f_{rep}. Long-term operation of the comb is therefore easier if the offset frequency does not require stabilization).

If the OFC used in the experiment is based on a Ti:sapphire laser instead of an Er-fiber laser, the measurement procedure remains the same but all optical frequencies are replaced by their frequency-doubled counterparts. That is, instead of (4a) we get

$$\nu_i = \tfrac{1}{2}[(f_{2s} - f_{2p}) + (m_{2p} - m_{2s}) f_{rep}], \quad (5b)$$

where f_{2p} and f_{2s} now refer to the beat frequencies between the comb and the second harmonic frequencies of the OPO pump and signal, respectively. The absolute frequency ν_i can be unambiguously determined also in this case. Although the pump laser frequency is already within the spectral coverage of the Ti:sapphire OFC, frequency doubling of both the pump and the signal is done in order to cancel out f_{ceo}.

III. Experimental Setup

A. Optical Frequency Comb

The MIKES frequency comb generator used in this work is described in detail in a previous publication [11]. A 15-THz wide comb centered at ~800 nm is generated in a mode-locked femtosecond Ti:sapphire laser. The ~1 GHz repetition rate of the laser is referenced to an active hydrogen maser (Kvarz CH1-75A). The offset frequency f_{ceo} does not have to be stabilized or monitored, since the measurement of ν_i is independent of it.

The maximum output power of the Ti:sapphire laser is 0.7 W when pumped with 5.5 W of 532-nm cw light from a frequency-doubled Nd:YVO_4 laser. For the experiments discussed here, the output power was divided into two

branches. Most (85%) of the power was coupled through a 1-m long microstructured fiber that broadens the comb over a full octave. This comb was optimized for 532 nm in order to measure the second harmonic frequency of the OPO pump laser. The remaining 15% of the total power of the Ti:sapphire laser was coupled into another microstructured fiber, which was selected so that the original comb produced by the laser was broadened to cover the wavelengths from approximately 760 nm to 860 nm. This branch was used to measure the second harmonic of the signal frequency.

B. Optical Parametric Oscillator

The cw optical parametric oscillator used in this work is schematically shown in Fig. 1. It is based on our previous design, which is described in detail in a previous publication [10]. The OPO is pumped at $\lambda_p = 1064$ nm with an Yb fiber laser system, which produces a maximum output power of 15 W. The output beam of the pump laser is linearly polarized and has an instantaneous linewidth of ~ 100 kHz. The pump beam is focused into a nonlinear crystal, which is placed within the OPO cavity. The size of the pump beam waist in the crystal is 42 μm, which corresponds to a focusing parameter of $\xi_p \sim 2.3$ [10]. The crystal is 5% MgO-doped periodically poled lithium niobate (MgO:PPLN). The poling period used in this work was 30.5 μm. Both ends of the crystal are antireflection coated for the pump, signal, and idler wavelengths. The crystal is 1 mm thick and 50 mm long, and it is placed in a copper oven. The oven temperature can be stabilized to any value between 15 and 200°C with 10 mK precision using Peltier elements.

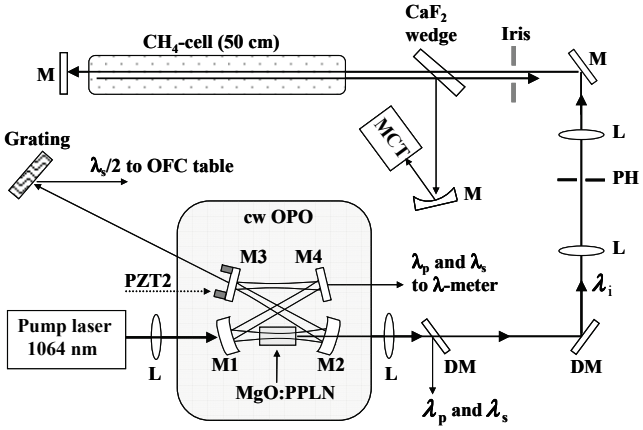


Figure 1. Schematic of the cw OPO (on gray background) and of the CH_4 spectrometer, which is described in Section III.C. The abbreviations are: DM – dichroic mirror, L – lens, M – (gold) mirror, and PH – pinhole (diameter 200 μm).

The OPO cavity is formed by 4 mirrors that are highly reflective for the signal wavelength, and transparent for the pump and idler wavelengths. The cavity is therefore singly resonant at the signal wavelength. A bow-tie ring geometry is used, and the dimensions of the cavity have been chosen so that the focusing parameter of the signal beam is matched with that of the pump beam ($\xi_s \sim \xi_p \sim 2.3$). The length of the cavity corresponds to a free-spectral range (FSR) of 580 MHz.

The first version of our experimental setup was designed for accurate spectroscopy at wavelengths between 3.2 and 3.4 μm (ν_3 band of methane). The OPO wavelength can be tuned within this range by adjusting the MgO:PPLN temperature. Gaps in the tuning range are unavoidable when using the temperature tuning method, but they can be easily covered with pump laser tuning [10]. Fast and mode-hop-free tuning of the pump laser by more than 90 GHz can be done by applying a voltage to a piezoelectric actuator that controls its cavity length. Fine tuning of the idler frequency can also be done through the signal beam by scanning the OPO cavity length using a piezoelectric actuator (PZT2 in Fig. 1), which is glued on one of the cavity mirrors. This method provides a maximum tuning range of approximately one FSR.

The single-mode output power of the OPO is > 0.5 W. Within the single-mode regime, the long-term frequency stability of the free-running OPO is typically better than 500 MHz [10]. The linewidth of the mid-infrared idler beam is of the order of 1 MHz, and therefore sufficient for sub-Doppler spectroscopy.

C. CH_4 Spectrometer

The output power of the OPO is one to three orders of magnitude higher than what can be achieved with pump-enhanced mid-infrared OPOs [9] or by DFG [5, 7]. The high output power simplifies the use of the OPO in high-resolution molecular spectroscopy. Several molecular transitions, even rather weak ones, can be measured with sub-Doppler resolution using a simple setup for saturated absorption spectroscopy. There is no need to use a power enhancement cavity. In this work, we have studied the ν_3 band of methane, which is centered at 3.3 μm. The saturation spectrometer used in the measurements is schematically depicted in Fig. 1. The mid-infrared output beam of the OPO is first separated from the residual pump and signal beams using dichroic mirrors, and subsequently spatially filtered using a pinhole. After the pinhole, the beam is collimated and passed through a 50-cm long sample cell that has Suprasil windows (ISI ASCR). The sample cell is filled with $^{12}CH_4$ to a pressure of ~1 mTorr. After the cell, the beam is reflected back and a fraction of it is directed to a MCT detector using an uncoated CaF_2 wedge. A rotatable polarizer is used to attenuate the power below the saturation level of the MCT detector. The focus of the beam is at the gold mirror that reflects the beam back. The $1/e^2$ diameter of the beam at the mirror was measured to be 2.7 mm.

D. Coherent Link Between the OPO and OFC

Doubled frequencies of the OPO pump and signal are needed in order to establish a phase-coherent link between the OPO and the Ti:sapphire OFC. To get the first one of these, a fraction of the output beam of the OPO pump laser is picked up before the OPO and frequency-doubled in a 3-mm long potassium titanyl phosphate (KTP) crystal. The crystal is cut for type II (eo-e) phase matching at room temperature; the cut angles are $\theta = 90^{o}$ and $\phi = 23.5^{o}$. The phase-matching bandwidth of the crystal is so large that no temperature control or tuning is required. With 2 W of pump power incident on the KTP, we get more than 1 mW of green 532-nm light using a simple single-pass doubling scheme. A single-mode optical fiber is used to transfer the frequency-doubled light beam to another optical table that accommodates the OFC. The beam is then combined with the green part of the OFC spectrum in a non-polarizing beam splitter. The overlapped beams are spectrally filtered with a grating and focused into a fast Si-photodiode to measure their beat frequency f_{2p}. A signal-to-noise ratio (SNR) of >35 dB was achieved in a 300-kHz resolution bandwidth (RBW) by using a long (1 m) distance between the grating and the photodiode [12].

Generation of the frequency-doubled beam of the OPO signal does not require an additional doubling crystal. A sufficient amount of optical power at $2\nu_s$ is produced in the OPO crystal as a result of a residual non-phasematched mixing process. Depending on the OPO wavelength, and with idler output powers of >0.3 W, the frequency-doubled power emitted through mirror M3 of the OPO is between 30 and 700 μW. This was measured from the -1^{st} order diffraction from a gold-coated grating, which was used to separate $\lambda_s/2$ from the other wavelengths, such as λ_p, also emitted through the mirror. After the grating, the beam was coupled into a single-mode fiber, and combined with the near-infrared part of the OFC in a 2×1 fiber coupler. The beams were then filtered using a grating before focusing them into a fast Si-photodiode for the detection of beat frequency f_{2s}. The SNR of the beat signal is 20 to 25 dB (300-kHz RBW), depending on the power of the frequency-doubled signal beam.

Both beat frequencies f_{2p} and f_{2s} can be simultaneously counted using two separate frequency counters. Alternatively, they can be mixed in an RF-mixer, in which case only one counter is required. A calibrated wavelength meter is, in general, needed to determine the integers m_{2p} and m_{2s} (Eq. 5b), i.e. the mode numbers of the OFC peaks used to measure f_{2p} and f_{2s}, respectively. Pump and signal beams leaking out through mirror M4 of the OPO are used for this purpose – they are coupled into the wavelength meter through a single-mode fiber, and optical band-pass/reject filters are used to select the wavelengths to be measured.

IV. Results

The setup described here can be used to measure the absolute frequency of the OPO idler beam anywhere between 3.2 and 3.4 μm. The long-wavelength side of the region is currently limited by the spectral coverage of the OFC´s near-infrared part. Even with the free-running OPO the setup can be used for mid-infrared spectroscopy referenced to the OFC. The resolution and accuracy of the spectroscopic measurements are, however, limited by the frequency stability of the OPO. These can be greatly improved by taking advantage of the high output power of the OPO. The OPO outputs enough power so that it can be used for saturation spectroscopy of several molecular transitions without a power enhancement cavity. The OPO idler frequency can therefore be stabilized on a Doppler-free absorption line (Lamb dip), after which it can be precisely measured with the OFC. As a first test we decided to study the P(7) transition of the ν_3 band of $^{12}CH_4$, located at 3.39 μm. This is a good test transition, since its frequency and other parameters are known from previous studies. Especially the P(7) $F_2^{(2)}$ line is well known, since it has been used for frequency stabilization of the HeNe laser. The 3.39-μm CH_4/HeNe laser is a common mid-infrared optical frequency standard, whose frequency has been determined with high accuracy. It provides a CPIM recommended radiation for the realization of the definition of the meter [13].

A scan over six tetrahedral fine structure components of the P(7) line of $^{12}CH_4$ is shown in Fig. 2. The spectrum was recorded using the spectrometer described in Section III.C, and the scan was done by tuning the frequency of the OPO pump laser. The etalon fringes in the spectrum are mainly caused by the OPO pump laser [14]. We have previously used a balanced-detection scheme to get rid of the fringes but decided not to use it here for simplicity. Instead, we are planning to solve the problem by replacing the seed laser (a fiber laser) of the pump system with a DFB diode laser.

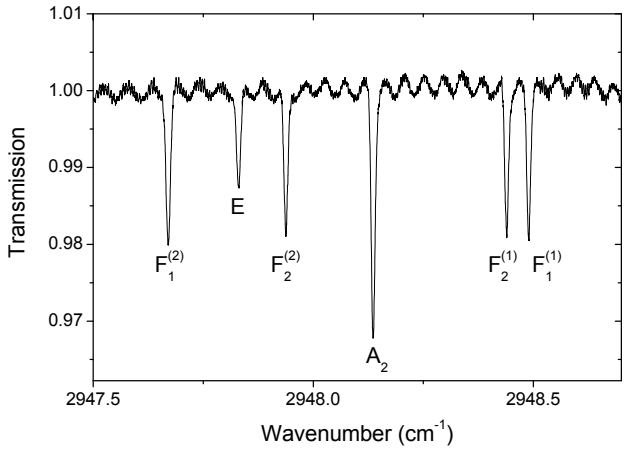


Figure 2. Absorption spectrum of the P(7) transition of $^{12}CH_4$.

Figure 3a shows a detail of the recorded spectrum, revealing the Lamb dip of the A_2 component of the transition. The observed linewidth of 4.3 MHz (full-width at half-maximum) is mainly due to the OPO linewidth and due to power broadening, which results from the high intensity used in the experiment (~ 10 times the saturation intensity). In addition to direct recording of the sub-Doppler spectrum, we used wavelength modulation spectroscopy with 3*f* detection. The frequency of the wavelength modulation was 1249 Hz, and it was done by dithering the OPO cavity length using PZT2. The modulation depth was ~2.5 MHz peak-to-peak. The third harmonic of the lock-in detected Lamb dip is shown in Fig. 3b. The time constant of the lock-in amplifier was 1 ms (24 dB) and the scan speed over the Lamb dip was 440 MHz/s.

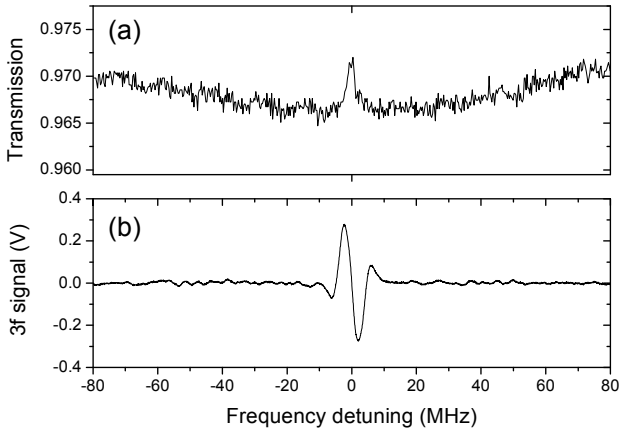


Figure 3. (a) Detail of Fig. 2, showing the Lamb dip of the A_2 component of the P(7) transition. (b) The third harmonic signal of the Lamb dip, as obtained from the output of a lock-in amplifier using wavelength modulation spectroscopy.

The 3*f* signal from the lock-in amplifier provides an error signal that can be used to lock the idler frequency to the center of a Lamb dip. We are currently carrying out the first absolute frequency measurements of the OPO idler frequency locked to a Lamb dip of the P(7) transition. The results of these measurements will be reported elsewhere.

V. Conclusion

We have presented a new method and instrumentation that can be used to measure the absolute frequencies of molecular transitions in the mid-infrared region. The method is based on the use of a high-power cw optical parametric oscillator, which can be applied to sub-Doppler molecular spectroscopy as a part of a simple saturated-absorption spectrometer. As an example, we have demonstrated sub-Doppler spectroscopy of the ν_3 band of methane at 3.4 μm. The optical parametric oscillator is linked to a Ti:sapphire optical frequency comb through its pump and signal beams. As the frequency comb is referenced to a microwave frequency standard, a phase-coherent link between the mid-infrared and the microwaves is established as well.

Acknowledgment

The financial support of the Academy of Finland is gratefully acknowledged.

[1] D. J. Jones, S. A. Diddams, J. K. Ranka, A. Stentz, R. S. Windeler, J. L. Hall, and S. T. Cundiff, "Carrier-Envelope Phase Control of Femtosecond Mode-Locked Lasers and Direct Optical Frequency Synthesis," Science, vol. 288, pp. 635−639 (2000).

[2] T. Udem, R. Holzwarth, and T. W. Hänsch, "Optical frequency metrology," Nature, vol. 416, pp. 233-237 (2002).

[3] F. Adler, K. C. Cossel, M. J. Thorpe, I. Hartl, M. E. Fermann, and J. Ye, "Phase-stabilized, 1.5 W frequency comb at 2.8–4.8 μm," Opt. Lett., vol. 34, pp. 1330-1332 (2009).

[4] N. Leindecker, A. Marandi, R. L. Byer, and K. L. Vodopyanov, "Broadband degenerate OPO for mid-infrared frequency comb generation," Opt. Express, vol. 19, pp. 6296-6302 (2011).

[5] K. Takahata, T. Kobayashi, H. Sasada, Y. Nakajima, H. Inaba, and F.-L. Hong, "Absolute frequency measurement of sub-Doppler molecular lines using a 3.4-μm difference-frequency-generation spectrometer and a fiber-based frequency comb," Phys. Rev. A, vol. 80, p. 032518 (2009).

[6] S. M. Foreman, A. Marian, J. Ye, E. A. Petrukhin, M. A. Gubin, O. D. Mücke, F. N. C. Wong, E. P. Ippen, and F. X. Kärtner, "Demonstration of a HeNe/CH_4-based optical molecular clock," Opt. Lett., vol. 30, pp. 570-572 (2005).

[7] P. Malara, P. Maddaloni, G. Gagliardi, and P. De Natale, "Absolute frequency measurement of molecular transitions by a direct link to a comb generated around 3-μm," Opt. Express, vol. 16, pp. 8242-8249 (2008).

[8] S. Borri, S. Bartalini, I. Galli, P. Cancio, G. Giusfredi, D. Mazzotti, A. Castrillo, L. Gianfrani, and P. De Natale, "Lamb-dip-locked quantum cascade laser for comb-referenced IR absolute frequency measurements," Opt. Express, vol. 16, pp. 11637-11646 (2008).

[9] E. V. Kovalchuk, T. Schuldt, and A. Peters, "Combination of a continuous-wave optical parametric oscillator and a femtosecond frequency comb for optical frequency metrology," Opt. Lett., vol. 30, pp. 3141-3143 (2005).

[10] M. Vainio, J. Peltola, S. Persijn, F. J. M. Harren, and L. Halonen, "Singly resonant cw OPO with simple wavelength tuning," Opt. Express, vol. 16, pp. 11141-11146 (2008).

[11] M. Merimaa, K. Nyholm, M. Vainio, and A. Lassila, "Traceability of Laser Frequency Calibrations at MIKES," IEEE Trans. Instrum. Meas., vol. 56, pp. 500–504, (2007).

[12] W.-K. Lee, D.-S. Yee, and H. Suhng Suh, "Direct frequency counting with enhanced beat signal-to-noise ratio for absolute frequency measurement of a He-Ne/I_2 laser at 633 nm," Appl. Opt. , vol. 46, pp. 930-934 (2007).

[13] T. J. Quinn, "Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001)," Metrologia, vol. 40, pp. 103-133 (2003).

[14] M. Vainio, M. Siltanen, J. Peltola, and L. Halonen, "Grating-cavity continuous-wave optical parametric oscillators for high-resolution mid-infrared spectroscopy," Appl. Opt., vol. 50, pp. A1-A10 (2011).

Applications of Electrodeless Discharge Rb Vapor Lamp for frequency stabilization of 1529 nm laser

Qinqing Sun, Xinyu Miao, Jingbiao Chen
School of Electronics Engineering & Computer Science
Institute of Quantum Electronics of Peking University
Beijing, P. R. China
sunqinqing@pku.edu.cn

Optical-optical double-resonance (OODR) Spectra for the $5^2P_{1/2}$ –$6^2S_{1/2}$ and the $5^2P_{3/2}$ –$4^2D_{3/2,5/2}$ transition of ^{87}Rb have been used for frequency stabilization in near-infrared region for many years. Though the method works well, it calls for locked 780nm lasers to pump Rb atoms into $5^2P_{3/2}$. We are trying to use an electrodeless discharge Rb vapor lamp to replace the complex locked pumping laser. The lamp's spectra show us that there are two spectral signals, 780nm and 795nm, are much larger than others. We are trying to make $5^2P_{3/2}$ level populated by shining the lamp light into a Rb vapor cell. Meantime, the lamp could absorb 1529 nm laser most seriously when operating in red mode. The absorption signal with high SNR but with several hundred megahertz line-width is hopefully to be used to lock the 1529 nm laser directly without 780 nm laser. The advantage of the lamp used in the frequency stabilized system is to reduce the size and cost of 1.5 μm wavelength standard for optical communication application.

I. INTRODUCTION

Frequency stabilized 1.5 μm laser is a necessary wavelength standard in optical communication [1-2]. Nowadays the frequency of 1.5 μm laser source can be stabilized by using saturated absorption spectroscopy of molecule[3-4] and double resonance spectroscopy[5-6]. The transition wavelengths between $5^2P_{3/2}$ –$4^2D_{3/2, 5/2}$ of rubidium atoms are 1529 nm. For $5P_{3/2}$ and $5P_{1/2}$ are the first excited states, double resonance optical pumping (DROP) and optical-optical double resonance (OODR)[7] are two useful ways to make these states populated. These methods work well, however they call for complex and expensive frequency stabilized 780 nm lasers.

We are trying to use an electrodeless discharge Rb vapor lamp in DROP and OODR methods to simplify the complex 1529 nm frequency locking system. The lamp's spectra show us that there are two spectra signal, 780nm and 795nm, are much larger than others[8]. So we consider that the lamp could replace the pumping laser in DROP system. Besides, for the lamp is pumped by radio frequency source, almost all the excited states of Rb atom are populated with different populations. So the lamp can absorb 1529 nm laser directly without 780 nm pumping laser. The absorption signal of 1529 nm laser could be used to stabilize the frequency of 1529 nm laser, which can be used as a wavelength standard for optical communication.

II. ELECTRODELESS DISCHARGE RB VAPOR LAMP

The Rb vapor cell in our lamp is a cylindrical glass cell winded with copper wire and can output light from the both (ends) sides of it. Details of the bulb are listed in Table 1.

Near-infrared region spectra (780 nm-1550 nm) of the lamp are measured when the lamp is working in red mode. The optical power of one side of the lamp has been measured by a pyroelectric power meter as 46 mW. The spectra of $5^2S_{1/2}$ –$5^2P_{3/2}$ and $5^2S_{1/2}$ –$5^2P_{1/2}$ transitions are much stronger than that between higher excited states. The results show that there are much more atoms in the first excited states than that in the higher ones[8].

TABLE I. PARAMETERS OF THE RUBIDIUM VAPOR LAMP

Parameters	Vapor Cell in the Lamp
Shape	Cylindrical
Diameter	3 cm
Length	3 cm
Filling	Nature Rb, 2 torr Xe
Rf signal frequency	178 MHz
Max temperature	200 ℃

III. EXPERIMENTS, RESULTS AND DISCUSSION

Firstly our experimental scheme for investigation is shown in Fig. 1. A low pass filter was used in front of the lamp trying to prevent atoms to be pumped into higher excited states. However, no obvious change of 1529 nm absorption signal was observed. The cylindrical shape of the bulb perhaps caused low pumping efficiency for the atoms on the path of light.

Project supported by National Natural Science Foundation of China (Grant No 10874009 and No 11074011)

In order to improve the pumping efficiency, the experimental scheme is changed as shown in Fig 2. Laser passed through the lamp and the nature Rb vapor cell orderly to let atoms on the light pass to touch lamp light closely. Besides, we could compare the result induced by the lamp with the traditional DROP spectra.

Laser at 780 nm is locked on the $5^2S_{1/2}$ (F=2)–$5^2P_{3/2}$ (F'=3) transition in an ^{87}Rb D2 line. Transmitted signal of 780 nm is detected when the frequency of 1529 nm laser is scanned. As shown in Fig 3, red line corresponds to the transmission signal of 780 nm laser obtained when the lamp is absence, and the blue line is measured when the lamp is working in ring mode. Actually, the traditional DROP spectra of the transition $5^2P_{3/2}$ (F=2)-$4^2D_{3/2}$(F'=2, 3) have 2 transition signals, and that of $5^2P_{3/2}$(F=2)-$4^2D_{5/2}$(F'=2, 3, 4) have 3 signals. The red line here fails to distinguish the hyperfine transition signals apart partially due to the copropagating beams of the two lasers. Thought the signal obtained corresponds to which transition is still not clear, due to the lamp, the magnitude of DROP signal is increased by 4 times with a line-width which doesn't changed so much. When the radio frequency source power supplied to the lamp is turned low and let the lamp operates in red mode, the transmitted signal of 780 nm laser disappears but the absorption signal of 1529 nm laser reaches the maximum value. In red mode, the first excited state of Rb in the lamp can gather most atoms and then make serious absorption of 1529 nm laser[9]. So the laser pass through the nature Rb vapor is too low to induce obvious DROP process. Until now, we are not very clear to what degree that the vapor is pumped by the lamp. More experiment should be done to research the underlying phenomenon.

On the other hand, since the lamp in red mode could generate an absorption signal of 1529 nm laser with big magnitude and big signal-to-noise ratio (Absorption process corresponds to transitions between higher excited state is generally weak, because that the intermediate state is not easily populated[7]), the lamp is useful to be used to lock the 1529 nm laser by the absorption signal directly. However, because of the buffer gas in the lamp, the hyperfine structures in the absorption signal are mixed together, and then induce the line-width of the signal widen to several hundred megahertz. The final stability of the laser depends on the balance of SNR and line-width. So the lamp is still hopefully can be used to reduce the size and cost of the stabilized 1.5 μm laser source for optical communication.

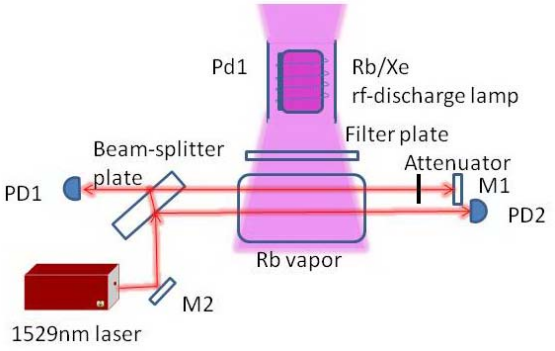


Fig 1. Experimental scheme for the first time. M1 and M2 are two reflection mirrors for 1.5 μm region. PD1 and PD2 are two Ge photodiodes, and PD1 obtains SAS signal, PD2 gets Doppler background signal. The filter plate lets only 780 nm region light pass.

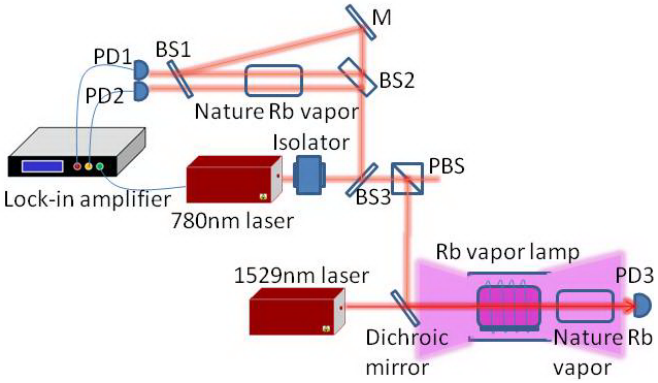


Fig 2. BS is beam-splitter mirror; PBS is polarized beam-splitter mirror; Dichroic mirror: reflection at 1529 nm, transmission at 780 nm; Filter is a high-pass mirror which cutoff frequency is 800 nm; PD1 and PD2 are Si photodiodes; PD3 is photo-detector, Thorlabs.Inc, PDA10A. At point A, the power of 1529 nm laser is 7.76 mW and, 780 nm, 1.7 mW.

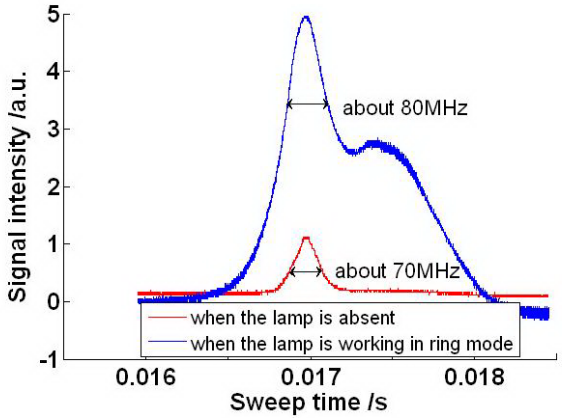


Fig 3. Detected transmitted signal of 780 nm laser

IV. Conclusion

We explored the potential applications of electrodeless discharged Rb vapor lamp in simplifying the frequency locking system for 1529 nm laser. Firstly, we checked the possibility of the lamp to replace the frequency stabilized 780 nm pumping laser, and there are still a lot of research we should do in the future. Moreover, we presented the possibility of locking the 1529 nm laser on the lamp simply by detect the absorption signal which passed through the lamp. Without the complex frequency locking 780 nm laser, the Rb vapor lamp could reduce the size and cost of a frequency stabilized 1529 nm laser as wavelength standard for optical communication application.

References

[1] J. Tatebayashi, M. Nishioka, and Y. Arakawa, "Over 1.5 um light emission from InAs quantum dots embedded in InGaAs strain-reducing

layer grown by metalorganic chemical vapor deposition," Appl. Phys. Lett, vol. 78, no. 22, pp. 3469-3471, May 2001.

[2] K. Nosu, H.Toba, and K. Iwashita, "Optical FDM transmission technique," J. Lightwave. Technol., vol. LT-5, pp. 1301-1308, September 1987.

[3] T. Yanagawa, S. Saito, Y. Yamanoto, "Frequency stabilization of 1.5 - μm InGaAsP distributed feedback laser to NH3 absorption lines," Appl. Phys. Lett., vol. 45, pp. 826-828, October 1984.

[4] M. de Labachelerie, K. Nakagawa, and M. Ohtsu, "Ultranarrow $^{13}C_2H_2$ saturated-absorption lines at 1.5 um," Opt. Lett., vol. 19, pp. 840-842, June 1994.

[5] H. S. Moon, W. K. Lee, L. Lee, and J. B. Kim, "Double resonance optical pumping spectrum and its application for frequency stabilization of a laser diode," Appl. Phys. Lett., vol. 85, pp. 3965-3967, November 2004.

[6] M. Breton, P.Tremblay, C. Julien, N.Cyr, M.Tetu, "Optically pumped rubidium as a frequency standard at 196 THz," IEEE Instru. Meas, vol. 44, pp. 162-165, April 1995.

[7] H. S. Moon, L. Lee, and J. B. Kim, "Double-resonance optical pumping of Rb atoms," J. Opt. Soc. Am. B., vol. 24, pp. 2157-2164, September 2007.

[8] Q. Q. Sun, Rongwu Sheng, and J. B. Chen, "Near-infrared region spectra of electrodeless discharge rubidium vapor lamp," National conference on optoelectronics and quantum electronics of China, pp. 225-227, March 2011.

[9] Q. Q. Sun, X. Y. Miao, Rongwu Sheng, and J. B. Chen, "Near-infrared spectra and distribution of population on excited states of electrodeless," unpublished.

Use of AOM-RN as a Phase Modulator in FM Sideband Heterodyne Technique in the Atomic Frequency Standards

Viacheslav Baryshev, Viacheslav Epikhin, Sergey Slyusarev

FGUP VNIIFTRI, Moscow Region , Mendeleevo, Russian Federation
E-mail: baryshev@vniiftri.ru

Abstract – **AOM-RN, an acousto-optic modulator designed to operate in the pure Raman-Nath diffraction regime, was applied as an optical phase modulator in FM spectroscopy of saturated absorption resonances, coherent population trapping resonances, sub-natural magneto-optical resonances in Cs vapor, i.e. such kind of optical resonances which are currently employed as the frequency references in metrology of the microwave and optical frequency standards and in magnetometry. We show experimentally that the spatial divergence of the AOM-RN frequency modulated output spectrum, the only seeming disadvantage of AOM-RN, is not an obstacle, as well, to obtain the error signals by means of Pound-Drever-Hall heterodyne technique with the use of high-finesse (up to tens of thousands) Fabry-Perot cavities as the frequency discriminators.**

I. INTRODUCTION

It has been shown in [1, 2] for the first time that phase-sensitive detection of the sub-Doppler saturated absorption resonances within D_2 line of Cs atom can be carried out by means of the conventional FM sideband heterodyne spectroscopy [3, 4] with an acousto-optic modulator operating in the Raman-Nath diffraction regime as an external phase modulator. This technique has been used to obtain the error signals for wide bandwidth, high-speed servo-loop for frequency control of the external cavity diode lasers (ECDLs), as well. Then, it has been reported in [5] on the production of AOM-RN (see fig. 1), a compact acousto-optic modulator operating purely in the Raman-Nath diffraction regime. The application of AOM-RN as a phase modulator in FM spectroscopy appeared to be reasonable, since in the case of weak AOM excitation, due to the phase modulation, its output spectrum consists of a strong carrier and two weak sidebands corresponding to the ±1st diffraction orders. It has been confirmed that FM spectroscopy with the use of AOM-RN is capable of measuring both the absorption and dispersion associated with the spectral features.

It has been shown in [5] that AOM-RN as an external phase modulator in FM spectroscopy has some features advantageous in comparison with the electro-optic modulators (EOMs) which are conventionally used for the same purposes. The important advantage is the possibility of obtaining error signal with a desired slope sign of its linear central part in a wide frequency range, which is easily achieved by varying the frequency of the amplified local oscillator signal without changing its power. This feature of AOM-RN has been utilized in a fast and wideband (about 10 MHz) ECDL locking system, developed for using in Cs fountain frequency standard at VNIIFTRI. It provides on-line tuning the ECDL frequency and optimization of the detection signals by varying the local oscillator frequency in a few tens of kHz range with corresponding shift of the saturated absorption error signal lock point in a few MHz range. Next, driving AOM-RN local oscillator signal power that is necessary to ensure a few-percent ratio of powers of each ±1st diffraction order sidebands to the carrier power did not exceed 100 mW. Actually, the existing EOMs do not produce a pure FM spectrum. Hence, one has to control the polarization of the input and output beams to remove the residual amplitude noise. When using AOM-RN, such control is not required, and the experimental setup is simplified. Also, the accurate spatial adjustment of AOM-RN is not required since, for sufficiently short acousto-optical interaction length, the Raman-Nath diffraction takes place at any incidence angle of the light on the acoustic beam. That distinguishes the Raman-Nath diffraction from Bragg diffraction which happens and is optimized at certain incidence angle satisfying Bragg condition.

FM spectroscopy in that sense as it was introduced by Bjorklund [1], when the AOM-RN modulation frequency is larger than the linewidth of the investigated transitions, enabled to get the error signals corresponding to the narrow coherent population trapping (CPT) resonances [5, 6]. It has been experimentally demonstrated [7, 8] in Cs vapor that the detection of sub-natural magneto-optical resonances in the Hanle configuration by means of FM spectroscopy with AOM-RN as an external phase modulator is possible. The electromagnetically induced absorption (EIA) error signal superimposed on the saturated absorption error signal becomes apparent when the last one corresponds to the dispersion component of

Fig. 1. Photographs of the "ISOMET" 1205-C2 AOM and couple of the "VNIIFTRI" AOM - RN.

the photo-detector beat signal formed by the frequency modulated probe radiation.

A closely related to the FM spectroscopy method the Pound-Drever-Hall method [9] is a most powerful technique for frequency stabilization of the existing lasers. Pound-Drever-Hall method introduced in practice a highly effective optical frequency discriminator: Fabry-Perot interferometer and based on that laser stabilization scheme. In our work, the Pound-Drever-Hall method with AOM-RN as an optical phase modulator has been used for frequency stabilization of the diode lasers. The only seeming disadvantage of AOM-RN is the spatial divergence of its frequency modulated output spectrum. In this paper, we will show experimentally that this divergence, as in the case of saturated absorption spectroscopy in the atomic vapor, is not an obstacle to obtain the error signals by means of FM sideband heterodyne technique with the use of high-finesse (up to tens of thousands) Fabry-Perot cavities as frequency discriminators.

II. EXPERIMENTAL SETUP AND RESULTS

The experimental setup which has been used for obtaining the error signals via Pound-Drever-Hall method is shown in Fig. 2. Developed and designed in VNIIFTRI AOM-RN provides fm spectrum of a single-mode ECDL radiation at λ=689 nm. This modulator has a center working frequency of 30 MHz and a modulation bandwidth of 20 MHz. Its acousto-optic interaction length is 2 mm. AOM-RN frequency modulated output spectrum consists of a strong carrier at frequency ω_{ECDL} and two weak sidebands at $\omega_{ECDL} \pm \Omega$ corresponding to the ±1st diffraction orders. Ω is the AOM-RN modulation frequency. The laser radiation at λ=689 nm is used in the second stage cooling of ^{87}Sr , ^{88}Sr atoms in the Sr atom optical lattice clock developed, at present, in VNIIFTRI. The second stage laser cooling of Sr atom in a optical lattice on 689 nm $^1S_0 - {}^3P_1$ transition requires an ECDL whose short-term linewidth is comparable with the 7.5 kHz transition natural width [10]. The spectrum width of the ECDL can be first narrowed by locking the laser to a stable optical cavity via the Pound-Drever-Hall technique. To ensure long-term stability, the ECDL should be next locked to the $^1S_0 - {}^3P_1$ resonance via saturated absorption spectroscopy. Again, this can be done by means of FM spectroscopy of saturated absorption resonances in the Sr atomic beam with the same AOM-RN as a phase modulator.

So, AOM-RN frequency modulated output spectrum consists of a strong carrier at frequency ω_{ECDL} and two weak sidebands at $\omega_{ECDL} \pm \Omega$ corresponding to the ±1st diffraction orders. Lenses L1 and L2 form a telescope system with AOM-RN in between. The Fabry-Perot input radiation represents three collimated and parallel light beams with size of 2 mm in the plane where all three beams lie. First, we used "Thorlabs" model SA200-5B confocal Fabry-Perot cavity with finesse about 200 and FSR=1.5 GHz. The total incident radiation power was 1.25mW. The zero background level error signal (curve *1*) and FP-cavity transmission (curve *2*) signals recorded

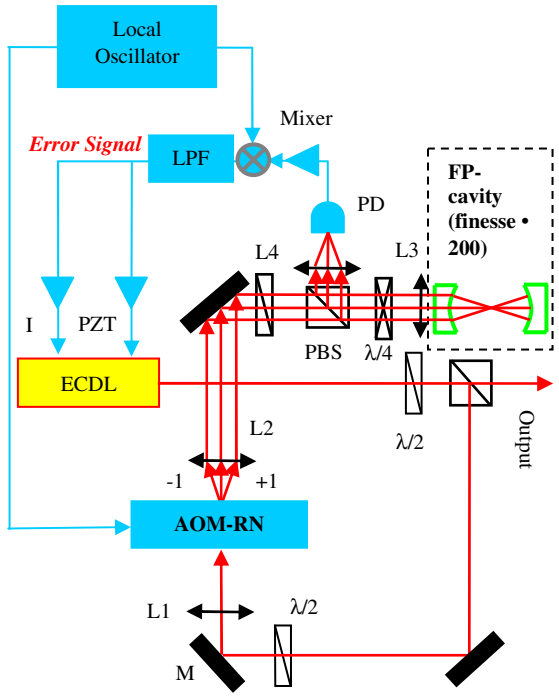


Fig. 2. The experimental setup. ECDL – external cavity diode laser; L – lens; M – mirror; LPF – low-pass filter; PZT – piezo-electric transducer; PBS – polarizing beam splitter; λ/2 and λ/4 – halve and quarter waveplates; PD – photo-detector.

simultaneously are shown in fig. 3. The AOM-RN modulation frequency was Ω = 20.25 MHz. The same signals are given in fig. 4 at Ω = 47.09 MHz. At continuous varying of the local oscillator output frequency Ω from 20 to 50 MHz, the error signal slope periodically changes its sign several times, as it can be seen in fig. 3 and 4.

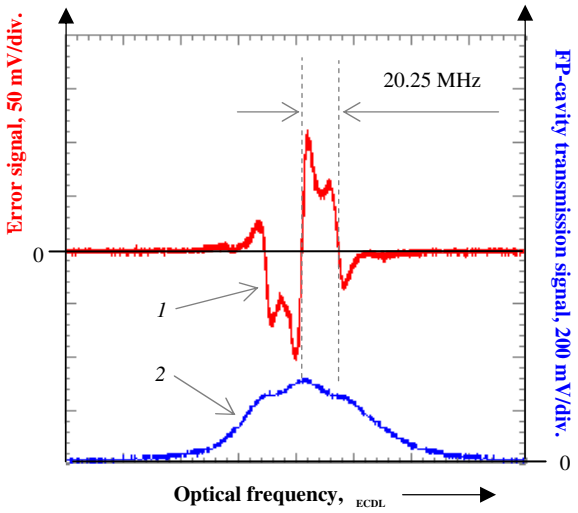


Fig. 3. The error (curve *1*) and SA200-5B confocal Fabry-Perot cavity transmission (curve 2) signals recorded simultaneously. Ω=20..25 MHz.

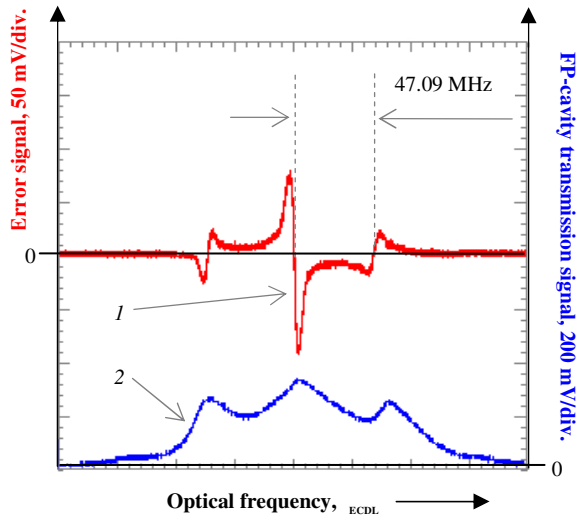


Fig. 4. The error (curve *1*) and SA200-5B confocal Fabry-Perot cavity transmission (curve 2) signals recorded simultaneously. Ω=47.09 MHz.

Since the photo-detector beat signal amplitude is proportional to the geometrical mean of the carrier intensity P_c and the intensity of some sideband P_s, the optimizing of the error signal and increase of its slope can be provided by varying the power of the driving AOM-RN local oscillator signal. Taking into account only sidebands of the ±1st diffraction orders, i.e. when the total laser power is $P_0 = P_c + 2P_s$, then the value $(P_c P_s)^{1/2}$ is maximum for $P_s / P_c = 1/2$ and $P_s = 1/4\ P_0$. Such a ratio of the power deflected into first diffraction order and carrier power several times larger than that ratio of a few percent which is sufficient to conduct the saturated absorption spectroscopy. However, it is not critical for operation AOM-RN in the Raman-Nath diffraction regime.

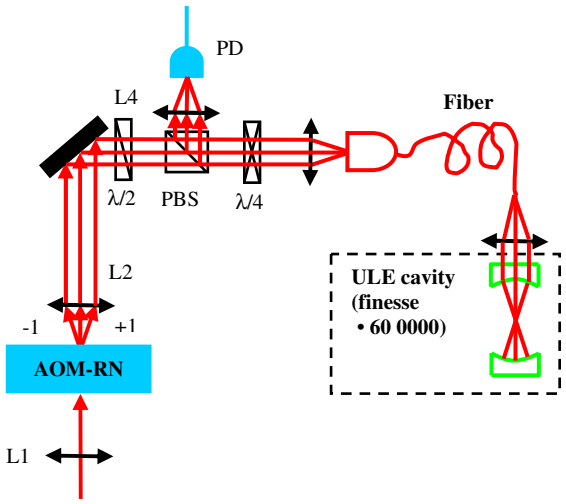


Fig. 5. Experimental setup with ULE-cavity.

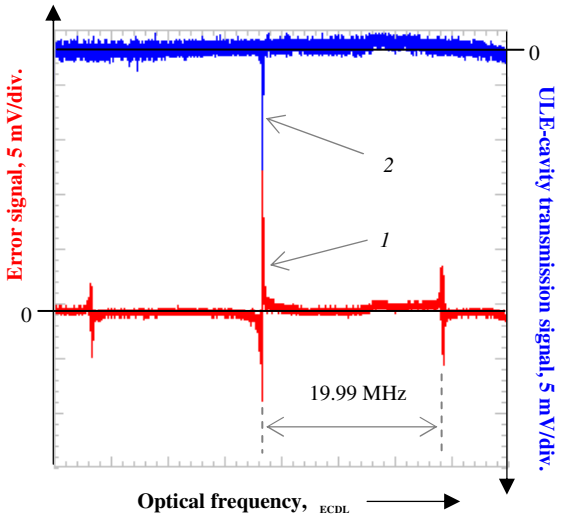


Fig. 6. The error (curve *1*) and ULE-cavity transmission (curve 2) signals recorded simultaneously. Ω=19.99 MHz.

In the same experimental arrangement, the ultra-low expansion (ULE) Fabry-Perot cavity with finesse about 60000 and FSR=1.92 GHz has been used as a frequency discriminator. A part of collimated AOM-RN output consisted, as before, of a carrier and two sidebands was fed via polarization maintaining fiber to the ULE-cavity as it is shown in fig. 5. The AOM-RN modulation frequency was Ω= 19.99 MHz. The total radiation power at the fiber output was P=0.34 mW. Three dispersion-like error signals can be observed on curve *1* in fig. 6 when the ECDL frequency ω_{ECDL} is scanned. Each of them arises when the frequencies of the upper, carrier and lower sidebands match consecutively the ULE-cavity resonance frequency.

III. CONCLUSION

AOM-RN, an acousto-optic modulator designed to operate in the pure Raman-Nath diffraction regime, was applied as an optical phase modulator in FM spectroscopy of saturated absorption resonances, coherent population trapping resonances, sub-natural magneto-optical resonances in Cs vapor, i.e. such kind of optical resonances, which are currently employed as the frequency references in metrology of the microwave and optical frequency standards. We showed experimentally that the spatial divergence of the AOM-RN frequency modulated output spectrum, the only seeming disadvantage of AOM-RN, is not an obstacle, as well, to obtain the error signals by means of Pound-Drever-Hall heterodyne technique with the use of high-finesse (up to tens of thousands) Fabry-Perot cavities as the frequency discriminators.

REFERENCES

[1] V.N Baryshev, Yu.S. Domnin, L.N. Kopylov, "Phase-sensitive detection of optical resonances by using an acousto-optic modulator in the Raman-Nath diffraction mode", *Russian J. "Quantum Electronics"*, vol. 37, (11), pp. 1006-1010, 2007.

[2] V. Baryshev, Yu. Domnin, and L. Kopylov, "External Cavity Diode Laser Frequency Control Using FM Sideband Technique with Acousto - Optical Modulator as an External Phase Modulator", *Proc. of the 21st EFTF*, pp. 107-110, 2007.

[3] G. Bjorklund, "Frequency-modulation spectroscopy: a new method for measuring weak absorptions and dispersions", *Opt. Lett*., vol. 5, pp. 15-17, 1980.

[4] J.L. Hall, L. Hollberg, T. Baer, H.G. Robinson, "Precision Spectroscopy and Laser Frequency Control Using FM Sideband Optical Heterodyne Techniques", *Laser Spectroscopy V*, ed. by A.R.W. McKellar, T. Oka, B.P. Stoicheff, Springer, Berlin, Heidelberg, New York, pp. 15-24, 1981.

[5] V.N. Baryshev, V.M. Epikhin, "Compact acousto-optic modulator operating purely in the Raman-Nath diffraction regime as a phase modulator in FM spectroscopy", *Russian J. "Quantum Electronics"*, vol. 40, (5), pp. 431-436, 2010.

[6] V. Baryshev, V. Epikhin, L. Kopylov, Yu. Domnin, "FM spectroscopy of CPT resonances with AOM operating purely in the Raman-Nath diffraction regime as optic phase modulator ", *Proc. of EFTF-IFCS* 2009, pp. 582-586, Besanson, France, 2009.

[7] V. Baryshev, "FM Spectroscopy of Nonlinear Magneto-Optical Resonances", *Proc. of 24th European Frequency and Time Forum*, *EFTF*2010, ESA/ESTEC, Noordwijk, Holland, P124, 2010.

[8] V.N. Baryshev, "Frequency-modulation heterodyne spectroscopy of nonlinear magneto-optical resonances", *in press*.

[9] R.W. Drever, J.L. Hall, F.V. Kovalski, J. Hough, G.M. Ford, A.J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator", *Appl. Phys. B*, vol. 31, pp. 97-105 1983.

[10] T. Mukaiyama, H. Katori,T. Ido, Y. Li, and M. Kuwata-Gonokami, "Recoil-limited laser cooling of 87Sr atoms near the Fermi temperature", *Phys. Rev. Lett*, vol. 90, p. 11300, 2003.

1.5-µm Cavity-Stabilized Laser for Ultra-Stable Microwave Generation

V. Dolgovskiy, S. Schilt, G. Di Domenico, N. Bucalovic, C. Schori and P. Thomann
LTF Laboratoire Temps-Fréquence
Université de Neuchâtel, Institut de Physique
Neuchâtel, Switzerland
vladimir.dolgovskiy@unine.ch

***Abstract*—Based on an extended-cavity diode laser locked to a high-finesse ultra-low thermal expansion Fabry-Perot cavity using the Pound-Drever-Hall (PDH) technique, we are developing an ultra-stable laser at 1.55 µm for the generation of low phase noise microwave. Our target short-term frequency instability is 10^{-15} at 1 s. The horizontally-mounted cavity has a resonance width of 8.1 kHz, experimentally determined from the cavity ring-down time. The analysis of the in-loop error signal of the laser stabilization to the cavity shows that the residual servo loop noise contributes only $<10^{-15}$ (@ 1 s) to the laser instability and no additional contribution to the laser linewidth is induced by the PDH stabilization. Consequently, the laser linewidth will be ultimately limited by the cavity noise only. The cavity will be stabilized at its zero coefficient of thermal expansion temperature, which has been experimentally determined to be 22.6°C.**

I. Introduction

In the development of an all-optical microwave oscillator based on an optical frequency comb (OFC), the transfer of the relative stability of an ultra-stable laser to a microwave signal is studied. The most critical component in this scheme is the ultra-low thermal expansion (ULE) reference cavity used for laser stabilization. Our target relative frequency stability in the low 10^{-15} at 1 s requires the same challenging stability in the cavity length, which must be made as little sensitive as possible to any external perturbation, such as temperature, pressure or vibration. We have designed and manufactured a 2-layer enclosure to hold the ULE cavity and minimize its temperature sensitivity. This enclosure, made of an internal copper "cradle" and an external aluminum thermal screen with gold-coated surfaces, placed in a vacuum chamber, aims at simultaneously achieving a high temperature stability (by efficiently attenuating external temperature fluctuations) and a high temperature homogeneity of the cavity. However, a much simpler 1-layer prototype enclosure made of an uncoated aluminum cradle without any temperature stabilization and little long-term temperature fluctuations filtering has been used so far to hold the cavity. Laser stabilization to a cavity resonance has been implemented and studied using this prototype enclosure.

II. Cavity-Stabilized Laser

A. Laser Frequency Noise

Our 1.55-µm extended cavity diode laser (ECDL) is stabilized to a 10-cm long, horizontally-mounted ULE cavity by the PDH stabilization technique [1]. The frequency noise properties of the free-running ECDL were first analyzed using a scanning Fabry-Perot interferometer with a resonance width of ~2 MHz. In order to use the linear side of the Fabry-Perot resonance as an optical frequency discriminator, the cavity was weakly locked to the laser by a side-of-fringe stabilization technique with a low servo-loop bandwidth of ~2 Hz.

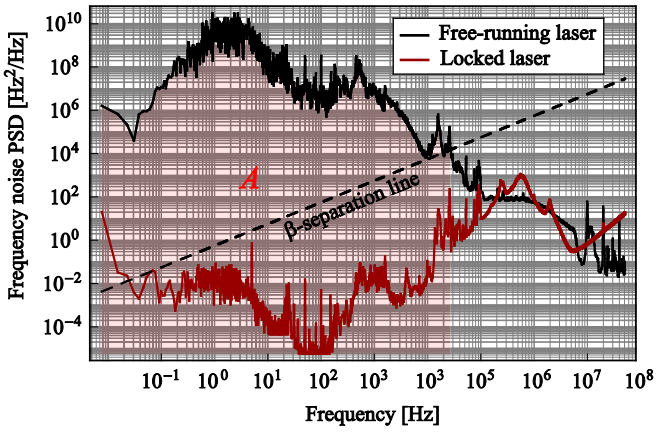


Figure 1. Frequency noise spectrum of the ECDL (free-running and locked to the ULE cavity). The free-running laser was measured using the resonance of a scanning Fabry-Perot interferometer as a frequency discriminator. Below 2 Hz, the measured frequency noise is reduced by the weak lock of the cavity to the laser. The noise of the locked laser is determined from the in-loop error signal. The β-separation line that is relevant for the determination of the laser linewidth [2] is also shown.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

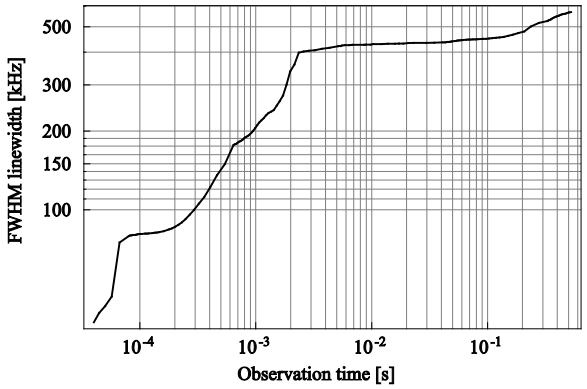


Figure 2. Linewidth of the free-running ECDL calculated from the frequency noise spectrum [2] as a function of the observation time.

The measured frequency noise power spectra density (PSD) of the free-running laser shows a general $1/f^{\alpha}$ trend between 1 Hz and 100 kHz (Fig. 1), with $\alpha \approx 1.8$. Additional noise observed in some spectral regions (e.g. the bump in the range 10^2-10^4 Hz) is of technical origin, mainly due to mechanical noise in the laser external cavity. The laser linewidth $\Delta\nu$ (FWHM) can be determined from the frequency noise spectrum according to $\Delta\nu = [8\ln(2)A]^{1/2}$ [2], where the surface A shown in Fig. 1 corresponds to the frequency noise PSD $S_{\delta\nu}(f)$ integrated over all spectral components for which the frequency noise exceeds the β-separation line given by $S_{\delta\nu}(f) = (8\ln(2)/\pi^2)\cdot f$. Due to the $1/f^{\alpha}$ type of noise, the linewidth of the free-running ECDL depends on the observation time (inverse of the low cut-off frequency considered in the frequency noise spectrum integration). The evolution of the laser linewidth as a function of the cut-off frequency is shown in Fig. 2.

Due to our lack of a suitable optical reference (e.g. another similar cavity or a second independent ultra-narrow linewidth laser) to quantify the frequency noise of the laser stabilized to the ULE cavity, only the in-loop error signal was considered to assess the performances of the stabilization loop. Absolute frequency noise of the laser (and true laser linewidth) can thus not be inferred from the analysis of this signal. However, Fig. 1 shows that the residual noise of the in-loop error signal lies well below the β-separation line in the entire analyzed frequency range. This means that the residual servo loop noise contributes only $<10^{-15}$ (@ 1 s) to the laser instability and no additional contribution to the laser linewidth is induced by the PDH stabilization. With a servo loop bandwidth of ~1 MHz, the laser thus faithfully follows the cavity and the laser frequency noise (and linewidth) will be ultimately limited by the cavity noise only.

B. Cavity Ring-Down Measurement

One of the important parameters characterizing the cavity is the finesse. The width of a cavity resonance has been determined experimentally from a ring-down time measurement. For this purpose, the laser was locked to a cavity resonance and the light transmitted through the cavity was monitored on a photodiode. Following a sudden laser switch off, the time decay of the photodiode signal was recorded on an oscilloscope (Fig. 3). Fitting the experimental

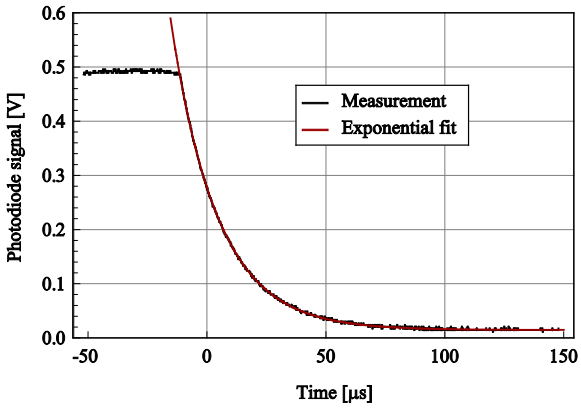


Figure 3. Cavity lifetime determined by a ring-down time measurement. A decay time of 19.6 µs is obtained from the exponential fit.

data by an exponential decay function gave a cavity decay time of τ = 19.6 µs, corresponding to a resonance width (FWHM) of 8.1 kHz. For the 1.5-GHz free spectral range of the cavity, this translates into a finesse of about 185'000.

III. CAVITY THERMAL EXPANSION

Length fluctuations of the Fabry-Perot cavity directly translate into frequency fluctuations of the laser, which tracks any change of the cavity resonance frequency. The use of ultra-low thermal expansion material allows to minimize cavity length fluctuations due to temperature variations. The temperature of the zero coefficient of thermal expansion (CTE) of the cavity (or cavity turning point) has been determined by measuring the tuning curve of the resonance frequency as a function of temperature. The resonance frequency was determined from the beat note between the ECDL locked to the ULE cavity and a fully stabilized OFC. The cavity temperature has been changed by heating or cooling the baseplate of the vacuum chamber using heaters or a set of 8 Peltier elements (for cooling). As the cavity temperature cannot be directly measured, it must be determined from the temperature measured on the surrounding aluminum cradle. The cavity temperature (T_{ULE}) is related to the cradle temperature (T_{cradle}) through the time constant τ of the heat flow between cavity and cradle:

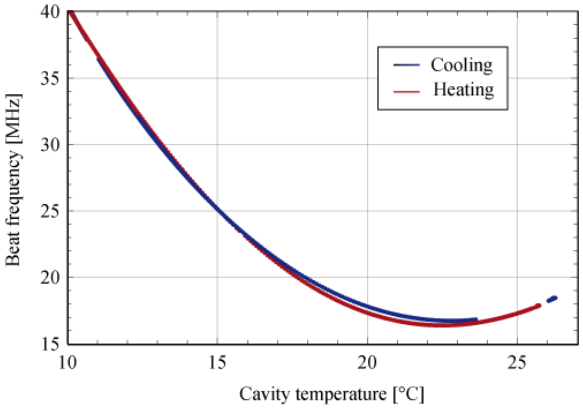


Figure 4. Determination of the cavity turning point from the beat frequency between the ultra-stable laser locked to the ULE cavity and a fully stabilized frequency comb (referenced to a H-maser) measured as a function of the cavity temperature.

$$T_{ULE}(t) = T_{cradle}(0)e^{-\frac{1}{\tau}t} + \frac{1}{\tau}\int_0^t e^{-\frac{1}{\tau}(t-\xi)} T_{cradle}(\xi)d\xi \,. \quad (1)$$

The time constant is unknown, but it can be determined if simultaneous measurements of the beat frequency and cradle temperature are performed once when the system is cooled and then when it is heated. Calculating the cavity temperature from (1) for both measurements, the time constant can be determined from the closest overlap between the heating and cooling curves [3]. Then the cavity turning point can be determined from the evolution of the beat frequency obtained as a function of the cavity temperature (Fig. 4). The turning point of our cavity was thus determined to be 22.6°C and the time constant 77'000 s. This value is achieved with a simple prototype cavity enclosure and a much higher value is expected for the final 2-layer enclosure.

IV. Conclusion and Prospective

A 1.55-µm ultra-stable laser has been developed by frequency-stabilization to a high-finesse ULE cavity. The horizontally-mounted cavity has a resonance width of 8.1 kHz, measured from the cavity ring-down time. The cavity zero CTE temperature has been determined to be 22.6°C, temperature at which the cavity will be stabilized when mounted in its designed 2-layer enclosure. A cavity time constant of 77'000 s has been achieved with a simple cavity enclosure.

The analysis of the in-loop error signal of the laser stabilization to the cavity shows that the residual servo loop noise contributes only $<10^{-15}$ (@ 1 s) to the laser instability and no additional contribution to the laser linewidth is induced by the PDH stabilization, meaning that the laser linewidth will be ultimately limited by the cavity noise only.

Acknowledgment

This work is funded by the Swiss National Science Foundation (SNSF), the Swiss Confederation Program Nano-Tera.ch scientifically evaluated by the SNSF and the Fondation pour un Laboratoire en Recherches Horlogères (FLRH).

References

[1] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward "Laser phase and frequency stabilization using an optical resonator," Appl. Phys. B. **31**, pp. 97–105 (1983)

[2] G. Di Domenico, S. Schilt and P. Thomann., "Simple approach to the relation between laser frequency noise and laser line shape," App. Opt. **49**, No. 25, pp. 4801–4807 (2010)

[3] S. A. Webster, M. Oxborrow, S. Pugla, J. Millo, and P. Gill, "Thermal-noise-limited optical cavity," Phys. Rev. A **77**, 033847 (2008)

Optical Carrier Phase Recovery and Locking in a Pulse-Position Modulation Communication Link

David C. Aveline, William H. Farr, and Nan Yu
Communication Architectures and Research Section
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California, U.S.A.
daveline@jpl.nasa.gov

***Abstract*—We have demonstrated optical phase recovery from a Pulse-Position Modulated (PPM) laser beam. Our goal is to develop and validate a laser communication and ranging method that utilizes optical coherence within the well-established scheme of PPM with photon-counting detection. Recovery of the carrier phase within optical PPM signals will enable precision Doppler tracking and ranging at optical wavelengths, benefitting spacecraft navigation and scientific measurements based on the propagation of light signals. It also has direct implications for ideal quantum-limited optical communication receivers.**

I. INTRODUCTION

Free-space optical communication is projected for future space missions requiring high data-rates and/or high efficiency. The high energy of optical photons allows detection of individual photons; thus photon-counting techniques can be employed to overcome thermal noise in detection electronics, leading to shot-noise limited performance [1]. For communication in deep space, current paradigms employ Pulse-Position Modulation (PPM) [2]. In this scheme, the light is transmitted and detected as a single laser pulse within a symbol comprised of multiple time slots. It utilizes pulses that are of uniform height and width but displaced in time from some reference position based on the amplitude of the input signal at the instant of sampling. PPM offers higher noise immunity than pulse amplitude modulation (PAM) and pulse duration modulation (PDM) because the receiver simply detects the presence of a pulse within a given time slot, while the exact duration and amplitude of the pulse are not specifically important.

The PPM approach is known for efficient data transmission, but generally it is not considered useful for precision Doppler tracking or ranging, as the phase of the laser light is not utilized by the direct detection scheme. It may seem that the irregular modulation of the laser intensity and the photon-counting detection could corrupt the carrier phase information, thereby squandering the coherent nature of the laser source. In this paper we show that the phase information is indeed preserved, and furthermore, we have demonstrated optical phase locking with a PPM laser signal.

Jet Propulsion Laboratory, California Institute of Technology under contract with the National Aeronautics and Space Administration

In this effort we built upon our earlier work including calculations and experiments designed to establish the feasibility of a deep space coherent optical transponder based on an optical phase lock loop (PLL). We modeled, developed and tested optical PLL's that operate at power levels as low as 40 femto–Watts [3]. Having demonstrated such excellent performance, a PLL–based optical transponder is a highly attractive prospect for deep space tracking with the inherent capability to outperform any microwave transponder [4].

Here we describe our investigation of carrier phase recovery in a PPM signal, with the goal of developing a hybrid laser communication and ranging scheme that utilizes optical coherence within the well-established PPM paradigm. We discuss our analyses in both frequency and time domains. We first developed code to generate pseudo-random PPM signals and model the spectrum via FFT analysis. We then measured spectra of PPM signals, and recovered the phase of a modulated radio frequency (RF) carrier. Lastly, we demonstrated optical phase locking in the presence of PPM.

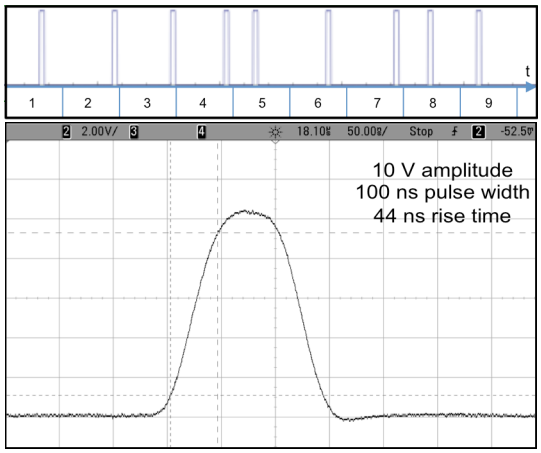


Figure 1. Top shows a sample of a code-generated PPM sequence with 9 symbols labeled below. Each symbol in this sample has 10 slots with the first empty to serve as an inter-symbol guard time reference. Below is s a scope trace of a single pulse from the signal generator.

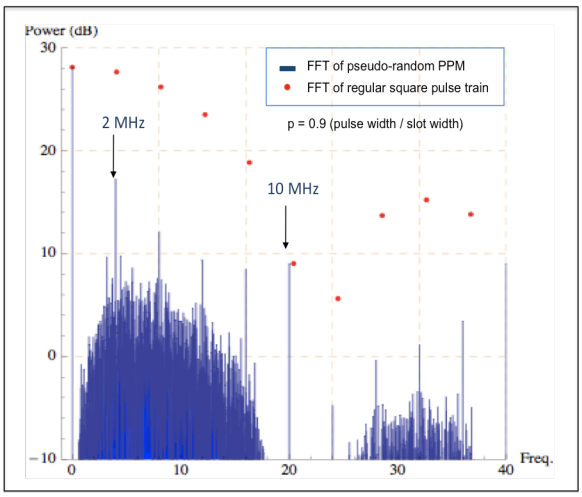


Figure 2. The plot shows the calculated FFT showing the spectrum of a 5-slot PPM with each pulse filling 90% of the slot width. For reference, in red is the FFT of a periodic train of square pulses, a pulse appearing in the same slot for every symbol. Note the apparent sinc-envelope and corresponding 2 MHz and 10 MHz features corresponding to the 500 ns symbol length, and pulse width of 100 ns.

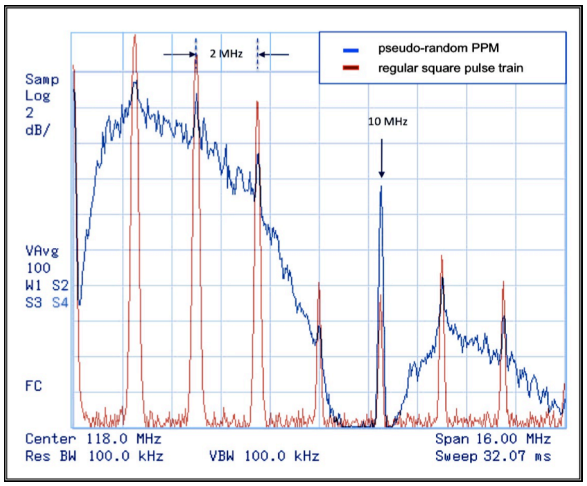


Figure 4. The measured power spectrum for a 5-slot PPM signal. Labeled is a large peak offset by 10 MHz corresponding to the pulse width of 100 ns. Also noted is the 2 MHz separation between peaks corresponding to the 500 ns symbol length. These peaks rise above a continuum in the frequency spectrum. The spectrum of a periodic pulse train is shown in light red, exhibiting the sinc-envelope.

II. PPM Spectral Analysis

We developed code to generate pseudo-random PPM data sequences for (1) numerical modeling of the PPM power spectrum, as well as (2) carrying out empirical spectral density measurements and (3) testing our carrier phase recovery techniques. The code enables adjustment of PPM parameters including total number of symbols, slots per symbol, slot-with, pulse height, pulse duration, resolution, rise-time, and reference slot position. The first slot of each symbol remains empty to serve as a reference, a standard "inter-symbol guard time" technique in PPM communication used to synchronize the receiver clock [5]. We generate randomly positioned pulses that fit within the PPM structure established by the code's input parameters (Fig. 1). These pseudo-random PPM sequences can then be uploaded to a signal generator for our laboratory tests, and/or analyzed with discrete Fourier transformation to calculate the frequency spectrum (Fig. 2). We can also tune the code to generate periodic pulse trains for comparison, and verify the validity of our simulations and experimental measurements.

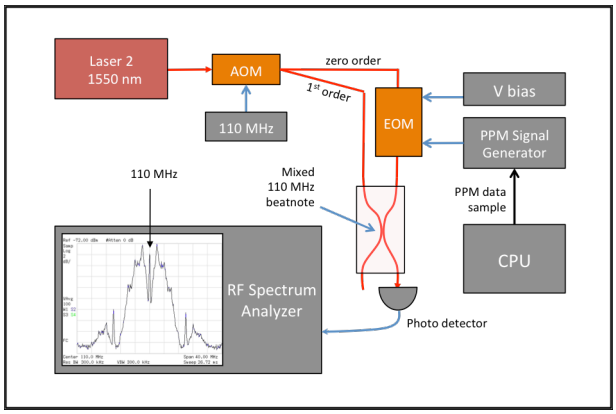


Figure 3. Diagram of the experimental set-up used to acquire the PPM signal's power spectrum. The spectrum is centered at 110 MHz, which is the RF driving the AOM shift. The fiber splitter's second channel can be utilized for real-time measurements as described in the next section.

To begin our investigation of PPM signals in the frequency domain, we measured the power spectral density of a PPM laser beam (Fig. 3). The RF spectrum was acquired by beating the PPM signal with a frequency-shifted carrier on a fast photo-detector and measuring with an RF spectrum analyzer. We directed a tunable 1550 nm DFB fiber laser through an acousto-optic modulator (AOM)). The frequency-shifted light was used as a reference, and the zero-order beam was intensity modulated by a Mach-Zehnder electro-optic modulator (EOM).

An arbitrary waveform generator, loaded with our pseudo-random PPM waveforms, controlled the modulation of the signal beam through the EOM. The EOM voltage bias was set to achieve the maximum extinction ratio for the pulsed signal. We directed this modulated optical signal through a fiber splitter/combiner, where it was mixed with the first-order shifted beam (110 MHz). We monitored the resulting power with a fast photo-detector, and analyzed the signal with a high-resolution RF spectrum analyzer. Two examples of measured power spectra are shown in Fig. 4, modulated by PPM and also a periodic pulse train for reference.

III. RF Carrier Phase Recovery

While a time-averaged frequency spectrum in the previous section exhibits mostly a continuum, underneath each pulse there exists optical carrier with instantaneous phase. Each pulse carries phase information that can be monitored and "stored" until the next pulse arrives. We

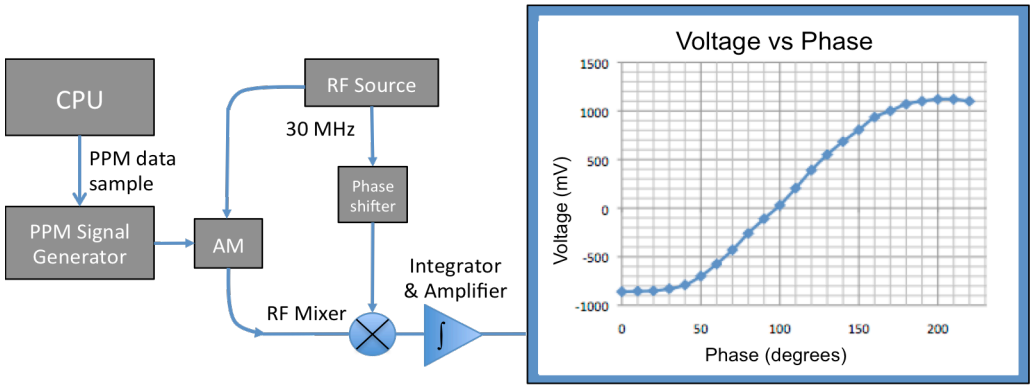


Figure 5. Diagram of the RF-based experiment designed to recover phase in a PPM signal. The mixed detection signal is integrated and amplified. We incremented the phase shift and measured the output voltage, recovering the expected sinusoidal behavior shown in the plot on the right.

accomplished this recovery by mixing the PPM signal with a reference carrier and integrating the resulting beat note. By choosing an appropriate time constant for the integration, the accumulated phase information can be acquired despite long "dark" times between pulses with low duty cycle.

To first demonstrate proof-of-principle for this approach, we operated in the RF domain with a 30 MHz carrier modulated by a slow PPM sequence. Fig. 5 illustrates this set-up. The PPM signal was mixed with the RF source reference, which could be phase-shifted through 2π radians. The resulting beat note was integrated and amplified to yield a voltage proportional to the average phase of the 30 MHz carrier. We were able to recover the phase of the pulsed signal, as shown in the plot on the right of Fig. 5. The same phase recovery procedure could be extended to optical carriers and higher frequency PPM as long as the bandwidth of the integrator is tuned accordingly.

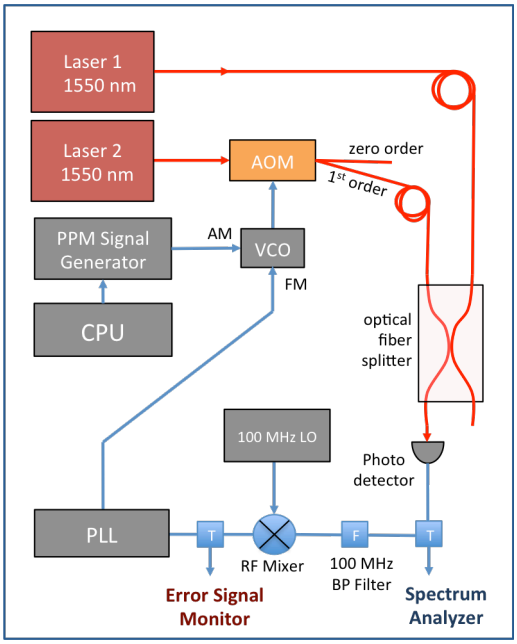


Figure 6. Diagram of the optical phase lock with PPM. Both the amplitude modulation (AM) and frequency modulation (FM) were accomplished with a voltage-controlled oscillator (VCO) driving an AOM, imparting the PPM signal on the optical beam.

IV. Optical Phase Lock with PPM

To demonstrate phase recovery in the optical domain, we phase-locked two lasers while one beam was modulated with a pseudo-random PPM code. We used a single AOM to amplitude modulate (AM) the first-order diffracted beam with the PPM waveform. A voltage-controlled oscillator (VCO) provided both amplitude modulation and frequency modulation (FM). In this way, we could feed back the frequency error signal to the PPM beam in order to phase lock it to the reference beam, a second free running fiber laser (Laser 1 shown in the set-up in Fig. 6). This approach is effectively the same as AM on one path while frequency tracking with the second. The two beams were combined in a fiber optic splitter and the beat note was measured on a fast photodiode. The resulting signal was monitored on a RF spectrum analyzer (Fig. 7). Most of the signal passed through a 100 MHz bandpass filter, and fed into an RF mixer, where it was mixed with a local oscillator (LO) of 100 MHz. This error signal was monitored on an oscilloscope and sent to the PLL. The PLL feedback controlled the FM of the VCO and served to phase lock the two laser signals. In this demonstration, we used a 500 ns symbol length containing five slots with 100 ns widths. This PLL set-up achieves optical phase locking between the two laser beams in the presence of a low-duty cycle PPM signal, thus demonstrating the preservation of optical carrier phase within a pseudo-randomly modulated signal.

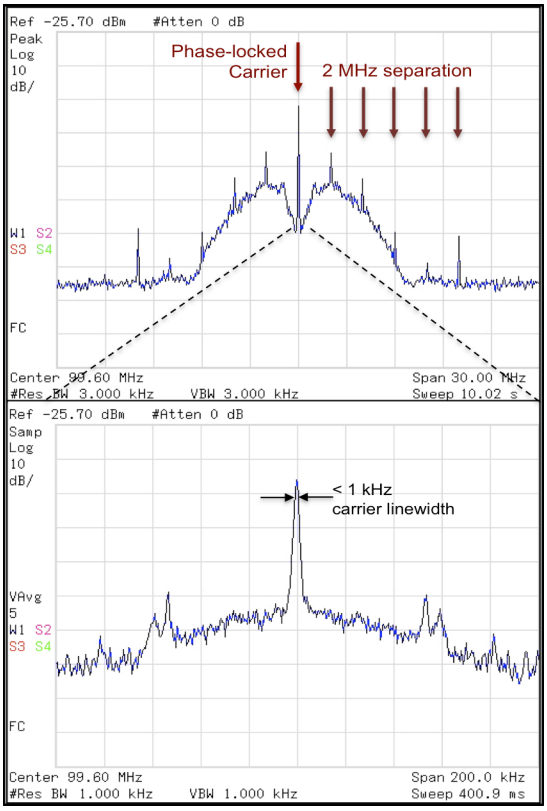


Figure 7. The measured power spectrum of a phase-locked PPM signal. The 500 ns symbol length is reflected in the 2 MHz spacing of sideband peaks (above). The zoomed-in spectrum (below) shows the phase locked laser linewidth less than 1 kHz.

V. CONCLUSIONS

We have conducted a series of simulated and laboratory experiments investigating the optical phase coherence of PPM signals. The experimental results show very good agreement with our simulations. We demonstrated that the carrier phase can be recovered from an RF-based PPM signal by employing a real-time technique to recover the phase information through the mixing of a PPM signal with a frequency-shifted local oscillator (LO) waveform. We also demonstrated that the optical carrier in the PPM can be tracked and recovered by implementing a phase lock between two laser sources, one of which was modulated with pseudo-random PPM. Future efforts will investigate these methods with photon-counting detection. We will also explore the lower limits of laser power with low duty cycles, in connection with our earlier optical transponder work demonstrating a phase lock at pico- and femto-Watt average power levels [3].

Both PPM and photon-counting are established techniques considered for optical communications in deep space [1], and carrier phase information is essential for sub-millimeter ranging measurements such as in a coherent optical link. When this phase is recovered, Doppler and range information become available for tracking similar to the current schemes with microwave signals, except with increased resolution due to the smaller wavelength of optical signals [4]. Micron-level range resolutions are readily achievable with a coherent optical link, which opens up new opportunities for performing scientific investigations based on the propagation of light signals. It also has direct implications for quantum-limited optical communication receivers.

ACKNOWLEDGMENTS

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology under a contract with the National Aeronautics and Space Administration.

REFERENCES

[1] Elmirghani, J.M.H, Cryan, R.A. "Analytic and numeric modelling of optical fibre PPM slot and frame spectral properties with application to timing extraction." IEE Proc., Commun. 141, 379, December 1994.

[2] Hemmati, H. "Deep Space Optical Communications." 1st ed. Hoboken, NJ: John Wiley & Sons, 2006.

[3] Dick, G.J, Tu, M., Strekalov, D., Birnbaum, K.,Yu, N. "Optical phase lock at femto-watt power levels for coherent optical deep space transponder." IPN Progress Report, 42-175, 1-17, 2008.

[4] Kinman, P.W., Gagliardi, R.M. "Doppler and range determination for deep space vehicles using active optical transponders" Applied Optics, 27, 4487–4493, 1988.

[5] Quirk, K.J. Gin, J.W. Srinivasan, M. "Optical PPM Synchronization for Photon Counting Receivers." IEEE Military Communications Conference, San Diego, CA. November 2008.

Whispering Gallery Mode Resonators as Optical Reference Cavities

Lukas Baumgartel[*,†], Rob Thompson[†], Dmitry Strekalov[†], Ivan Grudinin[†] and Nan Yu[†]
[*]Department of Physics, University of Southern California, Los Angeles, CA 90089. Email: lbaumgar@jpl.nasa.gov
[†]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.

***Abstract*—Highly stabilized lasers are an increasingly valuable tool for metrology. For many applications, however, existing Fabry Pérot systems are too bulky and cumbersome. We are investigating the use of miniature monolithic whispering gallery mode resonators as reference cavities for laser stabilization. We seek to exploit the benefit of small size and vibration resistance by suppressing thermally induced frequency fluctuations. We have theoretically investigated the viability of using a thin-film coating to achieve temperature compensation. We have experimentally investigated an active temperature stabilization scheme based on birefringence in a crystalline resonator. We also report progress of laser locking to the resonators.**

I. Introduction

Optical atomic clocks have attracted widespread interest due to their high stability and accuracy. These systems are facilitated by advances in laser technologies that enable optical trapping, cooling, and octave-spanning frequency combs, which provide a link from optical to microwave frequencies (see, e.g., [1]).

Limiting the short-term stability of optical frequency standards is the local oscillator – a highly stabilized laser capable of producing sub-hertz linewidths. Typical stabilization techniques utilize a Fabry-Pérot cavity fabricated from materials with extremely low thermal expansion coefficients. These cavities are further stabilized via operation under vacuum and with extensive thermal and vibration isolation. The net result is a large and complex local oscillator package.

Therefore, aiding the development of a portable and compact optical clock is a laser reference cavity that is both small and highly robust to mechanical vibrations and thermal fluctuations. At first glance, whispering gallery mode resonators (WGMRs) might not seem promising for such applications: because the light travels completely within bulk material, they are subject to higher levels of thermorefractive noise than is present in conventional Fabry-Pérot cavities. Furthermore, the most thermally stable materials, such as ultra-low expansion (ULE) and zerodur glass, are not suitable for WGMRs due to unacceptably high absorption at optical wavelengths. On the other hand, if the thermorefractive noise can be suppressed, WGMRs have potential to be high-performance cavities owing to their extremely high quality factor (Q) over a wide range of wavelengths, low sensitivity to mechanical noise, and small size, which enables precise and fast temperature control.

The dominant coupling mechanisms between temperature and resonant frequency in a WGMR are thermal expansion and thermorefraction. In case of the former, a change in resonator temperature leads to a change in its size – i.e., the cavity length – and a corresponding change in frequency. The latter refers to a temperature dependence of the host material's refractive index n, and is a quantity which varies widely from material to material. For WGMRs to be useful as reference cavities, both effects must be highly suppressed.

II. Thermal Compensation

We have explored two schemes for suppression of thermally induced frequency fluctuations: (1) a passive scheme whereby a material with negative thermorefractive coefficient is applied as a coating to a resonator to compensate for thermal expansion, and (2) an active scheme whereby a differential in frequency fluctuations between differently polarized modes in a birefringent crystalline resonator is used to stabilize its temperature.

A. Coated WGMR simulations

Coating a whispering gallery mode resonator with a thin-film of negative thermorefractive coefficient material could potentially be used to suppress the effects of thermal expansion. Such a compensation scheme was proposed by Han and Wang [2] and experimentally demonstrated by Jestin et al. [3], albeit with a resonator demonstrating relatively low Q. In such a heterogeneous resonator, the mode volume is shared between two materials so that part of the mode is contained in a medium demonstrating substantial negative thermorefraction (thermal expansion is always positive in the relevant materials).

We have performed simulations to investigate the viability of such a thin-film temperature compensation scheme following the procedure outlined in [2]. The method begins with a solution to Maxwell's equations over the geometry of the resonator, accounting for the different refractive index of resonator and coating (n_1 and n_2, respectively). The boundary condition demands continuity of tangential field components at resonator-coating and coating-air interfaces of the heterogeneous structure, yielding

$$\frac{1}{n_2}\frac{\chi_l'(z_4)}{\chi_l(z_4)} - \frac{C_l\psi_l'(z_3)+\chi_l'(z_3)}{C_l\psi_l(z_3)+\chi_l(z_3)} = 0 \tag{1a}$$

where the coefficient C_l is given by

$$C_l = \frac{\frac{n_2}{n_1}\psi_l(z_1)\chi_l'(z_2) - \psi_l'(z_1)\chi_l(z_2)}{\psi_l'(z_1)\psi_l(z_2) - \frac{n_2}{n_1}\psi_l'(z_2)\psi_l(z_1)}. \tag{1b}$$

978-1-61284-110-6/11/$26.00 ©2011 IEEE

The characteristic equation (1a) is derived in [4], and uses the following notation: The Riccati-Bessel functions are defined by $\psi_l(z) = zj_l(z)$ and $\chi_l(z) = zn_l(z)$, where l is the angular mode number; $j_l(z)$ and $n_l(z)$ are spherical Bessel and Neumann functions, respectively. The arguments are $z_1 = n_1 ka$, $z_2 = n_2 ka$, $z_3 = n_2 k(a+t)$, and $z_4 = k(a+t)$, and the prime denotes derivative with respect to the argument. Solution of equation (1a) depends on the radius of the disc, a, the thickness of the coating, t, and the refractive index of each material, all of which in turn depend upon temperature via

$$\begin{aligned} a &= a_0(1 + \alpha_l^{(1)}\Delta T) \\ t &= t_0(1 + \alpha_l^{(2)}\Delta T) \\ n &= n_0(1 + \alpha_n^{(i)}\Delta T) \end{aligned} \tag{2}$$

where $\alpha_l^{(i)}$ is the linear thermal expansion coefficient, $\alpha_n = (1/n_i)\partial n_i/\partial T$ is the thermorefractive coefficient, $i = 1, 2$ for the disc and coating, respectively, and the zero subscript indicates the initial ($\Delta T = 0$) value.

Thus, the temperature dependence of eigenfrequencies can be found from roots of (1a). We find these roots using a numerical solver in MATLAB, with an initial guess provided by asymptotic approximation. The derivatives are expanded using standard recursion relations (e.g., [5]). Furthermore, we make the approximation that both materials are non-magnetic and that the refractive index of air is identically 1. The characteristic equation has many roots for a given set of parameters; we choose the first one – corresponding to the "fundamental" whispering gallery mode.

First, we simulated the thermally induced frequency shift of a fused silica (SiO_2) resonator that has been coated with calcium fluoride (CaF_2). As shown in figure 1, we find that the coating has potential to substantially reduce the frequency shift when it is an optimal thickness. The shift is suppressed by about two orders of magnitude over a resonator where t varies $\pm$ 0.1 μm from this optimal value, and three orders of magnitude over the uncoated resonator (not shown for clarity). Furthermore, evidence of a turning point suggests the potential for extreme stability (see inset). However, a pitfall is the high sensitivity of the technique to coating thickness; an error of just 50 nm results in a ten-fold increase in thermal fluctuations, placing stringent tolerances on the fabrication process of such a device.

The simulations suggest that careful coating of an SiO_2 resonator could bring substantial stabilization. However, fused silica is an amorphous material. A crystalline resonator would offer several advantages, among them higher quality factor [6], and the possibility of birefringence (the benefit of which is described below). This motivates simulation of a crystalline magnesium fluoride (MgF_2) resonator coated with calcium fluoride. Such simulations suggest that the coating provides some improvement, though the gains are much less dramatic than in the case of the amorphous fused silica. Temperature induced frequency shift remains in the gigahertz per degree Celsius range, even for a coating of 5 μm (figure 2). Additionally, there does not exist a turning point for any reasonable combination of geometric factors and wavelength. This decreased improvement is due to the relatively low thermal expansion coefficient of SiO_2: 0.55×10^{-6}/°C [2] compared to 9.0×10^{-6}/°C for MgF_2 [7].

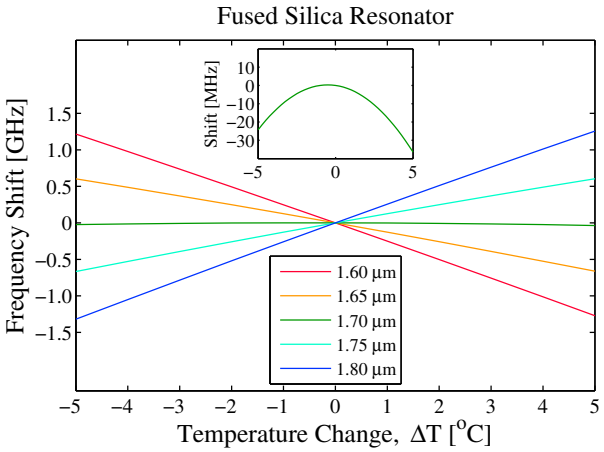


Fig. 1. Simulation results for an SiO_2 resonator coated with CaF_2. For a coating thickness of 1.70 μm, the frequency shift is considerably reduced. The inset has an expanded frequency axis for this optimal thickness, showing a turning point in thermal fluctuations.

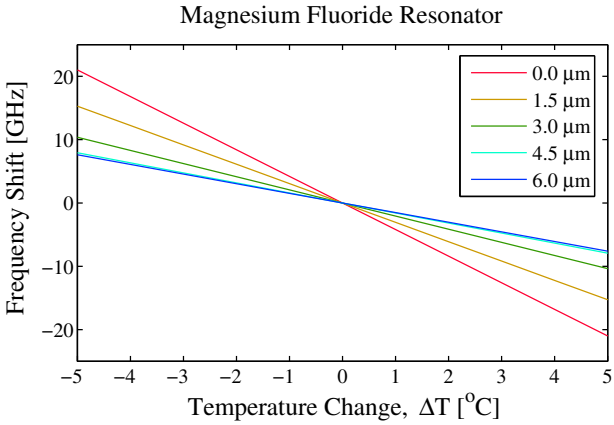


Fig. 2. Simulation results for a crystalline MgF_2 resonator with varying thicknesses of CaF_2 coating. Though some increase in temperature stability is predicted, the frequency shift remains monotonic and in the GHz/°C range.

The simulation results shown in figures 1 and 2 are performed for spherical resonators with a radius of 90 μm with light in the mode having a free-space wavelength of about 1 μm. This corresponds to an angular mode number l in the neighborhood of 800. While larger resonators can in principle be treated with the same simulation, the angular number gets very large if the 1 μm wavelength is maintained, exhausting computational resources.

B. Active Stabilization

We have developed an active temperature stabilization scheme that utilizes a differential in thermorefractive coefficients of the ordinary and extraordinary polarizations in

a birefringent resonator. Since the two polarizations have different thermal dependencies, a change in the frequency spacing between their modes corresponds to a change in the resonator's temperature. By selecting a TE and TM mode that are close together (in frequency) and measuring the difference between them in real time, one can measure temperature fluctuations. Feeding-back this measurement to a temperature control mechanism on or near the WGMR constitutes closed-loop temperature stabilization.

The experimental setup is as follows. Light from a fiber laser is coupled via an angle-polished fiber into a resonator made from magnesium fluoride. The laser frequency is swept over a few MHz, with a center frequency such that the two polarizations are simultaneously observable. Upon exiting the coupler, a polarizing beam splitter separates polarizations and each is detected with a separate photodiode (PD). Signals from the two detectors are acquired and processed digitally so that the frequency difference, Δf, between their peak centers can be precisely determined. The resulting error signal is passed through two proportional-integration stages to generate a correction signal. For temperature control, the resonator is placed in a house-made brass mount which contains two resistive heating elements, one above and one below the device, and in good thermal contact with it.

To predict the stabilization performance, we first determine the frequency differential's temperature dependence for a given, typical pair of modes. Such a pair is shown in figure 3. We make simultaneous, independent measurements of Δf and the resonator's temperature (with a precision thermistor) for several values of temperature. A fit to this data yields a value of $d(\Delta f)/dT \approx -90$ MHz/°C, a very sensitive function. Assuming we can make a shot-noise limited measurement of the mode locations (with a linewidth of 100 kHz), we can expect a frequency resolution of 0.1 Hz, corresponding to nK temperature resolution. This in turn yields a projected frequency stability of a few parts in 10^{14} [8].

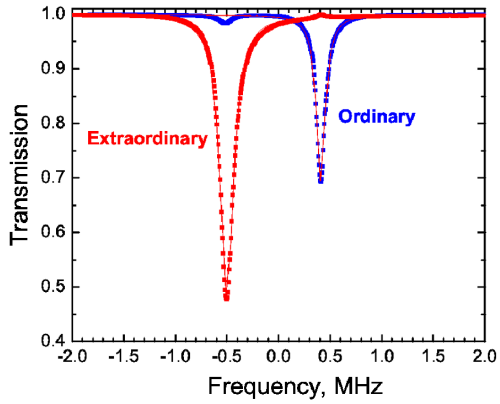


Fig. 3. Frequency spacing between the ordinary and extraordinary modes is a function of temperature and can be used for active temperature stabilization.

III. WGMR as a Reference Cavity

The ultimate goal of this work is to lock a laser to the temperature-stabilized WGMR. We have begun development of the locking system. Using the Pound-Drever-Hall (PDH) method [9], we have achieved preliminary stable locking. The magnesium fluoride cavity demonstrated Q=1.9×10^9 (corresponding to $\mathcal{F}$≈100,000), and is temperature stabilized by a conventional resistive heater and foam insulation.

The experimental setup is shown in figure 4 and described as follows. Light from a 1550 nm laser is coupled into the resonator via an angle-polished optical fiber, while a photodiode placed near the end of the fiber collects light from a single mode family. Our electronics employ a 1 MHz modulation frequency, which is split into two, with one portion being amplified and fed to an electro-optic phase modulator, while the other portion is used to demodulate the photodiode signal in an rf mixer. The resulting error signal (figure 5) is then fed to proportional-integrator-differentiator stages to generate the correction signal for the laser's piezoelectric tuning. So far, we have measured 3.2 kHz rms residuals over the loop bandwidth.

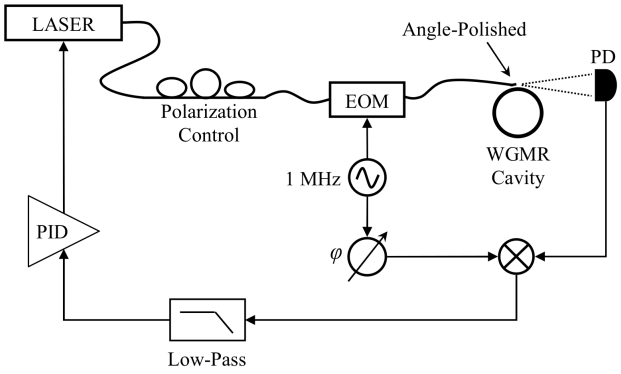


Fig. 4. Schematic of the experimental setup used for locking the Koheras fiber laser to the WGMR cavity. Locking has been achieved to an non-temperature stabilized resonator.

Fig. 5. Oscilloscope capture showing the photodiode signal (upper trace) and PDH error signal (lower trace). Sidebands are 1 MHz from the carrier.

IV. Conclusion

We have investigated temperature compensation schemes for whispering gallery mode resonators applied as frequency

reference cavities. Thermal compensation from a thin-film coating of material with negative thermorefractive coefficient is predicted to yield substantial improvements to an amorphous silica resonator, but insufficient improvement for a crystalline MgF_2 device. We note, however, that other material combinations may facilitate successful implementation with a crystalline resonator. A scheme utilizing the differential in thermorefractivity of differently polarized modes in a birefringent WGMR yields temperature resolution in the nK range and can be used for frequency stabilization. A cavity stabilized in this way has many advantages, chief amongst them being that it eliminates the need for thermal shielding, dramatically reducing the size of the apparatus. Our data support potential stabilization to a few parts in 10^{14} at 1 second. Furthermore, we believe that this technique is among the best performance of an active temperature control system operating near room temperature, and that it will have applications beyond laser stabilization.

V. Acknowledgment

This work was performed at Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract from NASA. Lukas Baumgartel acknowledges the College Doctoral Fellowship from University of Southern California.

References

[1] T. Udem, R. Holzwarth, and T. W. Hansch, "Optical frequency metrology," *Nature*, vol. 416, no. 6877, pp. 233–237, Mar 2002.

[2] M. Han and A. Wang, "Temperature compensation of optical microresonators using a surface layer with negative thermo-optic coefficient," *Opt. Lett.*, vol. 32, no. 13, pp. 1800–1802, Jul 2007.

[3] Y. Jestin, S. Berneschi, G. Conti, A. Chiappini, M. Ferrari, and G. Righini, "Micro resonator stabilization by thin film coating," in *Transparent Optical Networks, 2009. ICTON '09. 11th International Conference on*, 28 2009-july 2 2009, pp. 1 –4.

[4] I. Teraoka and S. Arnold, "Enhancing the sensitivity of a whispering-gallery mode microsphere sensor by a high-refractive-index surface layer," *J. Opt. Soc. Am. B*, vol. 23, no. 7, pp. 1434–1441, Jul 2006.

[5] G. Arfken and H. Weber, *Mathematical methods for physicists*. Elsevier, 2005.

[6] A. A. Savchenkov, A. B. Matsko, V. S. Ilchenko, and L. Maleki, "Optical resonators with ten million finesse," *Opt. Express*, vol. 15, no. 11, pp. 6768–6773, May 2007.

[7] A. A. Savchenkov, A. B. Matsko, V. S. Ilchenko, N. Yu, and L. Maleki, "Whispering-gallery-mode resonators as frequency references. ii. stabilization," *J. Opt. Soc. Am. B*, vol. 24, no. 12, pp. 2988–2997, Dec 2007.

[8] D. Strekalov, R. Thompson, L. Baumgartel, I. Grudinin, and N. Yu, "Temperature measurement and stabilization in a birefringent whispering gallery resonator," To be published.

[9] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator," *Applied Physics B: Lasers and Optics*, vol. 31, pp. 97–105.

Residual PM Noise Evaluation of Radio Frequency Mixers

C. A. Barnes, A. Hati, C. W. Nelson and D. A. Howe
National Institute of Standards and Technology
Boulder, CO U.S.A.

***Abstract*—Direct observation of phase-modulation (PM) noise is often difficult due to the high dynamic range that exists between the carrier and the modulated sidebands. A common tool used to reduce the dynamic range is the phase detector, which removes the carrier and down-converts its noise sidebands to baseband. The double balanced mixer (DBM) is the most widely used phase detector for high-resolution PM noise detection at most carrier frequencies. For Fourier offset frequencies close to the carrier, the residual flicker phase noise of the DBM is often the limiting factor of a PM noise measurement system. Careful evaluation of the phase detector under various operating conditions can lead to the optimization of a PM noise measurement system's sensitivity. This paper describes a survey of residual PM noise measurements for a variety of DBMs at 5 MHz. In order to attain quality measurements, careful attention is devoted to the reduction of ground loops during PM noise measurements. The input powers to the local oscillator (LO) and reference frequency (RF) ports of the mixers are varied to determine the optimal operating point of these devices.**

***Index Terms*— Flicker noise, mixers, phase detector, phase noise.**

I. INTRODUCTION

The measurement of close to carrier phase modulation (PM) noise of state-of-the-art oscillators is always challenging [1]. Quite often the residual noise of the phase detector used in these measurements is the source of difficulty [2], in particular, at Fourier offset frequencies below 100 Hz, where oscillator noise has a slope of f^{-3} and may be lower than the mixer noise floor, which follows a slope of f^{-1} [3]. The motivation for this paper is to locate mixers that offer low phase noise at 5 MHz for use in high resolution measurement systems. This paper describes a survey of PM noise measurements for 18 DBMs used as phase detectors [4, 5]. A study of the phase sensitivity of these detectors with operating power is also investigated. The measurement system used to characterize the detectors is described in section II, and the results of this survey are summarized in section III.

*This report summarizes the noise for various mixers by product name for completeness of this survey. No endorsements are implied.

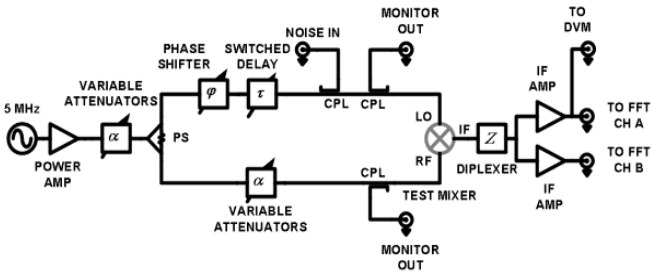


Figure 1. Block diagram of residual PM noise measurement system. PS: power splitter, CPL: directional coupler, IF AMP: low noise baseband amp, DVM: digital voltage meter, FFT: two-channel fast Fourier transform analyzer.

II. DESCRIPTION OF THE MEASUREMENT SYSTEM

The PM noise of the detectors for this survey is measured by use of a 5 MHz cross-correlated homodyne measurement system[6, 7, 8]. Fig. 1 shows a block diagram of the measurement configuration. The signal from a 5 MHz reference oscillator is amplified and split to drive the local oscillator (LO) and reference frequency (RF) ports of the detector. Variable attenuators are used to test the detectors at different power levels. A phase shifter is used to set a 90° quadrature condition between the LO and RF signals. In quadrature, the common-mode phase fluctuations of the oscillator and power amplifier cancel, and the residual noise of the detector is measured. The switched delay line and "Noise In" port are used to calibrate the system's phase sensitivity.

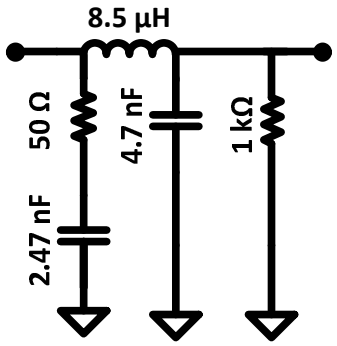


Figure 2. Mixer IF Diplexer schematic.

The detector is followed by a diplexer and a pair of baseband intermediate frequency (IF) amplifiers [9]. The diplexer, shown in Fig. 2, terminates and filters the high-frequency products of the mixing, while still maintaining a high voltage-to-phase conversion at baseband. This diplexer also allows for the baseband signals to be amplified without saturating the IF amplifiers with the 5 MHz and 10 MHz signals from the mixing process. The topology of the diplexer at the output of the phase detector plays a very important role in its performance [10, 5]. The pair of IF amplifiers is utilized to minimize the noise contribution of the fast Fourier transform (FFT) analyzer to the measurement. The uncorrelated input voltage noise of each IF amplifier is reduced by the use of cross-correlation. This allows measurement of the correlated residual detector noise [7]. This noise contribution of the IF amplifiers to the measurement is shown in Fig. 3.

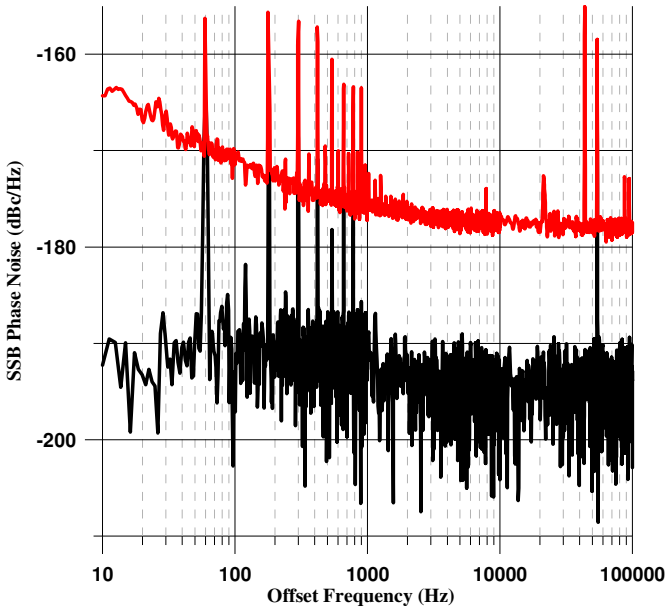


Figure 3 Plot shows that the PM noise of the phase detectors (Top) is higher than the noise floor of the IF amplifiers (Bottom).

In phase quadrature, the filtered output voltage of the detector is represented by

$$v = k_d \cos(\Delta\varphi + \pi/2) \quad (1)$$

where k_d is the mixer sensitivity in volts/rad and $\Delta\varphi$ is the differential phase deviation from quadrature between the LO and RF ports. For small phase deviations, k_d can be determined from a known phase shift introduced by the calibrated delay line. k_d is the ratio of DC voltage change to the introduced phase shift. This method of k_d calibration was also checked against single-sideband modulation and noise-injection methods [11, 12].

III. Measurement Results

For this survey 18 types of phase detectors have been compared. Each mixer was evaluated at two operating points.

1. Linear operation
   LO power = Nominal specified LO power
   RF power = 1 dB below the RF compression point.
2. Saturated operation
   LO power = Maximum mixer power 11 dBm.
   RF power = LO power.

Fig. 4 shows the range of PM noise for the measured devices on two separate plots. Due to the large number of devices evaluated, the measured residual PM noise of the detectors is reported at two offset frequencies; 10 Hz for specifying flicker PM noise and 100 kHz indicating thermal PM noise. The top and bottom plots show the 10 Hz and 100 kHz data respectively for each detector. Each offset frequency is displayed as a line with endpoints representing the linear and saturated measurement values.

Often the PM noise of the phase detectors decreased significantly when they were operated in saturated mode. As discussed in [1, 5], this is attributed mostly to the higher sensitivity, k_d, in saturated operation. For this survey, mixers with a nominal LO rating of +7 dBm and maximum rating of +17 dBm are classified as low-power. Mixers with a nominal LO power rating above +7 dBm are classified as high-power for this study. It can be seen in Fig. 4 that the variation of measured PM noise is greater in high-powered DBMs than for the low-power mixers. Of the commercial detectors, the HP10534A* has the lowest flicker and the ZAD-1H+* the lowest white noise in linear operation. The MCL ZRPD-1+* showed the largest improvement in its white noise between linear and saturated operation, and has the lowest white noise level overall. The PM noise of the best detectors surveyed in saturated mode are presented in fig. 5. The lowest level of flicker noise observed is in a custom mixer built and described in the following section.

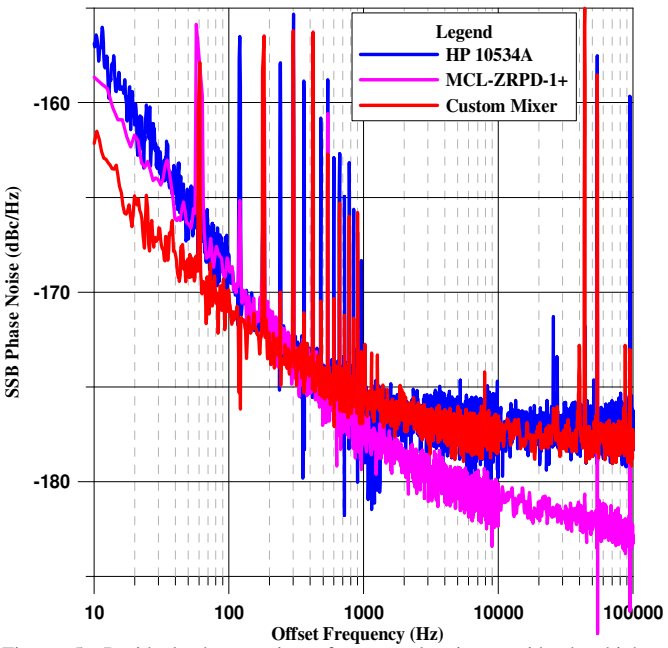


Figure 5. Residual phase noise of surveyed mixers with the highest performance in saturated operation.

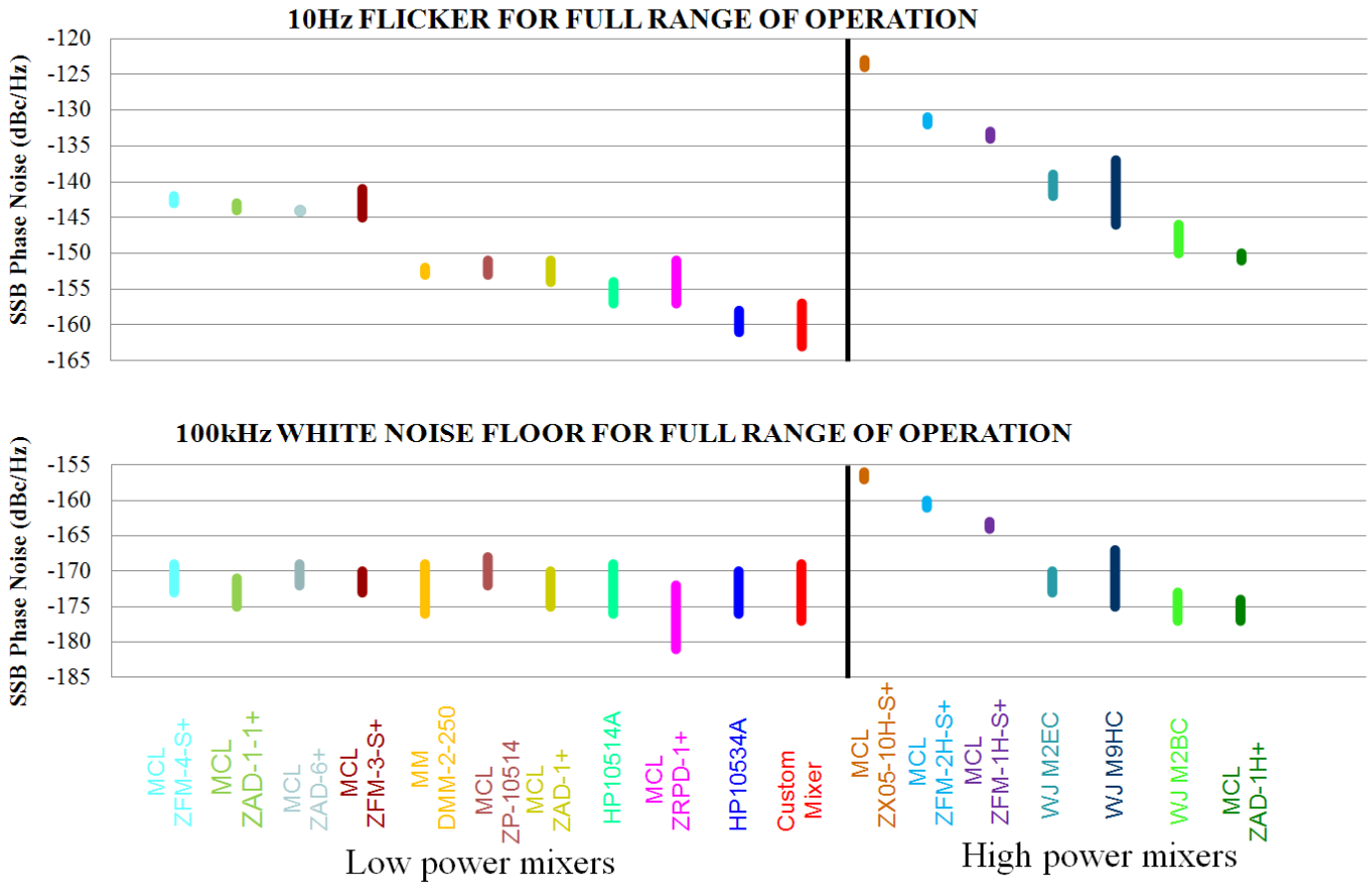


Figure 4. Residual PM noise summary of various phase detectors (endpoints indicate linear and saturated measurement limits).

IV. Custom Mixer Design

A. *Mixer Design*

The custom-built mixer listed in the summary of measurements was a simple double-balanced mixer with 2N2222A transistors in the diode ring. This design is discussed in a letter recently submitted for publication [13]. The custom detector is noteworthy because its flicker performance was the best in the detectors surveyed. Fig. 6 shows the general topology of the DBM. In this design there are four 2N2222A bipolar junction transistors (BJTs) used to construct a conventional double balanced diode ring.

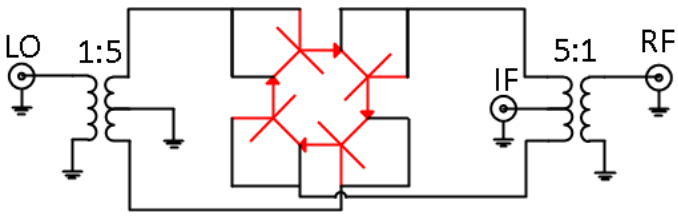


Figure 6. Double balanced mixer schematic. The diode ring in this double-balanced mixer is constructed by use of transistors with the collector input tied to the base.

We short the base to the collector of these BJTs and operate them as diodes by use of the emitter-base junction. The transformers used for this design are commercial off-the-shelf parts and have a 1:5 impedance ratio. This impedance ratio is chosen so that the input impedance of the reference frequency (RF) and local oscillator (LO) ports are nearly 50 Ω at 5 MHz. The mixer's input impedances are measured by use of the Smith chart display of a vector network analyzer.

B. *Mixer Characterization*

In order to measure the PM noise of the DBM, the nominal operating powers need to be found [3]. To determine the nominal LO power, a 4 MHz sinusoidal signal at -30 dBm is applied to the RF port of the mixer while the power of a 5 MHz signal at the LO port is varied. Conversion loss of the mixer is calculated by taking the power ratio of the 1 MHz beat at the intermediate frequency (IF) port to the RF signal at 5 MHz. A plot of conversion loss versus LO power is shown in Fig. 7. By use of this plot, the 1 dB compression point of the LO drive is found to occur at a level of +8 dBm. Operating the LO port in saturation is desired to reduce LO power fluctuation sensitivity. Taking the maximum allowed operating specifications of the transistors and transformers into account, a nominal LO drive of +11 dBm is selected. An

additional measurement of conversion loss versus RF power at the selected nominal LO is made and plotted in Fig. 8. The 1 dB compression point of conversion loss at the RF port occurs at +5 dBm. All conversion loss tests for the mixer were conducted with a 50 Ω load at the IF port.

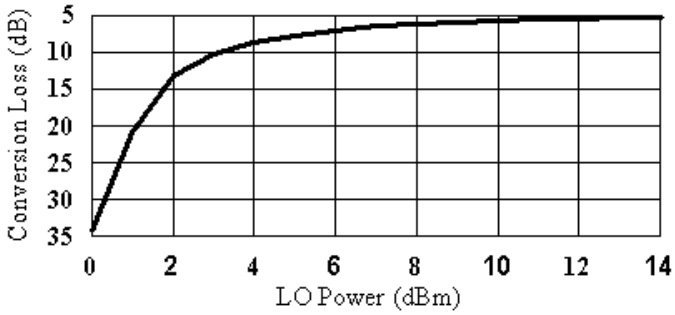


Figure 7. Conversion loss versus LO power of mixer with a -30 dBm RF power level. RF @ 4 MHz, LO @ 5 MHz, 50 Ω input impedance to FFT

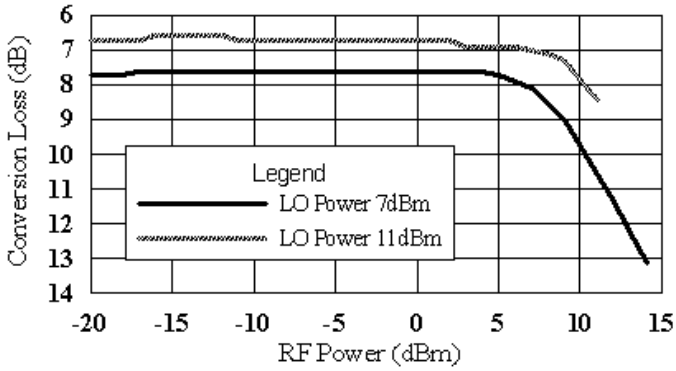


Figure 8. Conversion loss vs. RF signal power at constant LO drive powers. LO @ 4 MHz, RF @ 5 MHz, 50 Ω input impedance to FFT

Fig. 9 shows the results of the residual PM noise of this DBM design used as a phase detector. From these data it can be seen that the flicker noise of this device is L(10 Hz) = -161 dBc/Hz when the DBM is operated at LO = +11 dBm and RF = +5 dBm. In deep saturation (LO = RF = +11 dBm) the 10 Hz noise is improved to L(10 Hz) = -163 dBc/Hz. Both of these measurements show that this DBM has very low flicker noise as a phase detector.

V. Conclusion

We evaluated a number of phase detectors available in our lab at 5 MHz. A 40 dB range of flicker PM noise is observed among the surveyed detectors. Similar variation is also observed in the level of thermal PM noise. High-power mixers in general showed a higher level of flicker noise than that of the lower power detectors. We have found that for many phase detectors, the lowest noise floor is achieved while the device is operated in saturation. Since the sensitivity of the phase detectors depend on the impedance at the IF port, a further study of PM noise versus output termination design will be considered for future work. Specifically we will investigate the phase noise by use of non-absorptive terminations. Further investigation at other carrier frequencies will be conducted as well.

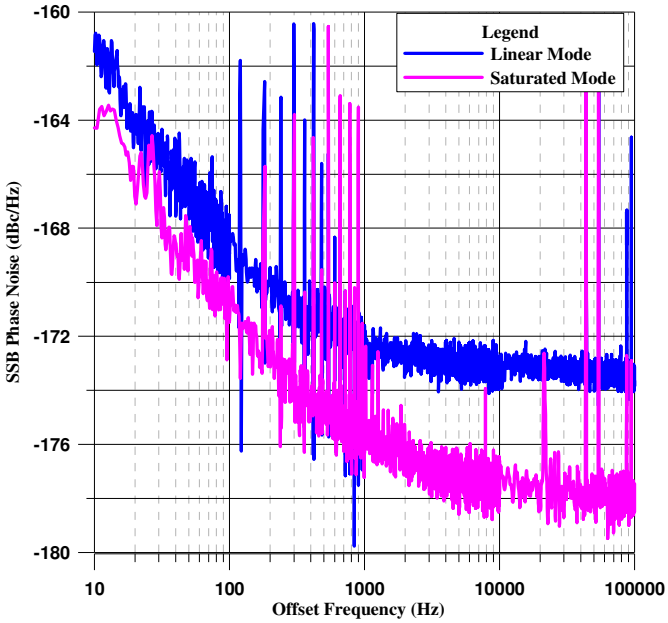


Figure 9. Cross-correlated residual PM noise floor of the 2N2222A based mixer in linear and saturated operation.

Acknowledgments

We thank Jesse Searle for the idea of constructing our own mixers, Enrico Rubiola for an IF amplifier design used in the noise floor measurements described in this paper [8], and Jason DeSalvo for theoretical discussions regarding measurement systems.

References

[1] E. Rubiola, *Phase Noise and Frequency Stability in Oscillators*. Cambridge University Press, 2010.

[2] U. Mavric and B. Chase, "Residual Phase Noise Measurements of the Input Section on a Receiver," *Microwave Journal*, no. March, pp. 94-102, 2008.

[3] L. M. Nelson, C. W. Nelson, and F. L. Walls, "Relationship of AM to PM noise in selected RF oscillators," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 41, no. 5, pp. 680-684, Sep. 1994.

[4] S. R. Kurtz, "Mixers as Phase Detectors," *Watkins-Johnson Tech-Notes*, vol. 5, no. 1, 1978.

[5] E. Rubiola, "Tutorial on the double balanced mixer," *physics/0608211*, Aug. 2006.

[6] F. L. Wall and E. Ferre-Pikal, "Measurement of Frequency, Phase noise and Amplitude noise," in *Wiley Encyclopedia of Electrical and Electronics Engineering*, 1st ed., vol. 12, 24 vols., Wiley-Interscience, 1999, pp. 459-473.

[7] W. F. Walls, "Cross-correlation phase noise measurements," in *Frequency Control Symposium, 1992. 46th., Proceedings of the 1992 IEEE*, 1992, pp. 257-261.

[8] G. K. Montress, T. E. Parker, and M. J. Loboda, "Residual phase noise measurements of VHF, UHF, and microwave components," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 41, no. 5, pp. 664-679, Sep. 1994.

[9] E. Rubiola and F. Lardet-Vieudrin, "Low flicker-noise amplifier for 50 Ω sources," *Review of Scientific Instruments*, vol. 75, no. 5, p. 1323, 2004.

[10] F. L. Walls and S. R. Stein, "Accurate Measurements of Spectral Density of Phase Noise in Devices," in *Proc. 3lst Ann. Frequency Control Symp.*, Atlantic City, NJ, 1977, pp. 335-343.

[11] F. L. Walls, “Correlation between upper and lower sidebands,” *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 47, no. 2, pp. 407-410, 2000.

[12] F. L. Walls, “Secondary standard for PM and AM noise at 5, 10, and 100 MHz,” *IEEE Transactions on Instrumentation and Measurement*, vol. 42, no. 2, pp. 136-143, Apr. 1993.

[13] C. A. Barnes, A. Hati, C. W. Nelson, and D. A. Howe, “Phase Noise Detector with Very Low Flicker at 5 MHz.,” *Submmited to Electronics Letters*.

A Flight-Like Optical Reference Cavity for GRACE Follow-on Laser Frequency Stabilization

R. Thompson, W. M. Folkner, G. deVine,
W. M. Klipstein, K. McKenzie, R. Spero, N. Yu
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109
E-mail: Robert.J.Thompson@jpl.nasa.gov

M. Stephens, J. Leitch, R. Pierce
Ball Aerospace and Technologies Corporation
PO Box 1062, Boulder, CO 80306-1062

T. T.-Y. Lam, D. A. Shaddock
Australian National University
Canberra, Australian Capital Territory 0200, Australia

***Abstract*—We describe a prototype optical cavity and associated optics that has been developed to provide a stable frequency reference for a future space-based laser ranging system. This instrument is being considered for inclusion as a technology demonstration on the recently announced GRACE follow-on mission, which will monitor variations in the Earth's gravity field.**

I. Introduction

Launched in 2002, the Gravity Recovery and Climate Experiment (GRACE) mission [1] monitors changes in the Earth's gravity field by measuring changes in the distance between two spacecraft induced by that variable field. The distance variation is measured with a microwave ranging system with sub-micron precision. [2] The ranging measurement accuracy is limited by the signal-to-noise ratio and by the frequency stability of the microwave signal referenced to an ultra-stable oscillator (USO). [3] For the upcoming GRACE follow-on mission a laser ranging system with precision better than the GRACE microwave ranging system is under consideration as a technology demonstration. A laser ranging system easily provides improved signal-to-noise ratio over the microwave system. [4] Laser frequency stability better than the GRACE USO stability has been demonstrated in several laboratories using thermally stabilized optical reference cavities. We are developing a space-qualifiable laser frequency stabilization system for the proposed GRACE follow-on mission. The system includes a ULE optical cavity with the associated optics and electronics and will be capable of providing a fractional frequency stability of better than 10^{-13} for averaging times from 0.01 to 1 s.

Jet Propulsion Laboratory, California Institute of Technology

II. Cavity and Thermal Enclosure Design

The goal for this development is to achieve a frequency noise power spectral density of less than 30 Hz/ $\sqrt{}$Hz over the frequency range of interest, which is 10 mHz to 100mHz. For an orbital speed of 7 km/sec, this frequency range corresponds to distance scales on the Earth's surface of 350 km to 35 km. Below this frequency range accelerometer noise is expected to dominate the measurements, while measurements at higher frequencies are limited by data sampling.

The cavity is based on a design available from Advanced Thin Films, Inc., which has been used successfully in a number of laboratories around the world. [5,6] It consists of a 77.5 mm ULE spacer to which high reflectivity mirrors are optically contacted. The mirror substrates are also fabricated from ULE glass. The design finesse was 10000. The cavity is mounted to a vacuum flange using titanium flexures, which are bonded to the cavity spacer. The flange is attached to a titanium vacuum enclosure which forms the first stage of a two stage thermal isolation system. Laser light is injected into the cavity via an optical fiber, which is mounted along with the required mode-matching optics on a zerodur optical bench. The bench, in turn, is mounted to the cavity via additional titanium flexures. A 90/10 beamsplitter picks off a portion of the light reflected from the cavity, which is steered through a window on the bottom of the vacuum enclosure to a photodiode. Figure 1 shows the reference cavity and optical bench mounted on the vacuum flange, while Figure 2 shows the sealed thermal enclosure, which is mounted, again with titanium flexures to a baseplate. Also shown is the outer thermal shield. The outer surface of the vacuum can has been coated with gold to reduce radiative heat transfer. Heaters and RTD temperature sensors mounted directly to the outer shield,

978-1-61284-110-6/11/$26.00 ©2011 IEEE

and also to an additional baseplate on which the entire cavity system is mounted are used to control the temperature at each end of the enclosure.

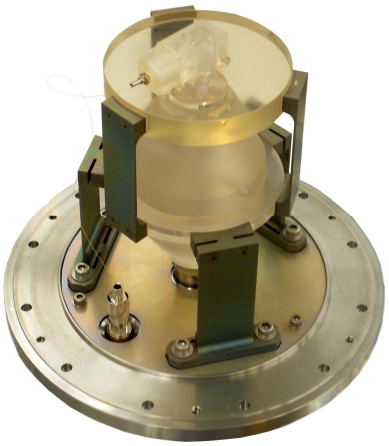


Figure 1. Reference cavity and optical bench mounted to a vacuum flange

Figure 2. Reference cavity assembly with outer thermal shield removed

III. PERFORMANCE TESTING IN A SIMULATED SPACE ENVIRONMENT

A neodymium:YAG laser with a non-planar ring oscillator (NPRO)configuration was used for performance tests of the cavity design. The laser was locked using a standard Pound Drever Hall technique [8], whereby the laser light is modulated (in this case at 5 MHz) and phase sensitive detection of the light reflected from the cavity generates the phase error signal as the frequency discriminator. A sketch of the setup is shown in Figure 3. A portion of the laser output was split off to beat with the output of a second, identical, laser, which was locked, to a second cavity in an independent vacuum chamber. The beat signal was detected on a high bandwidth photodetector and then mixed with a stable rf source to provide a signal within the 20 MHz bandwidth of a high accuracy digital phasemeter developed originally for GRACE-2 [9] and further developed for the LISA instrument [10]. Results of these performance tests are described in [7].

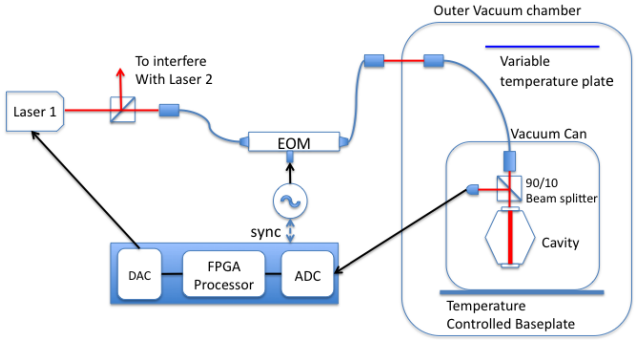


Figure 3. Test configuration for laser stabilization in simulated space environment

During space operations, the cavity will need to meet its performance requirements in a varying thermal environment, as the spacecraft flies in and out of the sun over a ninety-minute orbit. While the cavity will be mounted to a baseplate which is temperature controlled to within 0.1°C, some spacecraft surfaces will experience a temperature variation of not more than 1°C. To simulate this, the spacecraft is placed in a large vacuum chamber containing an 8" by 10" sheet of aluminum with a heater and temperature sensor attached. A function generator drives the heater such that the plate temperature varies by 2 °C over ninety minutes, a variation which is greater than that expected on the spacecraft. During these tests the thermal control was applied to the baseplate under the cavity only. Comparisons of the cavity under these conditions to a second independent cavity are shown in Figure 4. The results meet the current requirements for the GRACE follow-on mission.

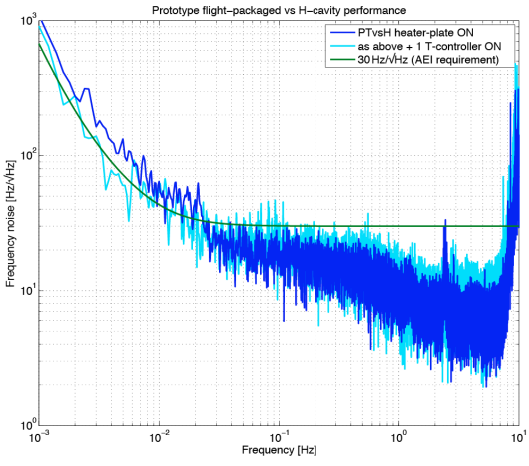


Figure 4. Power spectrum of difference frequencies of lasers locked to prototype and to a second cavity. The prototype is subjected a varying thermal environment as described in the text. Dark blue trace has no thermal

control applied to the prototype, light blue: with thermal control. Green: 30 Hz/√Hz requirement.

IV. FUTURE WORK

The prototype cavity assembly will be subject to thermal cycling and launch vibration tests in order to establish its readiness for consideration for a future space flight. A complete system test including an Interferometric Range Transceiver (IRT) developed under a previous IIP [4] will be performed.

ACKNOWLEDGMENT

This work was sponsored by the NASA Earth Science Instrument Incubator Program. This research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. T. T.-Y. Lam and D. A. Shaddock were supported under the Australian Government's Advanced Space Research Program.

REFERENCES

[1] B. Tapley, J. Ries, J, S. Bettadpur, D. Chambers, M. Cheng, F. Condi, B. Gunter, Z. Kang, P. Nagel, R. Pastor, T. Pekker, S. Poole, F. Wang, "GGM02 - An improved Earth gravity field model from GRACE", Journal Of Geodesy , vol. 79, pp. 467-478, 2005.

[2] C. Dunn, J. Kim, Y. Bar-Sever, S. Desai, B. Haines, D. Kuang, G. Franklin, I. Harris, G. Kruizinga, T. Meehan, S. Nandi, D. Nguyen, T. Rogstad, J. B. Thomas, J. Tien, L. Romans, M. Watkins, S.-C. Wu, S. Bettadpur, W. Bertiger, "Instrument of GRACE", GPS World, February 2003.

[3] J.B. Thomas, "An Analysis of Gravity Field Estimation Based on Intersatellite Dual Biased One-Way Ranging" JPL Pub. 98-15 May 1999, p B-3.

[4] M. Stephens, R. Craig, J. Leitch, and R. Pierce, R.S. Nerem, P. Bender, and B. Loomis, "Interferometric Range Transceiver for Measuring Temporal Gravity Variations", proceedings of the 2006 Earth Science Technology Conference, College Park, MD, 2006.

[5] J. Alnis, A. Matveev, N. Kolachevsky, Th. Udem, and T. W. Hänsch, Subhertz linewidth diode lasers by stabilization to vibrationally and thermally compensated ultralow-expansion glass Fabry-Pérot cavities, Physical Review A vol. 77, 053809-9, 2008.

[6] M. Nothcutt, L. S. Ma, J. Ye, J. L. Hall, "Simple and compact 1-Hz laser system via an improved mounting configuration of a reference cavity", Optics Letters, voi. 30, pp. 1815-1817, 2005.

[7] W. M. Folkner, G. deVine, W. M. Klipstein, K. McKenzie, R. Spero, R. Thompson, N. Yu, M. Stephens, J. Leitch, R. Pierce, D. Shaddock,T. Lam, "Laser frequency stabilization for GRACE-II", NASA Earth Science Technology Forum 2011, June 21-23, Pasadena, CA,

[8] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, "Laser phase and frequency stabilization using an optical resonator", Appl. Phys. B, vol. 31, pp. 97-105, 1983.

[9] B. Ware, W. M. Folkner, D. Shaddock, R. Spero, P. Halverson, I. Harris, T. Rogstad, "Phase Measurement System for Inter-Spacecraft Laser Metrology, proceedings of the 2006 Earth Science Technology Conference, College Park, MD, 2006.

[10] D. Shaddock, B. Ware, P. G. Halverson, R. E. Spero, and B. Klipstein, "Overview of the LISA Phasemeter", 6th International LISA Symposium, AIP Conference Series, vol. 873, pp. 654-660, 2006.

Improvement in Accuracy of a Single $^{40}Ca^{+}$ Optical Clock towards the 10^{-15} Level using a Magnetic Shield

Kensuke Matsubara, Ying Li, Shigeo Nagano,
Hiroyuki Ito, Masatoshi Kajita, Reiko Kojima,
Yuko Hanado, Kazuhiro Hayasaka,
and Mizuhiko Hosokawa
Space-Time Standards Group
National Institute of Information and Communications Technology
Koganei, Tokyo, Japan
matubara@nict.go.jp

***Abstract*— We report a frequency measurement of a single $^{40}Ca^{+}$ optical clock towards the 10^{-15} uncertainty level. In contrast with our previous measurement, in which the accuracy was limited by line broadening due to the magnetic field fluctuation, the use of a magnetic shield significantly reduced the spectrum linewidth of the clock transition. By a frequency comparison to an ^{87}Sr lattice clock, an Allan deviation of 7×10^{-16} was observed at an averaging time of 10^3 seconds. We preliminarily measured the clock transition frequency to be 411 042 129 776 397 (3) Hz with respect to the SI definition of the second. Further measurements are executed aiming at an accuracy of low parts in the 10^{-15} level.**

I. INTRODUCTION

As candidates for next generation definitions of the second, various efforts are being made to develop optical clocks with the uncertainty better than that of the cesium standards. The best uncertainty of 8.6×10^{-18} has been reported in a composite quantum system using $^{27}Al^{+}$ ions [1]. However, this technique requires sophisticated quantum logic operation and it is not easily accessible. Simpler clocks with modest accuracies have been also developed using Alkali-metal like ions ($^{88}Sr^{+}$, $^{177}Yb^{+}$). We are developing an optical clock using single $^{40}Ca^{+}$ ions [2]. The energy levels of $^{40}Ca^{+}$ that are relevant to the operation of the frequency standard are shown in Fig. 1. The lifetime of the $^2D_{5/2}$ state is about 1.2 seconds, which gives a high line Q to the $^2S_{1/2}$ - $^2D_{5/2}$ transition at 729 nm (411 THz). In 2008 we reported the absolute frequency of this transition with an uncertainty of 4×10^{-14}. In the next year, it was also reported by M. Chwalla *et al*. with an improved uncertainty of 10^{-15} level [3]. These two reports contributed to the acceptance of the $^{40}Ca^{+}$ transition to the 2009 CCTF list of recommended radiations. We measured this frequency to be 411 042 129 776 395 (5) Hz in 2009 [4]. These results are consistent with each other within the uncertainties. This paper reports on our improved frequency measurement towards the 10^{-15} uncertainty level, as well as the frequency stability of the clock laser locked to the $^{40}Ca^{+}$ clock transition.

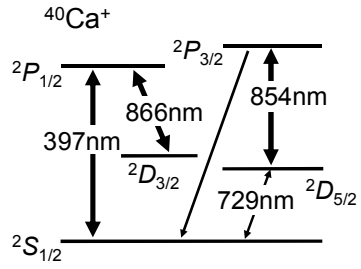


Figure 1. Partial term energy diagram of $^{40}Ca^{+}$ ions.

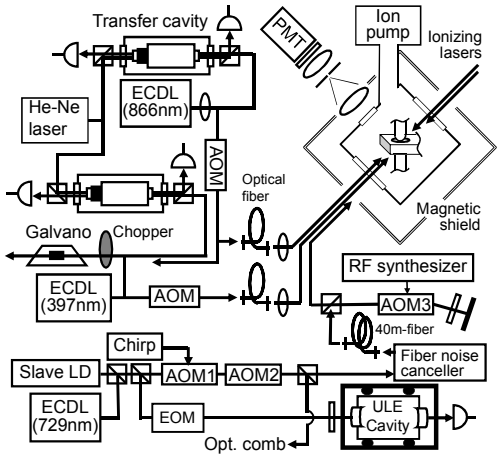


Figure 2. Experimental setup for Ca^{+} ions. AOM: acousto-optic modulator; EOM: electro-optic modulator; PMT: photo-multiplier tube; Galvano: Galvano tube. Roles of AOM1~3 are described in the text.

In this paper, we first outline the experimental setup, and then explain the clock transition measurement, frequency lock to the clock transition, and absolute frequency measurement. Finally, a conclusion is described.

II. Experimmental Setup

The experimental setup, which is based on that described in our previous paper, is shown in Fig. 2 [2]. The output of the 729-nm clock laser is distributed to the ion-trap chamber using a 40-m PM optical fiber. In order to minimize the phase noise due to the fiber transfer, we used a cancellation system [5]. The clock transition spectrum is observed by the electron shelving method [6] with a measurement sequence using laser pulses [2]. A stable dc magnetic field is used to resolve the clock transition into ten Zeeman components. In our previous measurement, the spectrum was broadened to a few hundred hertz mainly due to the ambient magnetic field fluctuation. Therefore, we have developed an ion trap chamber equipped with a two-layer magnetic shield. This shield decreases the amplitude of the ambient magnetic field by a factor of more than 25.

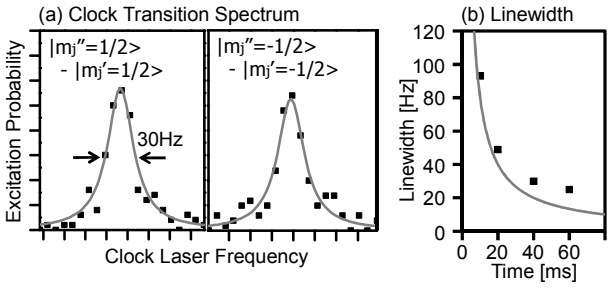


Figure 3. (a) Typical clock transition spectra and their Lorentz fitting. (b) Linewidth as the function of interrogating time of the clock laser. Gray line shows theoretical transit-time broadening using rectangular light pulses.

A typical transition spectrum is shown in Fig. 3 (a). It shows a pair of the Zeeman components of $|{}^2S_{1/2}, M_J=1/2\rangle$ - $|{}^2D_{5/2}, M_J=1/2\rangle$ and $|{}^2S_{1/2}, -1/2\rangle$ - $|{}^2D_{5/2}, -1/2\rangle$ transitions. Averaging of these two transition frequencies cancels the first order Zeeman shift. The linewidth is about 30 Hz when the interrogation time of the trapped ion by the 729-nm light was 40 ms. The measured linewidth at various interrogation times is shown in Fig. 3 (b). It was mainly limited by the transit-time broadening by rectangular 729-nm light pulses. Introducing the magnetic shield, the observed linewidth has been decreased to be about one digit smaller than that previously observed.

III. Frequency Stability Measurement

To certify the frequency stability of the $^{40}Ca^+$ optical clock, we stabilized the 729-nm laser frequency by the $^{40}Ca^+$ clock transition. The frequency was locked to the average of the two Zeeman components of the $|{}^2S_{1/2}, M_J=1/2\rangle$ - $|{}^2D_{5/2}, M_J=1/2\rangle$ and $|{}^2S_{1/2}, -1/2\rangle$ - $|{}^2D_{5/2}, -1/2\rangle$ pair, which was shown in Fig. 3(a). At the first we measured the stabilized frequency using an optical comb based on a femto-second pulse laser. A cryogenic sapphire oscillator (CSO) was used as the frequency reference. And next, we measured the frequency of the $^{40}Ca^+$ clock in comparison with that of an ^{87}Sr optical lattice clock.

In the short term, the clock laser frequency at 729 nm is stabilized using an ultra-low-expansion (ULE) optical cavity, which has a high finesse of about 2×10^5. Because of the material aging, however, its resonance frequency has a slow long-term frequency drift with a rate of about 0.03 Hz/s. In addition, very small thermal fluctuation of the optical cavity causes a non-linear frequency drift. The long-term linear drift was reduced using a frequency chirp with an acousto-optic modulator (AOM1 in Fig. 2). The residual drift was detected by the spectrum observation with the second AOM (AOM3). The residual drift was evaluated every 25 seconds and it was compensated using the third AOM (AOM2). Figure 4 (circles) shows the Allan deviation of the frequency stability measured using the CSO as the frequency reference. We believe that the Allan deviation of 1×10^{-15} at an averaging time of 10^3 seconds was limited by the stability of the CSO.

To measure the stability of the $^{40}Ca^+$ clock at the averaging time of more than 10^3 seconds, we compared the stabilized laser frequency to a 698-nm laser frequency that was locked to the ^{87}Sr clock transition. The ^{87}Sr clock was developed using the optical lattice technique [7]. The measured Allan deviation is also shown in Fig. 4 (squares). An Allan deviation of 7×10^{-16} at 10^3 s was measured. The Allan deviation was limited to about 5×10^{-15} at 10^3 s when we measured the stability without magnetic shielding. An improvement in the stability of about one digit over our previous measurements was attained.

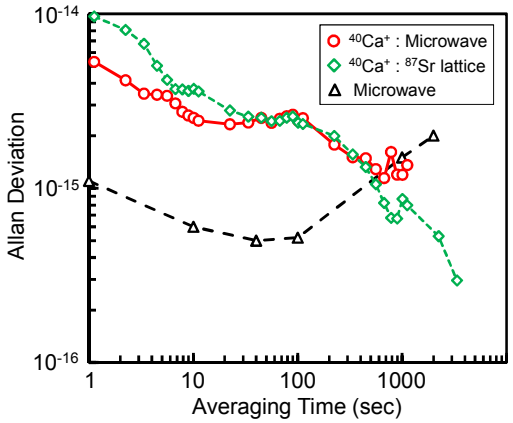


Figure 4. Frequency stability evaluated by Allan deviation between the $^{40}Ca^+$ clock, the ^{87}Sr lattice clock and the microwave frequency standard. Circles (red): Allan deviation between the $^{40}Ca^+$ clock and the microwave standard: Squares (green): between the $^{40}Ca^+$ and the ^{87}Sr clock; Triangles (black): Allan deviation of the cryogenic sapphire oscillator measured by the developer (University of West Australia).

IV. Absolute Frequency Measurement

The absolute frequency of the $^{40}Ca^+$ clock transition was evaluated from the laser frequency locked to the $^{40}Ca^+$ clock. It was measured by an optical comb with a hydrogen maser as the frequency reference. Here, the frequency of the hydrogen maser is evaluated every second by the Coordinated Universal Time (UTC) maintained in NICT. Time difference between the UTC(NICT) and the SI definition of the second on the

geoid surface (TAI) was corrected using the Circular T by the BIPM.

Systematic frequency shifts were evaluated as follows. In order to avoid the ac Stalk shift caused by the 397- and 866-nm light powers, we used a measurement sequence using the laser pulses [2]. A faint leakage of the 397- and 866-nm light powers during the interrogation time by the 729-nm laser causes a large ac Stark shift. Therefore, using a combination of a mechanical shutter and an AOM one we fully blocked the 397- and 866-nm light power, and we checked a negligible amount of frequency shift for our measurement by measuring the clock transition with various 397- and 866-nm powers. Black body radiation (BBR) frequency shift of 0.4 (0.5) Hz was evaluated from the room temperature. From the geographical height of the ion trap from the geoid surface, the gravitational shift was evaluated to be 3.4 (0.1) Hz [8].

TABLE I. SYSTEMATIC SHIFTS AND UNCERTAINTIES

Physical effect and drift	*Correction* (Hz)	*Uncertainty* (Hz)
Gravitational shift	3.4	0.1
Black body radiation shift	0.4	0.5
Electric quadrupole shift	0	2
Quadratic Zeeman shift (3μT)	0	0.1
ULE cavity drift	0	0.5
Magnetic field strength drift	0	0.5

The electric quadrupole moment of the D_J state interacts with the electric field gradient, causing an additional shift of a few hertz [9]. The value of the quadrupole shift ($\Delta\nu_Q$) at the M_J sublevel of the D_J state is expressed using

$$\hbar\Delta\nu_Q = \frac{dE_z}{dz}\Theta(D, j)\frac{J(J+1)-3M_J^2}{4J(2J-1)}(3\cos^2\beta-1). \quad (1)$$

where E_Z/d_Z is the electric field gradient along the symmetry axis of the trap potential (z), β is the angle between the axis z and the magnetic field, and $\Theta(D, j)$ is the strength of the quadrupole moment in terms of a reduced matrix element. From Eq. (1), the ratio of the quadrupole shift of the $|^2D_{5/2}, M_J = \pm1/2\rangle$ levels to the $|^2D_{5/2}, \pm3/2\rangle$ level and $|^2D_{5/2}, \pm5/2\rangle$ level is 4:1:-5. Therefore, the quadrupole shift can be canceled by averaging the transition frequencies of six Zeeman components of the $|^2S_{1/2}, \pm1/2\rangle$ - $|^2D_{5/2}, \pm1/2\rangle$ ($\Delta M_J = 0$ transition), $|^2S_{1/2}, \pm1/2\rangle$ - $|^2D_{5/2}, \pm3/2\rangle$ ($\Delta M_J = 1$), and $|^2S_{1/2}, \pm1/2\rangle$ - $|^2D_{5/2}, \pm5/2\rangle$ ($\Delta M_J = 2$). The uncertainty of 2 Hz was estimated from the standard deviation when we fitted the measured quadrupole shifts to the theoretical shift ratio of 4:1:-5. We considered other contributions to the transition frequency shift, such as the quadratic Zeeman shift, which were listed in TABLE I. The total correction was 3.8 (2.2) Hz.

We measured the absolute transition frequency on two separate days in this year, 16 February and 3 March in 2011. In each of the measurement, after observation of the transition spectrum, we stabilized the clock laser frequency to the average of the $|^2S_{1/2}, \pm1/2\rangle$ - $|^2D_{5/2}, \pm1/2\rangle$ transition pair. And then, we measured the clock laser frequency for about one hour by the optical comb. The Allan deviation was smaller than 3×10^{-15} at the averaging time of 10^3 seconds. Therefore it was considered that the long-term drift of the systematic shift including the quadrupole shift was smaller than 2 Hz for 10^3 seconds. The stabilized frequency was also measured for each of the $|^2S_{1/2}, \pm1/2\rangle$ - $|^2D_{5/2}, \pm3/2\rangle$ transition pair and the $|^2S_{1/2}, \pm1/2\rangle$ - $|^2D_{5/2}, \pm5/2\rangle$ one. The average frequency of these three measurements was used to evaluate the transition frequency. It was compensated for both of the first order Zeeman shift and the electric quadrupole shift.

The measured frequency was corrected using the Circular T. For the measurement on 16 February, we evaluated a correction of 0.1 (2.0) Hz using Circular T 278. For the measurement on 3 March, we evaluated a correction of 2.0 (2.0) Hz using Circular T 279. The uncertainty was estimated from the link uncertainty between the UTC(NICT) and the TAI. From the measured frequency, the correction using Circular T, and the shifts in TABLE I, we evaluated the absolute frequency of the $^{40}Ca^+$ clock transition to be 411 042 129 776 397 (3) Hz.

At the present, we have evaluated the absolute frequency from the measurements on two separate days. In 2009 M. Chwalla *et al*. reported the $^{40}Ca^+$ clock frequency to be 411 042 129 776 393.2 (1.0) Hz [3]. Although it overlaps with our measurement by the range of the frequency uncertainty, we will need further evaluation about the frequency difference between the two measurements. While we evaluated the BBR shift by the room temperature, it is preferable to measure the temperature at the ion trap electrode surface. Making full use of the advantage of the improved linewith and stability, we are now trying to re-evaluate the systematic shifts. The link uncertainty between the UTC(NICT) and TAI might be also a problem. Taking all effects into account, we will evaluate the absolute transition frequency, aiming at a total uncertainty of low parts in 10^{-15}.

V. CONCLUSION

We have improved our experimental setup using an ion trap equipped with a magnetic shield. The linewidth of the $^2S_{1/2}$ - $^2D_{5/2}$ clock transition was decreased to about 30 Hz. Stabilizing the clock laser frequency to the clock transition, the Allan deviance of 7×10^{-16} at 10^3 s was measured. We preliminarily evaluated the absolute frequency to be 411 042 129 776 397 (3) Hz with respect to the SI definition of the second. We now measure the transition frequency in order to evaluate the absolute frequency with an uncertainty of low parts in 10^{-15} level.

ACKNOWLEDGMENT

The ^{87}Sr optical lattice clock at NICT is developed by Dr. Atsushi Yamaguchi and Dr. Tetsuya Ido. The CSO was developed at the University of West Australia [10] and it is maintained by Dr. Motohiro Kumagai. The authors would like to express their appreciation to these collaborators.

REFERENCES

[1] C.-W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband, "Frequency Comparison of Two High-Accuracy Al^+ Optical Clocks," Phys. Rev. Lett., **104**, 2010, 070802.

[2] K. Matsubara, K. Hayasaka, Y. Li, H. Ito, S. Nagano, M. Kajita, and M. Hosokawa, "Frequency Measurement of Optical Clock Transition of $^{40}Ca^{+}$ Ions with an Uncertainty of 10^{-14} Level," Appl. Phys. Express, **1**, 2008, 067011.

[3] M. Chwalla, J. Benhelm, K. Kim, G. Kirchmair, T. Monz, M. Riebe, P. Schindler, A. S. Villar, W. Hänsel, C. F. Roos, R. Blatt, M. Abgrall, G. Santarelli, G. D. Rovera, and Ph. Laurent, "Absolute Frequency Measurement of the $^{40}Ca^{+}$ $4s^2S_{1/2}$–$3d^2D_{5/2}$ Clock Transition," Phys. Rev. Lett. **102**, 2009, 023002.

[4] K. Matsubara, Y. Li, S. Nagano, H. Ito, M. Kajita, R. Kojima, K. Hayasaka, and M. Hosokawa, "Frequency Stability Measurement of a $^{40}Ca^{+}$ Optical Clock," *Asia-Pacific Radio Science Conference*, Toyama, Japan, 2010, A3a-3.

[5] L.-S. Ma, Z. Y. Bi, A. Bartels, L. Robertsson, M. Zucco, R. S. Windeler, G. Wilpers, C. Oates, L. Hollberg, and S. A. Diddams, "Optical Frequency Synthesis and Comparison with Uncertainty at the 10^{-19} Level," Science, **303**, 2004, pp. 1843–1845.

[6] H. Dehmelt, "Mono-Ion Oscillator as Potential Ultimate Laser Frequency Standard," IEEE Trans. Instrum. Meas., IM–31, 1982, pp. 83–87.

[7] M. Takamoto, F. Hong, R. Higashi, and H. Katori, "An optical lattice clock", Nature, **435,** 2005, pp. 321-324

[8] M. Kumagai, H. Ito, M. Kajita, and M. Hosokawa, "Evaluation of caesium atomic fountain NICT-CsF1," Metrologia, **45**, 2008, pp. 139–148.

[9] C. F. Roos, M. Chwalla, K. Kim, M. Riebe, and R. Blatt, "`Designer atoms for quantum metrology," Nature, **443**, 2006, pp. 316–319.

[10] C. R. Locke, E. N. Ivanov, J. G. Hartnett, P. L. Stanwix, and M. E. Tobar, "Design techniques and noise properties of ultrastable cryogenically cooled sapphire-dielectric resonator oscillators" Rev. Sci. Instrum., **79**, 2008, 051301.

Design Considerations for a ^{87}Sr Optical Clock at VNIIFTRI

Sluysarev S., Kostin A., Baryshev V., Khabarova K., Pal'chikov V
National Research Institute for Physical-Technical and Radiotrchnical Measurements - VNIIFTRI,
Mendeleevo, Moscow Region, 141570, Russia
Email: serslu@mail.ru

***Abstract*— We report on progress made at the VNIIFTRI towards a ^{87}Sr optical clock. A detailed description of the main parts of the apparatus is given together with the preliminary measurements characterizing step by step the cooling and trapping of the strontium atoms. The Sr-MOT is operative and the first observation the cloud of cold atoms in the blue MOT is reported.**

***Index Terms*— optical lattice, frequency standard, cold atoms, magneto-optical trap**

I. Introduction

Inrecent years significant attention has been devoted to optical lattice clocks, in part because the prospects for a fractional frequency uncertainty of such clocks could achieve a level $10^{-17} - 10^{-19}$. On the other hand, the clock transition $^1S_0 - {}^3P_0$ in strontium atoms is being considered as one of the choices for a new definition of the SI second. Precise measurements of this transition at the best laboratories agree with uncertainties in the 10^{-15} range [1] -[4].

In the present paper we discussed the design of the strontium experiment at VNIIFTRI. The $^1S_0 - {}^3P_0$ transition with a hyperfine induced decay rate at λ_0 = 698 nm is used as the clock transition of ^{87}Sr atoms trapped in a 1D vertical optical lattice. The optical lattice formed by the standing wave of linearly polarized light at λ_L = 813 nm was kept on duration the experimental sequence. A preliminary cooling uses a MOT operating at the cycling transition $^1S_0 - {}^3P_1$ at 461 nm. Then the precooled atoms are trapped in the MOT using 698 nm diode laser. In order to probe the clock transition, an external cavity laser diode was stabilized to a high- finesse reference cavity made of ultralow extension (ULE) glass.

The preliminary results of the measurements are given.

II. Experimental Setup

Cooling and trapping of ^{87}Sr are conducted in two stages (Fig. 1). The energy levels related to the experiment are shown in the Fig. 2. All transition wavelengths, needed for cooling and trapping, are easily provided with commercial diode lasers. For example, to generate the 461 nm light we use a standard technique of doubling the frequency of a 922 nm Toptica diode/amplifier laser combination with a resonant cavity containing a BiBO non-linear crystal. The atoms coming out of an oven heated to 450-500°C are slowed down using a 30 cm long Zeeman slower. The trapping and oven chambers are evacuated by a 75 l/s and 10 l/s ion

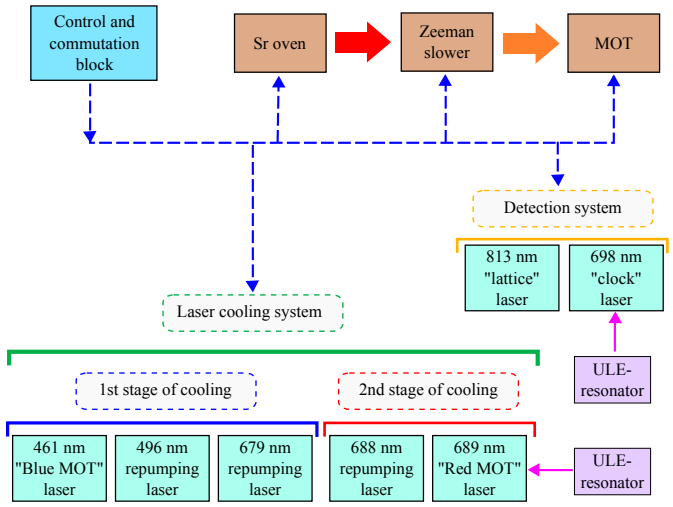


Fig. 1. A schematical drawing of our experiment.

pumps respectively, and vacuum levels $6.5 \cdot 10^{-11}$ mbar (oven chamber) and $1.6 \cdot 10^{-8}$mbar (trapping chamber) are achieved. Anti-Helmholtz MOT coils are made from 4mm OD copper tube, which ensures effective direct watercooling.

A vacuum chamber is mounted on vertically placed optical plate and put inside the system of the coils, intended to compensate the Earth's magnetic field, which have a form of a cube with sides equal 1 m (see Fig. 3). The measurements of the magnetic field generated by 3 pairs of these coils have been made; the currents, necessary for generation of compensatory magnetic field, have been determined for each pair of coils.

To stabilize "Red MOT" 689 nm and "clock" 698 nm lasers, two ULE-resonators are utilized [5]. Each one is placed inside vacuum chamber (vacuum is better than 10^{-7} mbar) and then mounted on 400×400 mm vibration-isolated platform. The characteristics of ULE-resonators are following: temperature excursions (without a laboratory climate control system) are smaller than 10 mK, finesse is 60000, free spectral range is 1.92 GHz.

The first stage MOT is operated on the $^1S_0 - {}^1P_1$ transition (with linewidth 32 MHz) at 461 nm. The number of atoms in the $^1S_0 - {}^1P_1$ MOT were increased by an order by using the 679 nm and 496 nm repumping lasers (see Fig. 4). The second stage cooling operates on the 689 nm $^1S_0 - {}^3P_1$ (linewidth 7.6 kHz) intercombination transition.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

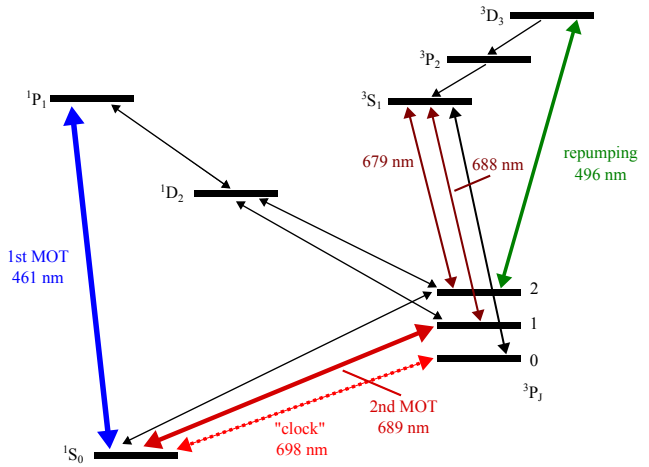


Fig. 2. Energy levels and transitions of Sr, used for spectroscopy.

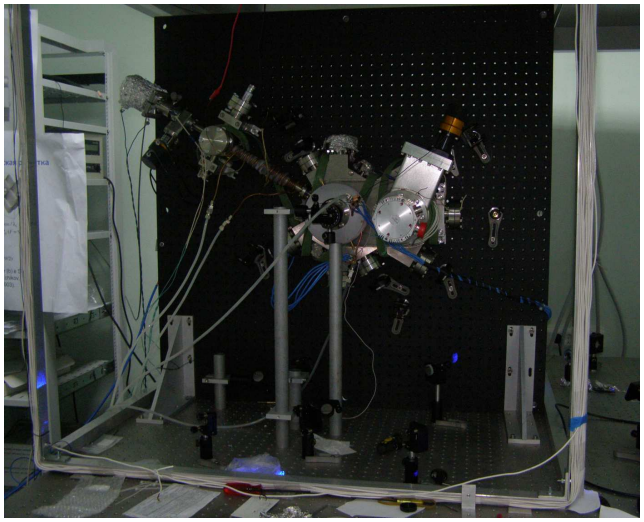
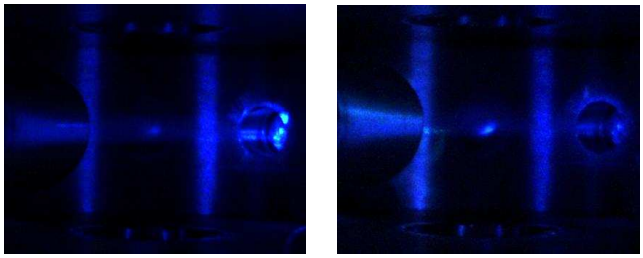


Fig. 3. The working chamber placed on optical table.

III. AOM IN RAMAN-NATH MODE

The second stage laser cooling of Sr atom in a optical lattice on 689 nm $^1S_0 - {}^3P_1$ transition requires an ECDL (external cavity diode laser) whose short-term line-width is comparable with the 7.5 kHz transition natural width. The line-width of the ECDL can be first narrowed by locking the laser to a stable optical cavity via the Pound-Drever-Hall technique. To date, a Pound-Drever-Hall method with AOM-RN, a new sample of acousto-optic modulator designed to operate in a pure Raman-Nath diffraction regime [6], as an optical phase modulator has been used for frequency stabilization of the 689 nm diode laser. To ensure long-term stability, the ECDL should be next locked to the $^1S_0 - {}^3P_1$ resonance via saturated absorption spectroscopy. Again, this can be done by means of FM spectroscopy of saturated absorption resonances in the Sr atomic beam with the same AOM-RN as a phase modulator.

The only seeming disadvantage of AOM-RN is the spatial divergency of its frequency modulated output spectrum. We show experimentally that this divergency, as in the case of saturated absorption spectroscopy in the atomic vapor [6], is not an obstacle to obtain the error signals by means of FM



(a) (b)

Fig. 4. A cloud of cold atoms trapped in Blue MOT (a) without any repumping (b) with 679 nm laser repumping.

sideband heterodyne technique with the use of high-finesse (up to tens of thousands) Fabry-Perot cavities as frequency discriminators.

The experimental set-up is shown in Fig. 5. Developed and designed in VNIIFTRI AOM-RN provides fm spectrum of a single-mode ECDL radiation at $\lambda = 689$ nm. This modulator has a center working frequency of 30 MHz and a modulation bandwidth of 20 MHz. Its acousto-optic interaction length is 2 mm. AOM-RN frequency modulated output spectrum consists of a strong carrier at frequency ω_{ECDL} and two weak sidebands at $\omega_{\mathrm{ECDL}} \pm \Omega$ corresponding to the 1st diffraction orders. Ω is the AOM-RN modulation frequency. Lenses L1 and L2 form a telescope system with AOM-RN in between. The Fabry-Perot (ULE) input radiation represents three collimated and parallel light beams with size of 2 mm in plane where all three beams lie. The ultra-low expansion (ULE) Fabry-Perot cavity with finesse about 60000 and $\mathrm{FSR} = 1.92$ GHz has been used as a frequency discriminator. A part of collimated AOM-RN output consisted, as before, of a carrier and two sidebands was fed via polarization maintaining fiber to the ULE-cavity. The AOM-RN modulation frequency was 19.99 MHz. The total radiation power at the fiber output was $\mathrm{P} = 0.34$ mW. Three dispersion-like error signals can be observed on curve 1 when the ECDL frequency ω_{ECDL} is scanned (see. Fig. 6). Each of them arises when the frequencies of the upper, carrier and lower sidebands match consecutively the resonance frequency of the ULE-cavity.

IV. CONCUSIONS

In summary, we have discussed the prospects of an optical clock operating on the $^1S_0 - {}^3P_0$ transition in ^{87}Sr isotope. As a first step for the realization of proposed optical clock, we have efficiently cooled and trapped ^{87}Sr atoms in a blue magneto-optical trap employing a Zeeman slower. The first stage MOT is operated on the $^1S_0 - {}^3P_1$ transition at 461 nm. We found, that the number of ^{87}Sr atoms were increased by an order by using the 679 nm and 496 nm repumping lasers. The experimental set-up with ULE-cavity and error signals are discussed.

REFERENCES

[1] A. D. Ludlow, T. E. Zelenvinsky, G. K. Campbell et al, Science 319, 1805 (2008)
[2] M. Takamoto,H. Katori, S. I. Marmo, V. D. Ovsiannikov, V. G. Pal'chikov, Phys. Rev. Letters 102, 063002 (2009)

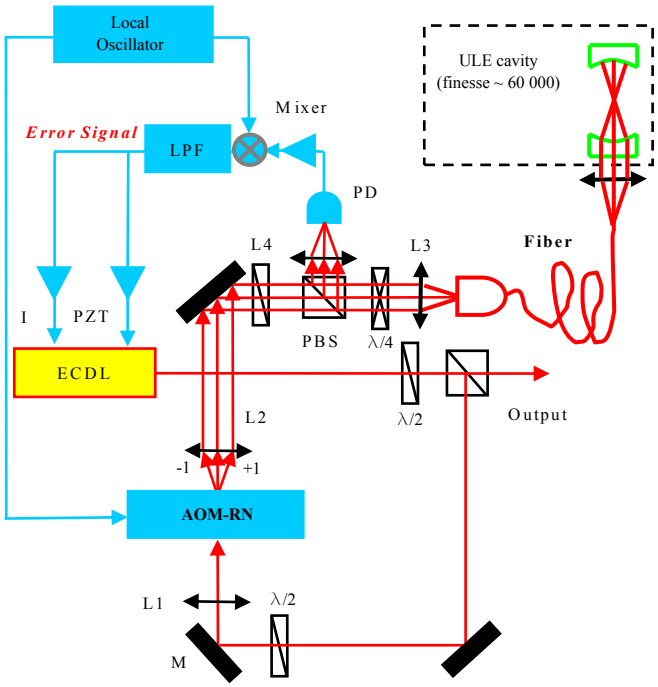


Fig. 5. The experimental set-up with ULE-cavity. ECDL - external cavity diode laser; L - lens; M - mirror; LPF - low-pass filter; PZT - piezo-electric transducer; PBS - polarizing beam splitter; $\lambda/2$ and $\lambda/4$ - halve and quarter waveplates; PD - photo-detector.

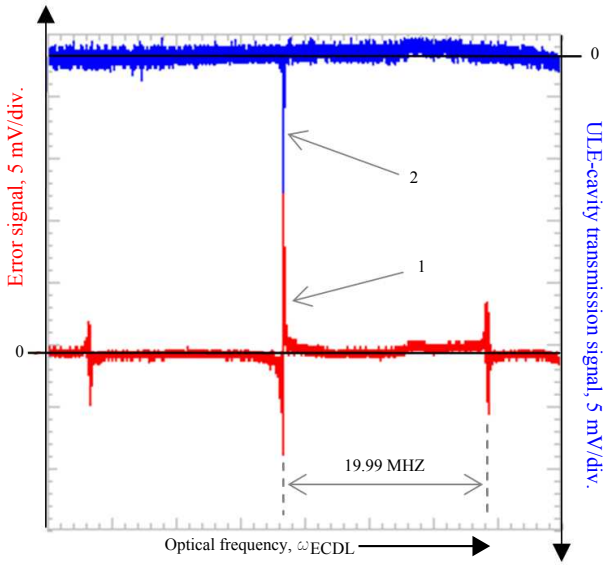


Fig. 6. The error (curve 1) and ULE-cavity transmission (curve 2) signals recorded simultaneously. $\Omega = 19.99$ MHz.

[3] H. Katori, K. Hashiguchi, E. Yu. Il'inova, V. D. Ovsiannikov, Phys. Rev. Letters 103, 153004 (2009)
[4] X. Baillard, M.Fouche, R. Le. Targat, et al, Eur. Phys. J. D48, 11(2008)
[5] J. Alnis, A. Matveev, N. Kolachevsky, Th. Udem, and T. W. Hansch, Phys. Rev. A 77, 053809 (2008)
[6] V. Baryshev, V. Epikhin, Russ. J. Quantum Electronics, vol.40 (5), 431-436 (2010)

Low Vibration Sensitivity Fiber Spools for Laser Stabilization

T. Li

Key Laboratory of Quantum optics, Shanghai Institute of Optics and Fine Mechanics, CAS, Shanghai, PRC

B. Argence, A. Haboucha, H. Jiang, J. L. Dornaux, D. Koné, A. Clairon, P. Lemonde, G. Santarelli

LNE-SYRTE, Observatoire de Paris, CNRS, UPMC, Paris, France

C. Nelson, A. Hati

National Institute of Standards and Technology, Boulder, Colorado, USA

E. Burt

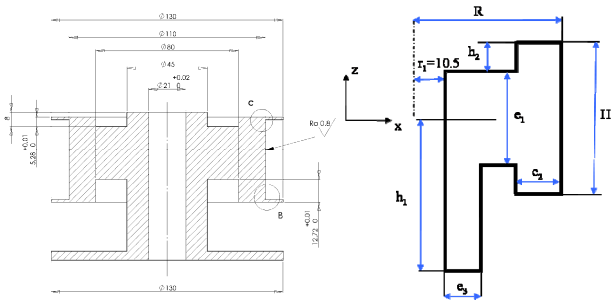
Jet Propulsion Laboratory, Pasadena, California, USA

***Abstract*—Mechanical vibration induced frequency noise is dominated at low Fourier frequencies in a fiber spool stabilized laser. Environmental vibration causes mechanical deformations in the fiber which induce phase fluctuations and then convert into excess frequency noise to the lasers. Therefore, the spool which supports the fiber plays a critical role in this frequency noise conversion. We have studied several different structures of spool. The preliminary results are about $3x10^{-10}/m\ s^{-2}$ for accelerations along the spool axis. In this paper, we describe the development of a spool design which is optimized for low vibration sensitivity along all spatial directions. Both simulations by Finite Element Modeling (FEM) and vibration sensitivity measurements are presented.**

I. INTRODUCTION

Very low noise lasers is a powerful tool which has a variety of important applications, such as optical frequency standards, gravitational wave detection and generation of low phase noise microwave signals. Currently, the lowest noise lasers are realized by stabilizing laser frequency onto an ultra stable high-finesse Fabry-Perot cavity with PDH method [1-3]. However, this approach requires stable and fine alignment of free space optical elements and high vacuum. A radical alternative is to use an optical fiber delay line as a frequency reference to stabilize laser frequency. Recently we demonstrated that the fiber-stabilized laser showed a comparable frequency noise to the one obtained by PDH locking to a high-finesse cavity for Fourier frequency from 40 Hz to 30 kHz and a level below $1Hz^2/Hz$ at 1Hz [4, 5].

The fiber stabilized laser is mainly sensitive to mechanic noise, temperature fluctuation, acoustic noise and air-flow. At low Fourier frequencies, the mechanical vibration is critical. First study on low vibration fiber spool was done by Huang et al [6]. They use an anti-symmetric structure to mount two identical fiber spools to obtain very low vibration sensitivity ($\sim 10^{-11}/ms^{-2}$). We pursuit the effort in this direction and propose to design a spool with similar level of vibration sensitivity or better.



(a) (b)

Fig. 1. (a) Mechanical draw of the spool. (b) Simplified schematic of the spool.

II. DESIGN AND SIMULATION OF LOW VIBRATION SENSITIVITY FIBER SPOOL

Before this design, we studied several configurations, such as simple cylindrical spools and cylindrical spool symmetrically held on a central post, and found that the vibration sensitivity depends strongly on the holding conditions. To overcome this problem, we develop a spool with asymmetric support configuration that would be independent on the holding constraints. The mechanical draw of the spool is shown in Fig. 1 (a). Because the spool has an axial symmetry structure, we restrict our analysis to the half section (Fig. 1 (b)). The spool is made of Titanium because of its stiffness and small mass compared with other metal materials. Some similar works about spool materials have been done and published [7]. The analysis is done by FEM with the hypothesis that fiber delay variation follows the support surface deformation. We test the consistency of the support surface deformation by varying the boundary

conditions. By optimizing the post length, in such a way, the surface variation under acceleration is independent from the boundary conditions applied on the post base.

This geometry shows another interesting feature. If we plot the FEM calculated vibration sensitivity with different h_2 parameter (see Fig.1 (b)) values as shown in Fig. 2, we can observe that the vertical vibration sensitivity can be zeroed with an optimized value of $h_2 \approx 5.2$mm. Other geometrical parameters (see Fig. 1 (b)) also influence the vibration sensitivity. For example, increasing the central width e_1 moves the null sensitivity point close to $h_2=0$ while increasing the radius R displaces this point towards opposite direction. The increasing of e_1 value reduces the vibration sensitivity coupling to the holding post length and the boundary conditions. The final geometry of the spool is optimized to hold more than 2km of fiber. In addition, a modal analysis is also performed and shows the first resonance of about 1.5 kHz.

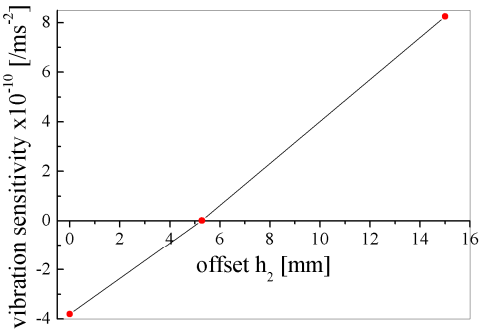


Fig. 2. Plot of FEM calculated vibration sensitivity with different h_2 parameter values

III. VIBRATION SENSITIVITY MEASUREMENT SETUP

The measurement is done in three steps. We first lock the laser source to the fiber interferometer and measure its frequency noise. Then, we shake the fiber spool vertically and measure the transfer function of frequency noise to vertical acceleration. And finally, the vibration sensitivity of the spool is calculated by the transfer function. The schematic of measurement is shown in Fig. 3. The interferometer part is enclosed by dash line.

A fiber laser at 1542.14 nm is used as the laser source, the laser beam goes to the Michelson interferometer passing through an acousto-optical modulator (AOM1). Another acousto-optical modulator (AOM2) is placed into the long arm of interferometer after a 300 m fiber spool and driven by a 70 MHz RF signal. Therefore, the heterodyne detection signal is 140 MHz. We use a home-made low phase noise tunable synthesizer to provide demodulation signal and a low-pass filter converts the demodulation signal into laser frequency correction signal. The correction signal simultaneously acts on a piezo-electric transducer (PZT) stretcher which provides a slow correction on large range and a voltage controlled oscillator (VCO) which drives AOM1 providing fast correction. The fiber spool stabilized laser is combined with a signal from PDH cavity stabilized laser to provide a beat note signal for frequency noise analysis. The fiber spool is housed in an aluminum cylindrical enclosure and placed on an active vibration isolation platform which can be used as a low level shaker by external modulation. In the same time, the other part of interferometer is placed on a passive vibration isolation platform and connected to the fiber spool by two output fibers. The transfer function of frequency noise to vertical acceleration $T_r(f)$ is measured by a vector signal analyzer HP89410. This instrument generates a chirp modulation signal injected in the active vibration isolation platform and then calculates the transfer function between acceleration noise $S_a(f)$ and the measured frequency noise $S_v(f)$ as follow:

$$T_r(f) = (S_v(f) / S_a(f))^{1/2} \quad (1)$$

The acceleration is measured by an accelerometer 356B18 (PIEZOTRONICS INC.) placed on the top of the aluminum enclosure. The vibration sensitivity of the spool $\Gamma(f)$ can be presented as:

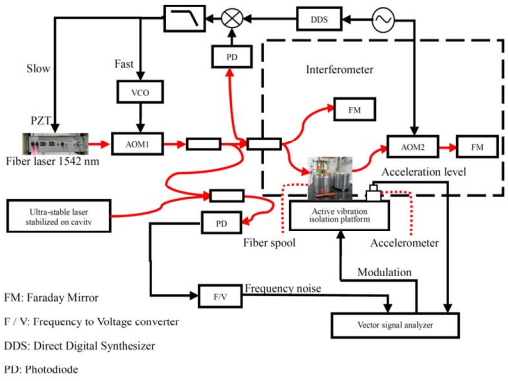


Fig. 3. Schematic of vibration sensitivity measurement, the optical signals are drawn in red while the electronic signals drawn in black

$$\Gamma(f) = \Delta L / (L\Delta V) = \Delta\tau / (\tau\Delta V) = T_r(f) / \nu_0 \quad (2)$$

Where ΔL, L, $\Delta\tau$, τ, ΔV, ν_0 represents length variation, total length, delay variation, total delay of the fiber spool and acceleration variation, optical frequency, respectively. Additionally, it should be noticed that the acceleration level is quite low (1mg RMS maximum level). At frequency higher than 60 Hz the acceleration rolls off of about 20 dB up to 110 Hz because of the response of the active vibration isolation platform.

IV. THE RESULTS AND DISCUSSIONS

In our experiment, the reference fiber is standard single mode fiber (SMF-28). Two output fibers of about 0.4 m are used as connections between the fiber spool and other optical components in interferometer. These output fibers are bare fiber (8/125/250μm) and can be looked as "free" fiber because

they are only fixed by few points. We notice that the vibration sensitivity of the fiber spool has dependency on the way to fix these fibers. As a consequence, we measure the vibration sensitivity of the fiber spool with different tensions on the "free" fiber. The results are shown in Fig. 4. The vibration sensitivity of new-design spool wound with 300 m fiber is around $5\times10^{-11}/ms^{-2}$. However, it is strongly influenced by the constraints on the "free" fiber. The Young's modulus of SMF-28 fiber is about two orders of magnitude lower than that of spool materials. Therefore, the vibration sensitivity of the "free" fiber should be same order higher than that of the fiber wound on the spool. Some studies about fiber vibration sensitivity have been doing in NIST [8]. Moreover, different constraints on the "free" fiber will change its vibration transmission and result in vibration sensitivity variation.

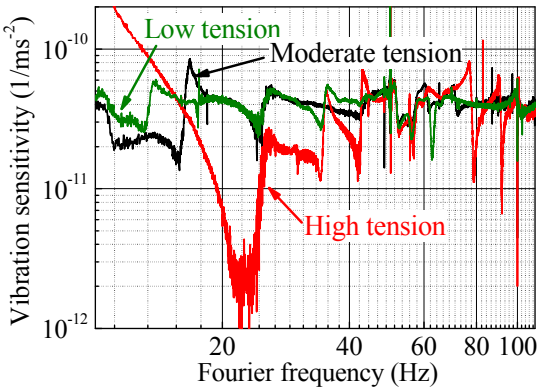


Fig. 4. Vibration sensitivity of fiber spool with different tensions applied on the "free" fiber

The impact of the "free" fiber on the overall vibration sensitivity is still being studied. Nevertheless, several ideas can be carried out to reduce its contribution, such as increasing the fiber length wound on the spool while reducing the "free" fiber length, removing the optical connectors instead of optical splicing. In addition, the studies of vibration sensitivity contributions from other optical components (AOM, FM, etc.) will be performed in the future. In this way, we want to demonstrate a compact ultra-low noise laser without vibration isolation platform that could be used for a variety of precision measurement applications.

REFERENCES

[1] A. D. Ludlow, X. Huang, M. Notcutt, T. Zanon-Willette, S. M. Foreman, M. M. Boyd, S. Blatt, & J. Ye, Compact, thermal-noise-limited optical cavity for diode laser stabilization at 1e-15. Opt. Lett., (2007) 32, 641–643.

[2] J. Millo, D. V. Magalhães, C. Mandache, Y. Le Coq, E. M. L. English, P. G. Westergaard, J. Lodewyck, S. Bize, P. Lemonde, and G. Santarelli. Ultrastable lasers based on vibration insensitive cavities. Phys. Rev. A (Atomic, Molecular, and Optical Physics), (2009) 79, 053829.

[3] S. A. Webster, M. Oxborrow, S. Pugla, J. Millo, and P. Gill, Thermal-noise-limited optical cavity. Phys. Rev. A (Atomic, Molecular, and Optical Physics), (2008) 77, 033847.

[4] F. Kefelian, H. Jiang, P. Lemonde & G. Santarelli, Ultralow-frequencynoise stabilization of a laser by locking to an optical fiber-delay line. Opt. Lett., (2009a). 34, 914–916.

[5] H. Jiang, F. Kefelian, P. Lemonde, A. Clairon & G. Santarelli, An agile laser with ultra-low frequency noise and high sweep linearity. Opt. Express, (2010)18, 3284–3297.

[6] S. Huang, M. Tu, S. Yao & L. Maleki, A 'turn-key' optoelectronic oscillator with low acceleration sensitivity. FCS'2000 , CL00–0626.

[7] J. Taylor, C. Nelson, A. Hati, N. Ashby & D. Howe. Vibration-induced pm noise measurements of a rigid optical fiber spool. Frequency Control Symposium, 2008 IEEE International , 807–810.

[8] A. Hati, C. Nelson & D. Howe, Vibration sensitivity of optical components: a survey. Joint Conference of the IEEE International Frequency Control Symposium and European Frequency and Time Forum, 2011.

Coherent detection of an active mode-locked terahertz quantum cascade laser

G. Santarelli
LNE-SYRTE,
CNRS, UPMC, Observatoire de Paris, Paris, France
S. Barbieri, M. Ravaro,P. Gellie, C. Manquest, C. Sirtori
Laboratoire Matériaux et Phénomènes Quantiques,
Université Paris 7, CNRS
Paris, France

S. P. Khanna, E. H. Linfield, and A. G. Davies
School of Electronic and Electrical Engineering,
University of Leeds,
Leeds, United Kingdom

Abstract— We demonstrate coherent detection of the emission from a terahertz frequency quantum cascade laser which is actively mode-locked by modulating its bias-current with a microwave synthesizer. Coherent sampling of the terahertz pulse train is achieved by phase-locking the repetition rate and the carrier frequency of the terahertz quantum cascade laser to the repetition rate of the of a commercial erbium-doped mode-locked femtosecond fibre laser. We show that the terahertz pulses from the quantum cascade laser are transform limited, and reconstruct the full emitted waveform.

Introduction

Quantum cascade lasers (QCLs) are unipolar semiconductor lasers based on electronic intersubband transitions in the conduction band of multi-layered heterostructures [1]. In contrast to interband diode lasers, where the lifetime of the electrons populating the upper laser state is determined by electron-hole recombination, the electron lifetime in QCLs is dominated by non-radiative phenomena, and in particular optical-phonon scattering. This leads to fast relaxation times, on the picosecond (ps) timescale, which have an important impact on the dynamical properties of this family of devices [2]. In particular, modulation bandwidths up to several tens of GHz have been demonstrated, which were recently used to show that the frequency separation, or roundtrip frequency, between the longitudinal modes of multimode terahertz (THz) frequency QCLs [3-7] can be stabilised using an external RF-synthesiser [8-11]. Under these conditions the laser is expected to operate in a regime of active mode-locking, producing a train of short pulses. The lack of suitable detection techniques in the THz range has prevented the measurement of these pulses. Difficulty in demonstrating QCL mode-locking has previously been encountered at mid-infrared (mid-IR) frequencies, and arises in part from the inefficiency of second-order autocorrelation techniques in the mid-IR (and THz) ranges, hindering the measurement of the emitted pulse train. The first reports of passive and active mode-locking of mid-IR QCLs were later identified as coherent dynamic instabilities resulting from the fast gain recovery time (~1 ps) preventing the formation of stable mode-locked pulses. Recently, however, mode-locking of a QCL operating at 6.3 μm has been demonstrated using 2nd order auto-correlation with a two-photon quantum well detector. Compared to mid-IR QCLs, THz QCLs based on bound-to-continuum active regions have generally longer non-radiative relaxation times of the upper laser state, of ~5-10 ps, as a result of the laser transition energy lying below the optical phonon energy.

We show that THz QCLs can be operated in a regime of active mode-locking. The measurement proof of mode-locking operation is obtained by sampling the electric field amplitude emitted by the QCL using a mode-locked fs-laser. In our work, we use off the shelf mode-locked erbium-doped fs-laser emitting at 1.55 μm (193 THz), for the asynchronous sampling of a mode-locked QCL emitting at 2.5 THz [7]. Mode-locking of the QCL is obtained by direct modulation of its bias current close to the laser roundtrip frequency[11]. Asynchronous sampling is made possible by phase-locking the QCL source to the fs-laser,[12] i.e. the pulse repetition rate of the QCL at ~13.3 GHz and its carrier frequency at 2.5 THz are simultaneously phase-locked to a harmonic of the repetition rate of the fs-laser-comb at 96.5 MHz. Furthers details in Refs [12,13]

REFERENCES

[1] R. Paiella (ed.), Intersubband transitions in quantum structures (Mc-Graw Hill Nanoscience and Technology, 2006.

[2] R. Paiella, et al. High-frequency modulation without the relaxation oscillation resonance in quantum cascade lasers. Appl. Phys. Lett., Vol.79, pp. 2526-2528, 2001.

[3] R. Köhler, et al. "Terahertz semiconductor-heterostructure laser," Nature, Vol. 417, pp.156-159, 2002.

[4] G. Scalari, et al. THz and sub-THz quantum cascade lasers. Laser Photon. Rev., Vol. 3, pp.45–66, 2009.

[5] B. S. Williams. "Terahertz quantum cascade lasers," Nature Photon., Vol. 1, pp. 517-525, 2007.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

[6] S. Kumar, Q. Hu, and J. Reno. "186 K operation of terahertz quantum-cascade lasers based on a diagonal design," Appl. Phys. Lett., Vol. 94, pp. 131105-131107, 2009.

[7] S. Barbieri, et al. "2.9 THz quantum cascade lasers operating up to 70 K in continuous wave," Appl. Phys. Lett., Vol. 85, pp. 1674-1676, 2004.

[8] S. Barbieri, et al. "13GHz direct modulation of terahertz quantum cascade lasers," Appl. Phys. Lett., vol 91, pp. 143510-143512, 2007.

[9] W. Maineult, et al."Microwave modulation of THz quantum cascade lasers: a transmission−line approach," Appl. Phys. Lett., Vol. 96, pp. 021108, 2010.

[10] M. C. Wanke, et al. "Monolithically integrated solid-state terahertz transceivers," Nature Photon., Vol. 4, pp. 565-569, 2010.

[11] P. Gellie, et al "Injection-locking of terahertz quantum cascade lasers up to 35GHz using RF amplitude modulation Linewidth and tuning characteristics of terahertz quantum cascade lasers," Opt. Express, Vol 18, pp. 20799-20816, 2010.

[12] S. Barbieri, et al. "Phase-locking of a 2.7-THz quantum cascade laser to a mode-locked erbium-doped fibre laser," Nature Photon., Vol 4, pp. 636-640, 2010.

[13] S. Barbieri, et al. "Coherent sampling of active mode-locked terahertz quantum cascade lasers and frequency synthesis," Nature Photon., Vol 5, pp. 306-313, 2011.

GHz AlN Lateral Overmoded Bulk Acoustic Wave Resonators with a $f \cdot Q$ of 1.17×10^{13}

Songbin Gong, Nai-Kuei Kuo, and Gianluca Piazza
Penn Micro and Nano Systems Lab
University of Pennsylvania
Philadelphia, USA 19104
Email: {songbin, kuo1, piazza}@seas.upenn.edu

***Abstract*— This work presents a new class of devices, AlN-based lateral overmoded bulk acoustic wave resonators (LOBAR), to investigate the Q limit imposed by the intrinsic loss of thin film AlN. The LOBAR devices utilize a rectangular AlN thin film plate with a small percentage (0.57%) inter-digitated transducer (IDT) fingers coverage to minimize the piezoelectric-metal interface loss. This design experimentally pushes the $f \cdot Q$ product value (1.17×10^{13}) for thin-film AlN to magnitude level close to the theoretical limit predicted by the Akhizer effect (AKE) for AlN ($2.5\text{-}5 \times 10^{13}$). To study the performance of the LOBAR design, the IDT to rectangular plate coverage ratio was varied. The results have shown that the Q of the LOBAR device scales nonlinearly and inversely with coverage ratio. Measurement of the highest Q LOBAR device for temperatures varying between 300 and 400 K have indicated that the AlN LOBAR has a temperature coefficient of frequency (TCF) of -16.25 ppm/K, and phonon-phonon dissipation (AKE) is not yet the dominant loss mechanism..**

I. Introduction

Driven by the increasing needs for high performance, multi-functional wireless services, several researcher groups have been working on the implementation of intelligent systems that realize more efficient ways of utilizing the increasingly crowded RF spectrum. One of the most popular approaches involves sensing the ambient radio spectrum and seeking for opportunistic use of idle RF channels [1]. Making such systems feasible for handheld mobile applications, requires the implementation of low cost and power spectrum analyzers. A very promising way of synthesizing such a system with low complexity is to detect the RF power density across the spectrum using passive narrowband channelizes and power detectors. Such approach has strong potential to enable low cost deployment of cognitive radios [2]. The same channelizer can also be used to assist feature-base spectrum sensing to significantly reduce the device count and power requirements. However, the implementation of such narrowband channelizers, which requirehigh Q ($>$ 10,000) and low motional impedance resonators (50Ω), remains a grand challenge. Micro-machined electrostatic and piezoelectric resonators have already demonstrated their capabilities to potentially meet these requirements at high and very high frequencies(3-300 MHz) [3][4]. However, at higher RF frequencies (∼3 GHz), such task becomes significantly more difficult due to $f \cdot Q$ limit imposed by phonon-phonon dissipation [5]. In the past, numerous research efforts have been focused on bringing the experimentally

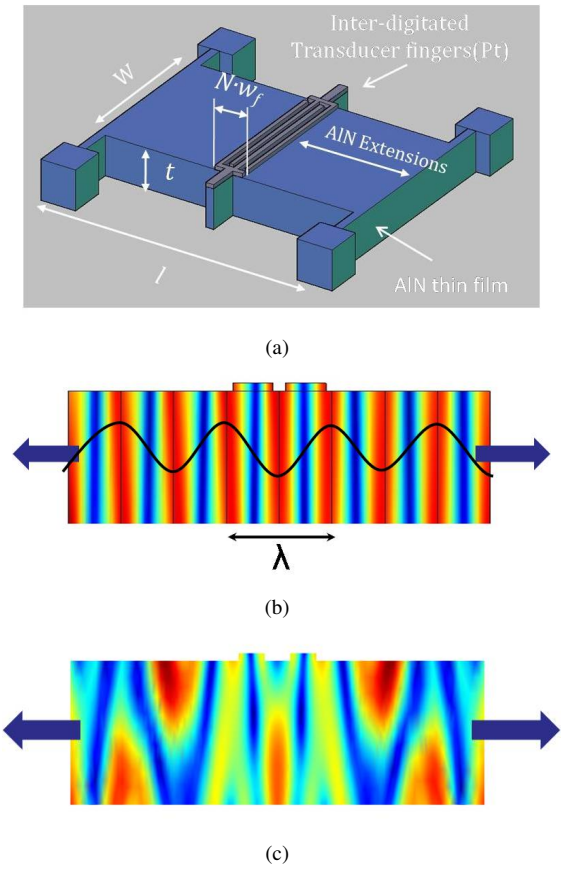


Fig. 1. (a) Mockup view of the AlN-LOBAR resonator (b) COMSOL simulation showing the lateral vibration generated in the AlN-LOBAR devices (c) COMSOL simulation showing the S1 lamb-mode (3.3 GHz) in the same device.

demonstrated $f \cdot Q$ product of different materials close to the their theoretically predicted values, and simultaneously enhancing the understanding of various loss mechanism in play [6]. Among these low loss materials, it has been show that Si can reach $f \cdot Q$ values (2-3 $\times$ 10^{13}) [7] [8] at GHz frequencies that are close to the theoretical limit(3.9×10^{13}) [5]. Instead, AlN-based resonators have an experimental $f \cdot Q$ that is about one order of magnitude lower than its theoretical limit[3][9][10][11].

This work reports on a new class of devices, the AlN

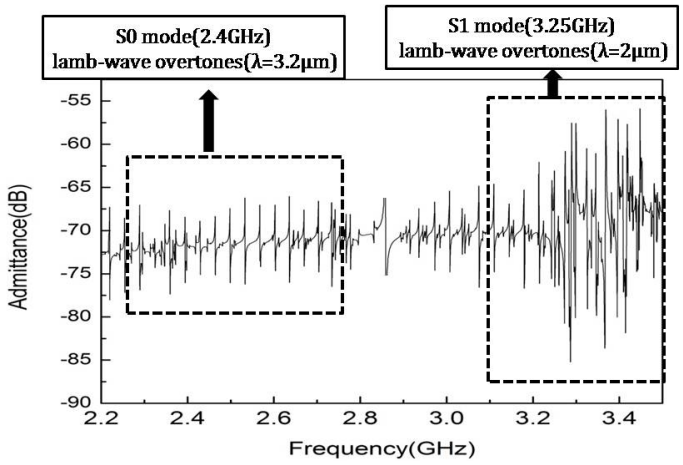


Fig. 2. COMSOL simulated admittance performance of one port AlN-LOBAR device having the dimensions shown in Table 1. A Q of 10,000 was assumed for AlN and a Q of 2,000 was assumed for Pt electrodes.

LOBAR, which permits to explore the ultimate $f \cdot Q$ in thin AlN piezoelectric films . Owing to the partial electrode coverage of the AlN LOBAR design, the device minimizes the piezoelectric-metal interface induced loss. Therefore, a high $f \cdot Q$ can be achieved. Effectively, the LOBAR device can be classified as a general instrument that can be readily adopted for experimentally studying the theoretical $f \cdot Q$ limits of many low loss acoustic materials (SiC [12], Si [8], Diamond [13], $LiNbO_3$ [14], and sapphire). In this paper we will demonstrate and study the LOBAR performance in AlN thin films.

II. DEVICE DESIGN AND MODELING

A. Device Operation

As shown in Fig. 1, the AlN LOBAR device, similarly to high overtone bulk acoustic wave resonators [15], utilizes a small volume piezoelectric transduction to induce vibrations in a larger resonant cavity. The prototype device has IDT placed in the center of the AlN thin film plate covering a small portion ($< 2\%$) of overall resonator. In operation, instead of launching thickness mode vibrations vertically into the substrate like the HBAR [16], the AlN LOBAR excites laterally S0 and S1 overtone vibrations into the AlN extensions as shown in Fig. 1(b) and 1(c). The center frequency of the S0 mode is defined by the lateral dimension of the IDT fingers. The S1 mode of vibration is also excited by the same IDT, but its center frequency is set primarily by the thickness of the AlN thin film plate. The LOBAR design incorporates multiple advantages. First, the piezoelectric-metal interface loss is greatly minimized by limiting the metal coverage on AlN. This is in line with what has been shown experimentally for AlN contour mode resonators [17]. Second, the well defined acoustic boundaries of the LOBAR in the lateral and vertical dimensions should yield lower energy loss into the supporting substrate (differently from the HBAR whose lateral boundaries are not well defined). Third, the device is easily fabricated with well-establish surface micromachining processes [18][19].

TABLE I

THE DEVICE DIMENSION FOR THE COMSOL SIMULATION.

$W(\mu m)$	$l(\mu m)$	$t(\mu m)$	$w_f(\mu m)$	t_f(nm)[1]	N[1]	CR[1]
48	560	1	1.6	50	2	0.57%

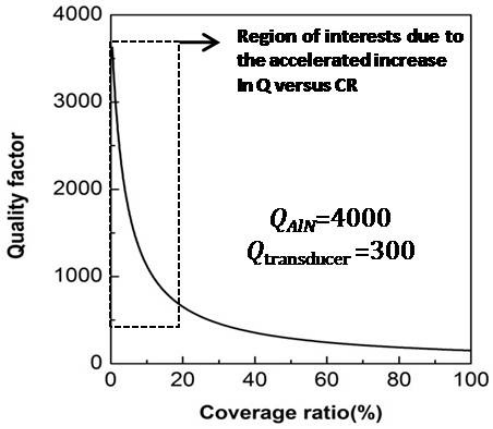


Fig. 3. Theoretical analysis of the quality factor versus coverage ratio. Arbitrary assumptions (close to what seen experimentally in GHz devices) on the transducer and AlN Q were made.

B. COMSOL Analysis

A one port device with the dimensions shown in Table I was simulated in COMSOL to predict its performance across the spectrum. The simulated results, shown in Fig. 2, indicate that two groups of overtones exist. The first group of overtones centered around 2.4 GHz are S0 (longitudinal) vibrations, whose center frequency is set by the lithographically defined IDT finger width (1.6 μm). The second group of overtones appearing between 3.0 and 3.5 GHz are the S1 mode lamb wave vibrations, of which center frequency (3.25 GHz) is determined by the thickness of the AlN thin film (1 μm in this case). The mode shape of one of the overtones (3.4 GHz) is shown in Fig. 1(c) to clearly show that most of the vibrations occur across the thickness of the film, although they propagate laterally.

C. Damping Model

In order to understand how the coverage ratio of the electrodes (IDT to total rectangular plate length ratio) affects the overall Q of the device, a simple linear damping model is used. Damping coefficients, ζ, are assigned to the transducer and extension section (uncovered portion of the AlN plate) of the LOBAR device.

$$\zeta_{unmetalized-AlN} = \frac{1}{2 \cdot Q_{unmetalized-AlN}} \quad (1)$$

$$\zeta_{Transducer} = \frac{1}{2 \cdot Q_{Transducer}} \quad (2)$$

As shown in Eq. 3, the overall damping coefficient of the device is related to the damping coefficients of the transducer

[1]where N is the number of fingers, and t_f is the thickness of the IDT electrodes, and CR is the coverage ratio.

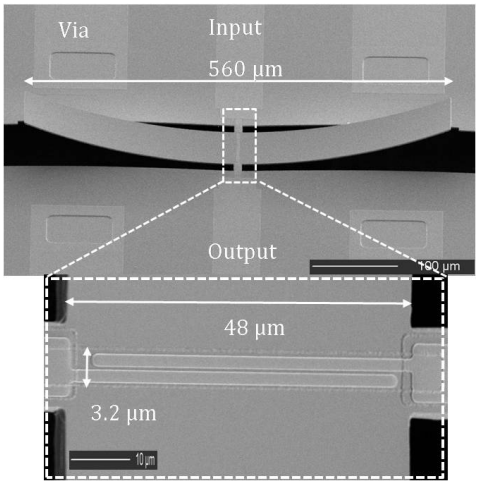


Fig. 4. The SEM of the fabricated AlN resonator with a coverage ratio of 0.67%. The device is made out of a 48×560 μm thin film AlN plate with 2 IDT fingers, each having a width of 1.6 μm. The anchors are placed in the center of the two sides of plate to provide signal routing and mechanical support.

and AlN extension via the coverage ratio.

$$\zeta_{Total} = \zeta_{unmetalized-AlN} \cdot (1 - CR) + \zeta_{Transducer} \cdot CR \quad (3)$$

$$Q_{Total} = \frac{1}{2\zeta_{Total}} \quad (4)$$

Fig. 3 shows that the quality factor of the overall device increases as the coverage ratio decreases for a given Q value of the transducer and extension sections. Due to the nonlinear relationship between Q and coverage ratio, an accelerated increase in Q is observed as the coverage ratio is reduced to below 10%. In order to reveal the Q limit imposed by the AlN material at high frequencies, a coverage ratio study should be performed for low percentages ($< 2\%$) so as to minimize the piezoelectric-interface loss.

III. EXPERIMENTAL RESULTS

Multiple devices are simultaneously fabricated using a fabrication process similar to the one described in [19]. One of the fabricated devices, with the same dimension as the one in the COMSOL model, is shown in the Fig. 4. The zoomed-in view shows excellent definition of the IDT fingers. The released AlN thin film exhibits upward curvature due to stress gradient accumulated in the fabrication process.

A. Transmission Response

The fabricated devices were measured at room temperature and atmospheric pressure with a probe station and network analyzer. On-chip TRL calibrations were performed for all these measurements. The measured transmission response, plotted in Fig. 5, shows multiple overtone resonances across the spectrum from 1.5 to 3.5 GHz with an even spacing of 16MHz. As predicted by the COMSOL simulations, the lateral mode (S0) overtone resonances center around 2.4 GHz. These S0 mode overtones ranging from 1.5 to 2.6 GHz attain average

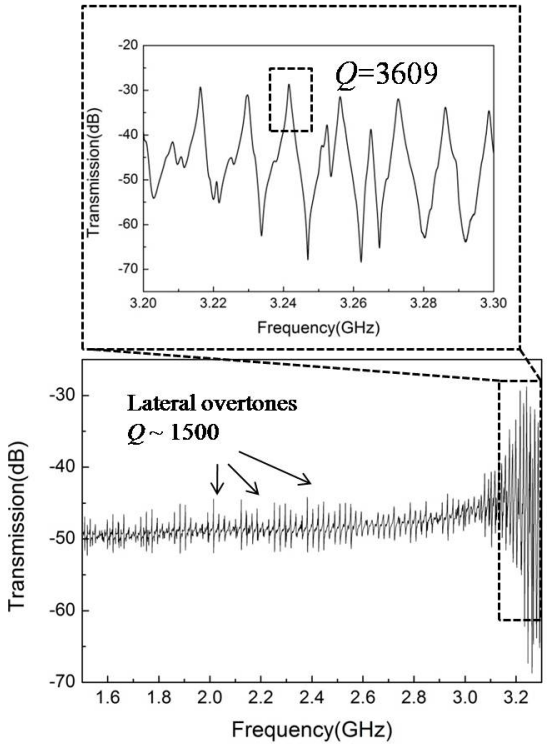


Fig. 5. Measured transmission response of the two-port AlN LOBAR device.

TABLE II
DIMENSIONS OF THE DEVICES USED IN THE COVERAGE STUDY AND THE MEASURED Q AND EXTRACTED R_m OF THESE DEVICES.

$Device$[1]	$W(\mu m)$	$l(\mu m)$	N	CR	R_m	Q
A	48	560	2	0.57%	3062	3609
B	48	480	2	0.67%	2912	3294
C	48	400	2	0.8%	3100	2935
D	48	320	2	1%	3021	2624
E	48	480	4	1.14%	3108	2284
F	48	320	4	2%	3087	1600

Qs of about 1500. S1 mode vibrations are also excited around 3.3 GHz with multiple overtone resonances. In comparison with the S0 mode, the S1 mode resonances have a smaller insertion loss due to the higher value of the electromechanical coupling coefficient (due to d33 $>$ d31). The highest Q among these overtones was recorded experimentally at 3.244 GHz. A value of 3609 was measured, resulting in a $f \cdot Q$ product of 1.17×10^{13}. This is the highest $f \cdot Q$ ever demonstrated for AlN-based micro-resonators (Fig. 6). This result is also very close to the theoretical limit for AlN [5]. This demonstration has effectively verified that AlN is indeed a high Q material at GHz frequencies, has great potentials in achieving higher Qs, and expand the use of AlN-based micro-resonators into new applications.

B. Coverage Ratio Study

[1]The AlN thickness in these devices is kept to 1 μm.

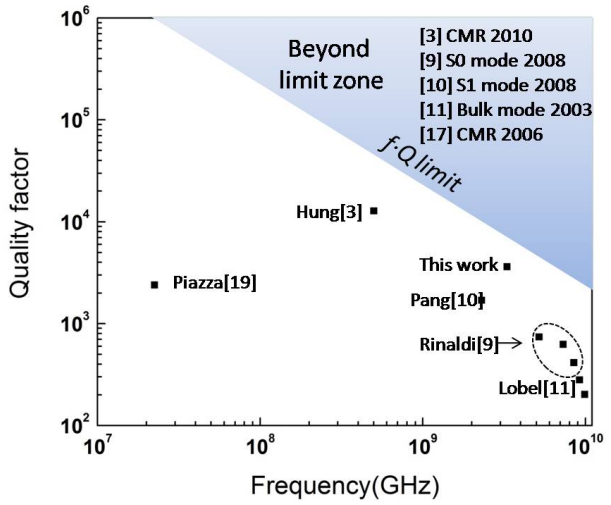


Fig. 6. Comparison of $f \cdot Q$ achieved by the state-of-art AlN-based micro-resonators.

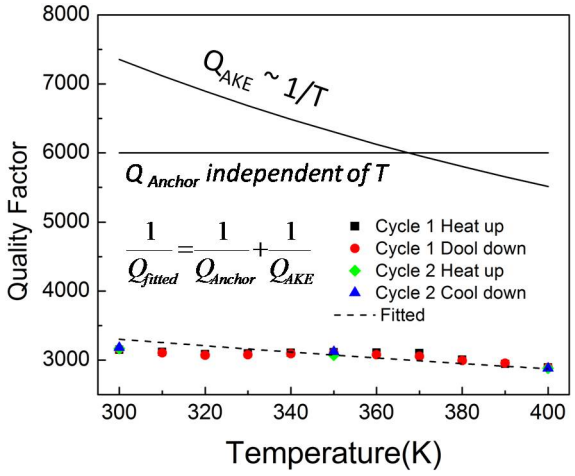


Fig. 7. The Measured Q of the LOBAR device versus temperature. Fitting of the measurements is done assuming phonon scattering and anchor loss are the two major damping mechanisms.

To experimentally investigate the effect of piezoelectric-metal interface induced loss on the quality factor of the AlN LOBAR devices, a coverage ratio (CR) study consisting of 6 devices with different dimensions and CR was performed. The device parameters are listed in Table II. The lateral dimensions of device A, B, C, and D are fixed to be 48 μm while their longitudinal dimension varies from 560 to 320 μm to attain coverage ratios from 0.57 to 1 %. Device E and F, despite having the same dimensions as device A and B, further increase the coverage ratio to 1.14 and 2 % by increasing the number of IDT fingers to 4. The measured Q values show an inverse scaling versus coverage ratio as predicted by the damping model. The extracted motional resistance R_m remains roughly constant as the coverage ratio varies. Therefore, although reducing the CR of the electrodes results in high Q, a trade-off between Q and electromechanical coupling, k_t^2 has to be made.

C. Temperature Study

The fabricated LOBAR devices were tested from 300 to 400 K under atmospheric pressure using the Lakeshore probe station. The devices were cycled between 300 and 400 K three times before each measurements. The device shown in Fig. 4 exhibits a TCF of -16.25 ppm/K as opposed to the -25 ppm/K commonly demonstrated by AlN contour mode resonator [19]. This is because the composition of AlN LOBAR is mostly un-metalized AlN and the TCF of these devices should be close to the theoretical value derived for sputtered AlN thin films [19]. The measured Q of the device shows little change versus temperature. As previously reported, phonon scattering due to Akhieser effect imposes a lower limit on Q than thermoelastic dissipation [5] in high frequency, extensional mode resonators,. The theoretical limit of Q thus can be formulated as:

$$Q_{AKE} = \frac{\rho c^2}{2\pi\gamma^2 c_v T}\frac{1}{f} \quad (5)$$

where γ is the Gruneisen's parameter, c is the acoustic velocity, c_v is the heat capacity per unit volume, T is the temperature, and f is the frequency of operation. As shown in Fig. 7, the Q of the LOBAR shows little dependence on temperature between 300 and 400 K. This indicates that other damping mechanisms independent of temperature (differently from the Akhizer's effect) are still contributing to setting the Q of the overall device. It is believed that anchor loss is playing a key role. Future studies on a larger set of devices will help in clarifying the origin of these damping mechanisms and engineer ways to is required in order lift the LOBAR Q limit close to Q_{AKE}.

IV. CONCLUSION

To summarize, a new class of piezoelectric micro-resonators, the LOBAR, has been demonstrated using sputtered AlN thin films. The fabricated devices exhibit multiple S0 and S1 mode overtone resonances across the GHz frequency spectrum. The highest Q of 3607 is observed for one of the S1 mode overtones, resulting in a $f \cdot Q$ product of 1.17×10^{13}. This is the highest $f \cdot Q$ product ever demonstrated for state-of-art AlN-based micro-resonators. These devices have moved one step closer to experimentally reach the theoretical $f \cdot Q$ limit imposed by the AlN material. A study on the electrode coverage ratio also validated the damping model used for understanding the working principle behind LOBAR type devices.

ACKNOWLEDGMENT

The authors would like thank DARPA for funding support through the chip scale spectrum analyzer project (CSSA) and Nipun Sinha for help with the device fabrication.

REFERENCES

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201 – 220, feb. 2005.

[2] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 116 –130, quarter 2009.

[3] L.-W. Hung and C.-C. Nguyen, "Capacitive-piezoelectric AlN resonators with Q $>$ 12,000," in *Micro Electro Mechanical Systems (MEMS), 2011 IEEE 24th International Conference on*, jan. 2011, pp. 173 –176.

[4] J. Clark, W.-T. Hsu, M. Abdelmoneum, and C.-C. Nguyen, "High-Q UHF micromechanical radial-contour mode disk resonators," *Microelectromechanical Systems, Journal of*, vol. 14, no. 6, pp. 1298 – 1310, dec. 2005.

[5] S. Chandorkar, M. Agarwal, R. Melamud, R. Candler, K. Goodson, and T. Kenny, "Limits of quality factor in bulk-mode micromechanical resonators," in *Micro Electro Mechanical Systems, 2008. MEMS 2008. IEEE 21st International Conference on*, jan. 2008, pp. 74 –77.

[6] M. Weinberg, R. Candler, S. Chandorkar, J. Varsanik, T. Kenny, and A. Duwel, "Energy loss in MEMS resonators and the impact on inertial and RF devices," in *Solid-State Sensors, Actuators and Microsystems Conference, 2009. TRANSDUCERS 2009. International*, june 2009, pp. 688 –695.

[7] E. Hwang and S. Bhave, "PN-diode transduced 3.7-GHz silicon resonator," in *Micro Electro Mechanical Systems (MEMS), 2010 IEEE 23rd International Conference on*, jan. 2010, pp. 208 –211.

[8] M. Palaniapan and L. Khine, "Micromechanical resonator with ultra-high quality factor," *Electronics Letters*, vol. 43, no. 20, pp. 1090 –1092, 27 2007.

[9] M. Rinaldi, C. Zuniga, and G. Piazza, "5-10 GHz AlN contour-mode nanoelectromechanical resonators," in *Micro Electro Mechanical Systems, 2009. MEMS 2009. IEEE 22nd International Conference on*, jan. 2009, pp. 916 –919.

[10] W. Pang, R. Ruby, R. Parker, P. Fisher, M. Unkrich, and J. Larson, "A temperature-stable film bulk acoustic wave oscillator," *Electron Device Letters, IEEE*, vol. 29, no. 4, pp. 315 –318, april 2008.

[11] H. P. Loebl, M. Klee, C. Metzmacher, W. Brand, R. Milsom, and P. Lok, "Piezoelectric thin AlN films for bulk acoustic wave (baw) resonators," *Materials Chemistry and Physics*, vol. 79, no. 2-3, pp. 143 – 146, 2003.

[12] S. Gong, N. Kuo, and G. Piazza, "A 1.75 GHz SiC lateral overtone bulk acoustic-wave resonator," in *Solid-State Sensors, Actuators and Microsystems Conference, 2009. TRANSDUCERS 2011. International*, june 2011.

[13] J. Wang, J. Butler, T. Feygelson, and C.-C. Nguyen, "1.51-GHz nanocrystalline diamond micromechanical disk resonator with material-mismatched isolating support," in *Micro Electro Mechanical Systems, 2004. 17th IEEE International Conference on. (MEMS)*, 2004, pp. 641 – 644.

[14] D. Gachon, E. Courjon, J. Masson, V. Petrini, J. Rauch, and S. Ballandras, "LiNbO3-LiNbO3 high overtone bulk acoustic resonator exhibiting high Q.f product," in *Ultrasonics Symposium, 2007. IEEE*, oct. 2007, pp. 1417 –1420.

[15] W. Pang, H. Zhang, J. J. Kim, H. Yu, and E. S. Kim, "High Q single-mode high-tone bulk acoustic resonator integrated with surface-micromachined FBAR filter," in *Microwave Symposium Digest, 2005 IEEE MTT-S International*, june 2005, p. 4 pp.

[16] H. Zhang, W. Pang, H. Yu, and E. S. Kim, "High-tone bulk acoustic resonators on sapphire, crystal quartz, fused silica, and silicon substrates," *Journal of Applied Physics*, vol. 99, no. 12, p. 124911(5), may 2006.

[17] Z. Hao and B. Liao, "An analytical study on interfacial dissipation in piezoelectric rectangular block resonators with in-plane longitudinal-mode vibrations," *Sensors and Actuators A: Physical*, vol. 163, no. 1, pp. 401 – 409, apr 2010.

[18] M. Rinaldi, C. Zuniga, C. Zuo, and G. Piazza, "Super-high-frequency two-port AlN contour-mode resonators for RF applications," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 57, no. 1, pp. 38 –45, jan. 2010.

[19] G. Piazza, P. Stephanou, and A. Pisano, "Piezoelectric aluminum nitride vibrating contour-mode MEMS resonators," *Microelectromechanical Systems, Journal of*, vol. 15, no. 6, pp. 1406 –1418, dec. 2006.

Tunable Silicon Bulk Acoustic Resonators with Multi-Face AlN Transduction

Roozbeh Tabrizian and Farrokh Ayazi
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, USA
roozbeh@gatech.edu; ayazi@gatech.edu

***Abstract*—This paper presents tunable width-extensional mode bulk acoustic resonators that are piezoelectrically-actuated and sensed using thin layers of AlN on the sidewalls as well as the top surface. By using both longitudinal and transverse piezoelectric effects of conformally-sputtered AlN layers on sidewalls and top surface of a 20 μm thick resonator, a low motional resistance of ~35 Ω was achieved for a 100 MHz silicon resonator operating in air. The motional resistance is improved by at least 10x compared to similar devices with capacitive transduction. Furthermore, it is shown that the resonance frequency of these piezoelectrically-transduced devices can be tuned by varying the electric signal power from 0 to 7 dBm.**

I. Introduction

Lateral bulk acoustic wave (BAW) resonators implemented in single crystal silicon (SCS) are of great interest for signal processing applications. Since the resonance frequencies of such resonators are defined lithographically, devices with multiple frequencies can be implemented in the same batch and integrated with CMOS circuitry. Furthermore, superior acoustic properties of SCS such as high bulk acoustic wave velocity, small intrinsic dissipation in different frequency regimes [1, 2] and extended linear-elastic region alleviate the implementation of resonators with small form factor, improved power handling, and high quality factor (Q). Although advances in micromachining techniques (i.e., fabrication of small capacitive gaps, deposition of high quality piezoelectric thin films, etc.) have enabled the realization of high-frequency SCS BAW resonators [2], large motional resistance of these devices remains a major obstacle to utilize them in low insertion loss filters.

Most transduction techniques that have been used to actuate and sense silicon BAW resonators [3-5] require a relatively large DC polarization voltage (V_p) to provide sufficient electromechanical coupling required for low motional resistances. Piezoelectric transduction, on the other hand, has the advantage of high electromechanical coupling without requiring V_p. However, all demonstrations of lateral micromechanical resonators have so far used the transverse piezoelectric coefficient e_{31} of a planar piezoelectric layer such as AlN [6-9]. While low motional resistances have been demonstrated using this technique for AlN-on-Si resonators [6], achieving low resistances (< 50 Ω) has been challenging due to limited transverse piezoelectric coupling, which is mainly a result of smaller e_{31} compared to e_{33} and small transduction area. To overcome this limitation, high-order modes are used in high-frequency lateral BAW resonators to increase the transduction area, which in turn leads to a larger die area [8]. Moreover, the efficiency of transverse piezoelectric transduction of silicon lateral BAW resonators is significantly degraded by the increase in proportional thickness of silicon in the resonator stack. However, thick silicon substrates are desirable for improvement of Q, power handling and linearity. Furthermore, the resonance frequency of BAW devices with transverse piezoelectric transduction cannot be tuned without sacrificing transduction area and considerably degrading the motional resistance.

Using the technique presented here, in addition to the AlN layer on top surface of the resonator AlN layers on the vertical sidewalls of the resonator are simultaneously used to employ the larger longitudinal piezoelectric coefficient e_{33} to actuate and sense bulk acoustic waves through the resonator sidewalls. In this configuration since Mo electrodes on the sidewalls connect equi-stress areas of AlN layers, charge cancellation is substantially reduced. Moreover, since transduction occurs on two sidewalls in addition to the top surface, it is scalable with resonator thickness which makes this method preferable for high-frequency resonators with small width (frequency-determining dimension). Because the sidewall AlN layers are mainly responsible for actuation and sensing, the top AlN layer can be dedicated to tuning purposes without a considerable reduction in effective transduction area.

II. Multi-Face AlN Transduction

Figure 1 shows the cross-section of a silicon BAW resonator with multi-face AlN transduction. Since the top and bottom Mo electrodes are conformally deposited over the resonator surface, an input voltage signal results in electric fields E_{Top} and E_{SW} perpendicular to the top surface and sidewalls, respectively. E_{Top} and E_{SW} induce transverse and longitudinal mechanical stress in top and sidewall AlN layers, respectively (Fig. 1a), which results in excitation of the width-extensional bulk acoustic mode (Fig. 2). This resonance mode has an amplified strain in the AlN films on

978-1-61284-110-6/11/$26.00 ©2011 IEEE

top and sidewall of the resonator and induces electric charge on the Mo electrodes (Fig. 1b).

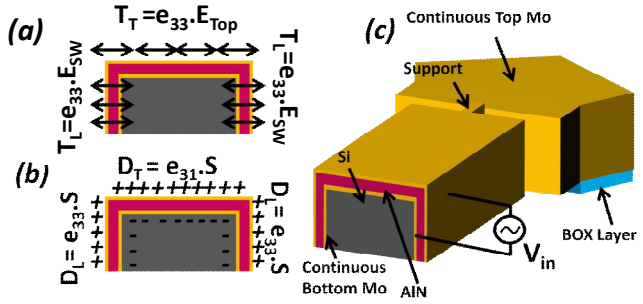


Figure 1: Silicon Bulk Acoustic Resonator (SiBAR) with Mo/AlN/Mo on top surface and sidewalls: (a) actuation mechanism; (b) sense mechanism; (c) resonator cross-section; E_{Top} and E_{SW} are applied electric fields and D_T and D_L are excited electric-displacement on top and sidewall AlN layers, respectively.

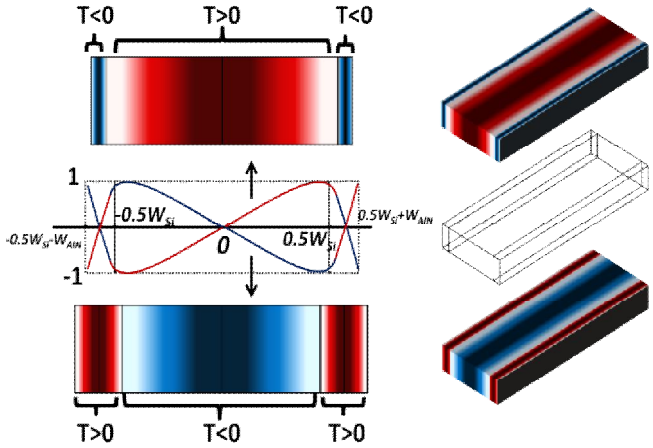


Figure 2: Stress field (T) and deformation in two half-cycles of resonance for a SiBAR with multi-face AlN transduction.

III. Modeling

Considering each AlN layer (two sidewall layers and one top layer) as a separate electromechanical transducer, a SiBAR with multi-face AlN transduction can be modeled as a three-port device. Figure 3 shows the general form electrical equivalent circuit of the resonator around the first width-extensional mode.

In this model, transformers represent the two piezoelectric transduction mechanisms with longitudinal (η_1) and transverse (η_2) coupling coefficients.

In the equivalent circuit of Fig. 3, D_1, K_1, and M_1, and D_2, K_2, and M_2 represent equivalent damping, stiffness and mass of the SiBAR with top AlN and sidewall AlN, respectively. C_{SW} and C_{Top} are the capacitances of sidewall and top AlN layers, respectively. The ratio between η_1 and η_2 can be estimated from (1):

$$\frac{\eta_1}{\eta_2} = \frac{|e_{33}|}{|e_{31}|} \cdot \frac{A_{SW}}{A_{AlN,Top}} \qquad (1)$$

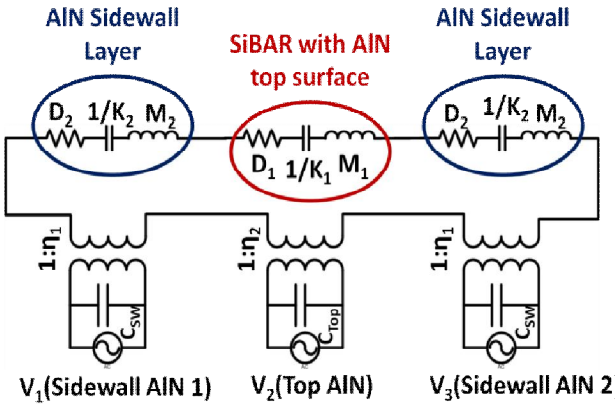


Figure 3: Electrical equivalent circuit for general SiBAR with multi-face AlN considering three ports for electromechanical transduction. V_1, V_2 and V_3 are shorted if the top Mo layer is continuous over the sidewall and top surfaces of the device, resulting in a one-port resonator.

where A_{SW} is the sidewall surface area and $A_{AlN,Top}$ represents the area of the vertical cross-section of top AlN layer.

Since the top Mo layer is conformally deposited on the sidewall and top surfaces (Fig. 1), all three ports are electrically connected together; hence, the resonator can be treated as a one-port device.

The resonator motional resistance R_m can be estimated from (2):

$$R_m = \frac{D_{tot}}{(2\eta_1 + \eta_2)^2} \qquad (2)$$

where D_{tot} represents the total mechanical damping in the resonator. Since $|e_{33}| \approx 3|e_{31}|$ and A_{SW} is much larger than $A_{AlN,Top}$, η_1 is much larger than η_2 (in this work $\eta_1 \approx 100\eta_2$), implying that actuation and sensing of the bulk acoustic resonance is more efficient using piezoelectric films on the sidewalls rather than the top surface. Thus, bulk acoustic resonators with sidewall piezoelectric transduction should provide a considerably lower R_m than those employing top surface AlN actuation and sensing. Additionally, since η_1 is directly proportional to A_{SW}, R_m can be further reduced by increasing resonator thickness. This is a superior advantage of sidewall transduction over top surface transduction, where a piezoelectric layer with large surface area is required to achieve small values of R_m.

IV. Fabrication Process

The fabrication process of one-port SiBARs with multi-face AlN layers consists of two masks and three steps (schematically shown in Fig. 4). First, the resonator body is patterned in the silicon device layer of a thick SOI wafer using the Bosch DRIE process. This is followed by thermal oxidation and annealing in N_2 ambient at 1100°C to improve the smoothness of sidewalls surface after Bosch DRIE. The BOX layer is then partially etched in HF to prevent deposition of Mo/AlN on the BOX layer in the next step. This is done to allow access of HF to the BOX layer during the final release step. Silicon structures are then covered with RF sputter deposition of a thin AlN layer sandwiched between top and

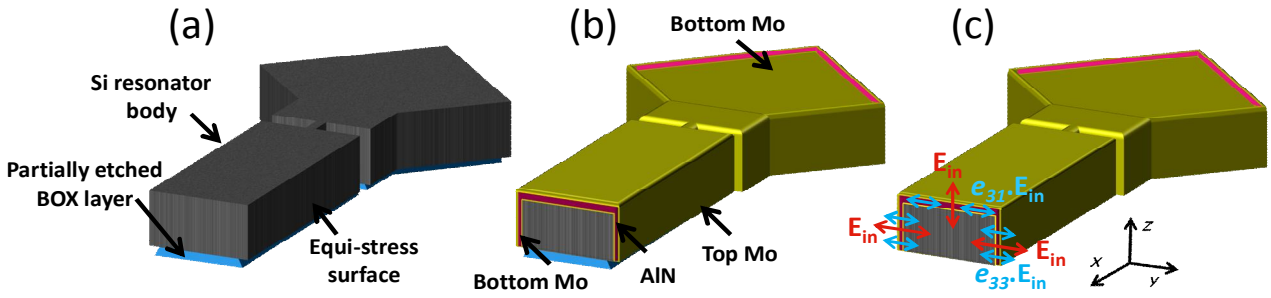


Figure 4: Resonator fabrication process: (a) patterning Si device layer and partially etching the BOX layer; (b) deposition of Mo/AlN/Mo layers, etching top Mo and AlN on one pad to expose ground contact; (c) releasing device in HF.

bottom Mo electrodes on the top surface and sidewalls. Then, the top Mo and AlN layers are etched from one pad to access the bottom Mo which serves as an electrical signal ground. Finally, the device is released by etching the BOX layer in HF. Figure 5 shows an SEM image of a fabricated one-port SiBAR with multi-face AlN and schematic of the electrical interface for measurement.

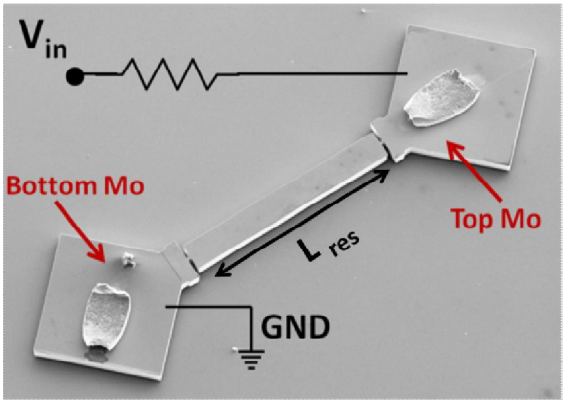


Figure 5: SEM image of the resonator and electrical interface scheme for measurement.

Figure 6 is the cross-sectional SEM picture of the resonator before release in HF. Figure 7 shows the continuity of AlN and Mo films from the top surface to sidewalls at the edges of the resonator, with sidewall AlN thicknesses approximately half of the top surface AlN.

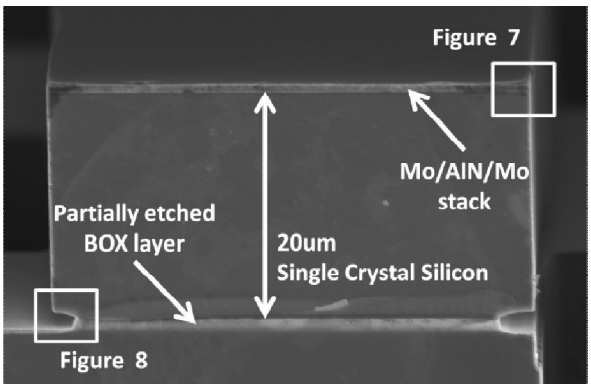


Figure 6: SEM image of the device cross-section; BOX layer partially etched to enable full release in HF.

Figure 8 shows the quality and uniformity of sidewall AlN towards the bottom of the resonator. Uniform thickness of the AlN sidewall film results in efficient actuation of the desired width-extensional resonance mode over the entire thickness of the resonator and eliminates charge cancellation in the sensing mechanism.

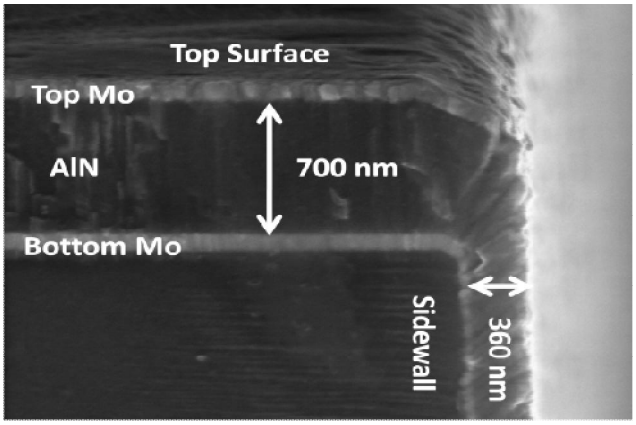


Figure 7: Continuity of AlN and Mo films from the top surface to sidewalls at the resonator edges.

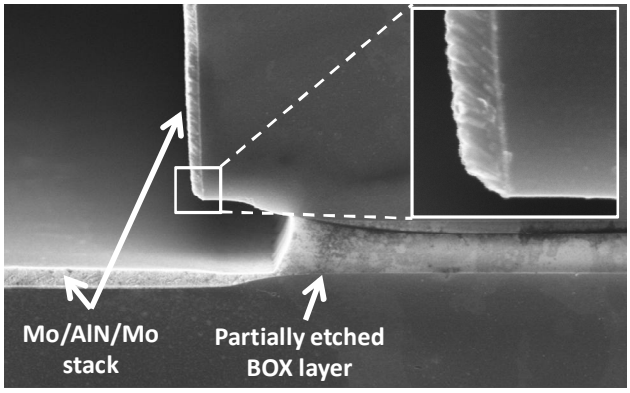


Figure 8: Close-up SEM image, detailing the quality and uniformity of sidewall AlN towards the bottom of the resonator.

V. Measurement Results

Experimental results are obtained using a vector network analyzer. Fig 9 shows the measured reflection coefficient

(S_{11}) as well as the resonator impedance extracted from S_{11} response of the one-port resonator of Fig. 5. A quality factor (Q) of 3500 can be extracted from [10]:

$$Q = \frac{\omega}{2R} \cdot \frac{\partial X}{\partial \omega} \qquad (3)$$

where ω is the angular frequency and R and X are the real and imaginary components of resonator impedance, respectively. An R_m of 35 Ω has been achieved at 99.85 MHz, which is at least ten times less than its capacitive and transverse piezoelectric counterparts [3-6].

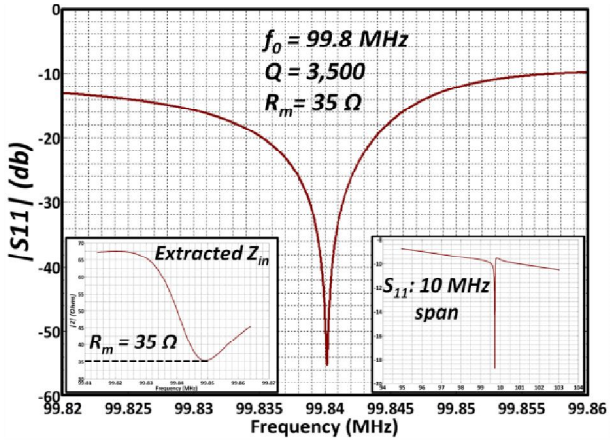


Figure 9: Reflection coefficient ($|S_{11}|$) of the resonator in Fig. 5; (insets) $|S_{11}|$ resonator impedance and response over a 10 MHz span.

Figure 10 shows the resonator frequency shift for input powers higher than 0 dBm, where leakage current increases in sidewall AlN thin layers [11]. This leakage current changes the electric field boundary condition in the top AlN layer as a result of the finite resistance of the excitation source (schematically shown in the equivalent circuit of Fig. 10 inset). As shown in Fig. 10, tuning does not degrade the amplitude of reflection coefficient ($|S_{11}|$) and the motional resistance of the one-port device.

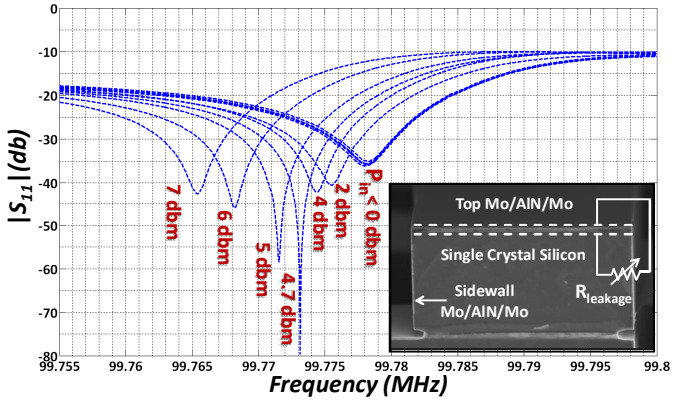


Figure 10: Tuning of AlN-on-Si resonator with input power levels from 0 to 7 dBm; inset shows tuning concept.

VI. Conclusion

This work presents very low motional resistance tunable silicon bulk acoustic resonators with piezoelectric AlN layers on the sidewalls and top surface of the resonator. Exploiting the larger longitudinal piezoelectric coefficient of sidewall AlN layers as well as the substantially reduced charge cancellation as a result of the electrode configuration, this technique demonstrates very high transduction efficiency for silicon bulk acoustic resonators. The electrical equivalent model is presented to demonstrate the advantage of this transduction technique as well as the possibility of efficiency improvement by using thick silicon substrates. Resonators with R_m values as low as 35 Ω have been implemented at 100 MHz using a two-mask, three-step fabrication process. The feasibility of tuning these devices without degrading R_m has been demonstrated by sweeping the applied signal power from 0 to 7 dBm.

Acknowledgment

This work was supported by Integrated Device Technology (IDT). The authors wish to thank the OEM Group for AlN film deposition and the staff at the Nanotechnology Research Center (NRC) at the Georgia Institute of Technology for assistance with microfabrication.

References

[1] R. Tabrizian, M. Rais-zadeh and F. Ayazi, "Effect of Phonon Interactions on Limiting the f.Q product of micromechanical resonators", *IEEE International Conference on Solid-state Sensors, Actuators and Microsystems (TRANSDUCERS'09)*, Denver, CO, June 2009, pp. 2131-2134.

[2] F. Ayazi, L. Sorenson, and R. Tabrizian, "Energy Dissipation in Micromechanical Resonators," in Proc. SPIE, 2011, pp. 1-13.

[3] A. K. Samarao and F. Ayazi, "Quality Factor Sensitivity to Crystallographic Misalignments in Silicon Micromechanical Resonators", *Solid-State Sensors, Actuators, and Microsystems Workshop (Hilton Head 2010)*, Hilton Head Island, SC, June 2010, pp. 479-482.

[4] S. A. Bhave and R. T. Howe, "Internal electrostatic transduction for bulk-mode MEMS resonators", *Solid-State Sensors, Actuators, and Microsystems Workshop (Hilton Head 2004)*, Hilton Head Island, SC, June, 2004, pp. 59-60.

[5] A. Rahafrooz, *et al.*, "Thermal actuation, a suitable mechanism for high-frequency electromechanical resonators", *Proc. 23st IEEE MEMS, Hong Kong, China*, Jan. 2010, pp. 200-203.

[6] W. Pan and F. Ayazi, "Thin-film piezoelectric-on-substrate resonators with Q enhancement and TCF reduction", *23rd IEEE MEMS*, pp 727-30, 2010.

[7] M. Rinaldi, *et al.*, "Super-high-frequency two-port AlN contour-mode resonators for RF applications", *Ultrasonic, Ferroelectrics and Frequency Control IEEE Transactions,* vol. 57, pp. 38-45, 2010.

[8] B. P. Harrington, *et al.*, "Toward ultimate performance in GHz MEMS resonators: Low impedance and high", *Proc. IEEE Micro Electro Mechanical Systems Conference*, Jan. 2010, pp. 707–710.

[9] G. K. Ho, *et al.*, "Piezoelectric-on-Silicon Lateral Bulk Acoustic Wave Micromechanical Resonators", *IEEE Journal of Microelectromechanical Systems, Vol. 17, No. 2,* April 2008, pp. 512-520.

[10] C.G. Montgomery, *et al.*, "Principles of Microwave Circuits", New York, McGraw-Hill (1948).

[11] F. Martin, P. Muralt, M. A. Dubois, and A. Pezous, "Thickness Dependence of Properties of Highly C-Axis Textured AlN Thin Films", J. Vac. Sci. Technol. A 22, 361 (2004).

Voltage-Controlled Tuning to Optimize MEMS Resonator Array-Composite Output Power

Mehmet Akgul, Zeying Ren, Clark T.-C. Nguyen
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA, USA

***Abstract*—A voltage controlled electrical stiffness tuning method has been demonstrated to correct phase and amplitude mismatches between the constituent resonators in a half-wavelength (λ/2) mechanically coupled array-composite towards maximizing its output power. Via tuning, a nine-disk array-composite using 3 output resonators achieves an output current 2.91× larger than that of a single one of its constituent resonators, and only a bit short of the 3× theoretical maximum. Without tuning, the array-composite achieves only 2.78× the current of a single device, and the deviation from ideal is expected to increase with the number of resonators in the array. The amount of tuning available can be tailored in numerous ways, from sizing of electrode-to-disk gap spacing, to specifying the number of devices in the array involved with tuning, to simple variation of voltages across selected electrode-to-resonator gaps. By raising the power output of a high-*Q* micromechanical disk-array composite resonator, the method and design of this work stand to greatly lower the phase noise of oscillators referenced to such devices.**

I. Introduction

Among the various transduction methods [1]-[4] used by micromechanical resonators, capacitive transduction has historically achieved the highest *Q*'s, with values reaching past 160,000 at VHF [3] and 14,600 at 1.2 GHz [4]. Such high *Q*'s, combined with small size and the potential for on-chip integration with CMOS, make this technology very attractive for various applications, from very narrowband low insertion loss filters for RF frequency gating spectrum analyzer implementation [5]; to low phase noise, low power oscillators for radar and communications [6].

Still, adaptation of capacitive micromechanical resonators has so far been slowed by their higher than conventional impedances and limited power handling capability (governed by linearity) relative to other technologies, such as piezoelectric resonators. However, there is no theoretical reason why the impedances and linearities of capacitive resonators cannot achieve the desired levels. In fact, a 60 MHz capacitive resonator recently attained an impedance as low as 140Ω on par with piezoelectric resonators, but with a much higher *Q* >70,000 [7]. The power handling and linearity performance of capacitive transducers has also been studied extensively [8][9], with single 156 MHz capacitively transduced disk resonators demonstrating measured IIP_3's of 19.5dBm that meet GSM requirements.

If even better linearity is needed for more demanding applications, mechanically coupled array-composite resonators, cf. Figure 1, that sum the output currents of identical mechanically coupled resonators designed to resonate in phase, have been shown to improve performance dramatically over stand-alone devices [3][10]. Specifically, arraying reduces motional resistance by increasing the total output current for the same input voltage amplitude; and simultaneously improves linearity by reducing the resonator displacement required to source a given amount of power, thereby reducing the percent of the electrode-to-resonator gap traversed by the mechanical structure. Ideally, arraying improves the linearity and motional resistance by factors equal to the number of resonators used in the array [3]. However, due to fabrication non-idealities, random variations in the resonator and coupling beam dimensions across the array often compromise the degree to which output currents actually add. Because of this, previous array-composites have fallen short of their expected impedance and linearity improvements, especially when the number of resonators used in the array is large [3][10].

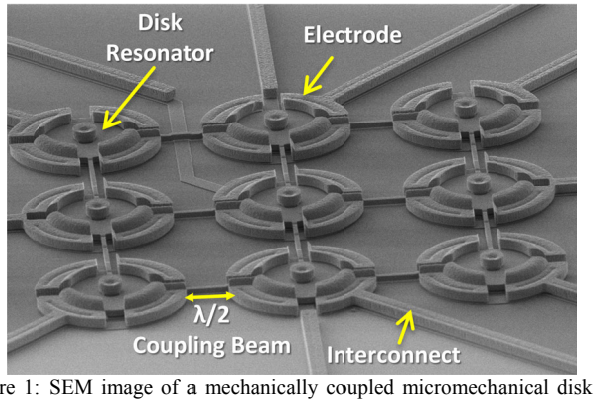


Figure 1: SEM image of a mechanically coupled micromechanical disk array-composite with individual electrode access to each of the resonators constituting the array.

This work attempts to better understand the effect of mismatches between the elements of an array-composite resonator by investigating how stiffness tuning some of the resonators in the array influences the amplitude and phase matching of resonators being summed. In particular, voltage-controlled electrical stiffness tuning applied to several resonators in the array-composite generates maxima in the individual and combined output currents measured off three resonators of the array when 1) their phases are matched; and 2) their phases are close to that of the input drive signal. Tuning over a 7V range induces a measured 16% change in combined array-composite output current, which corresponds to a 1.3dB change in output power that can impact the linearity of a channel-select filter or

The authors would like to thank DARPA for supporting this work.

$$R_x = \frac{f_o \gamma \rho}{2 Q h \varepsilon_o^2} \frac{d_{gap}^4}{V_P^2} \quad (1)$$

f_o : Resonance frequency
h : Resonator thickness
ε_o : Permittivity of vacuum
V_P : DC polarization voltage
γ : Dynamic-to-physical mass ratio
m_{re} : Dynamic mass $m_{re} = \gamma \pi R_d^2 h \rho$
d_{gap} : Electrode-to-resonator gap spacing
ρ : Density R_d : Disk radius Q : Quality factor

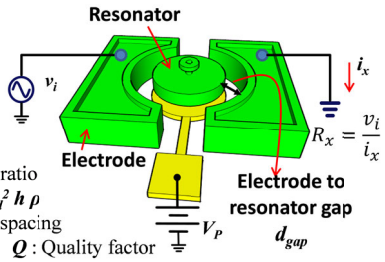


Figure 2: Schematic of a contour mode disk resonator, identifying important design variables and providing the expression for the effective motional resistance R_x of the device.

the phase noise of a high-Q MEMS-based oscillator.

II. High Frequency, High Q, Low Impedance Capacitively Transduced Mechanical Resonators

This work utilizes capacitively transduced radial-contour mode micromechanical disk resonators [11], such as shown in Figure 2. This device consists of a conductive disk suspended by a center stem 700 nm above an electrically contacted ground plane and surrounded by electrodes spaced only 80 nm from the disk edges. A voltage applied to the ground plane, i.e., V_P in the figure, is effectively also applied to the disk itself and creates voltage drops across the disk and each of its electrodes that then activates the capacitive transducer. Under normal operation, a combination of a dc-bias voltage V_P and an ac excitation voltage v_i is applied across one electrode (the input electrode), as shown in the figure. When the frequency of v_i matches the resonance frequency of the disk, the disk begins to vibrate, generating a dc-biased (with V_P) time-varying capacitance at the output electrode that then sources an output current i_o.

Although often maligned for supposedly insufficient electromechanical coupling compared with other transducer choices, e.g., piezoelectric, capacitive transducers are actually among the strongest of the bunch when dimensions scale down to tens of nanometers. To see this, one can employ the expressions for electrode-to-resonator overlap capacitance C_o, motional capacitance C_x, and resonance frequency f_o, for a radial-contour mode disk resonator, given by [11],

$$C_o = \frac{\varepsilon_o A_o}{d_{gap}} \;,\; C_x = \frac{V_P^2}{\omega_o^2 m_{re}} \frac{\varepsilon_o^2 A_o^2}{d_{gap}^4} \;,\; f_o = \frac{\alpha}{R_d} \sqrt{\frac{E}{\rho}} \qquad (2)$$

where A_o is the electrode-to-resonator overlap area, ω_o is resonance frequency in radians, α is a mode shape dependent constant that equals 0.342 for the fundamental contour mode shape in polysilicon, E is the Young's modulus, and other variables are defined in Figure 2. Combining these yields the equivalent k_t^2 for a capacitively transduced disk resonator:

$$k_t^2 = \frac{\pi^2}{2} \frac{C_x}{C_o} = \frac{1}{4 f_o} \frac{\varepsilon_o}{\gamma \alpha} \frac{1}{\sqrt{E \rho}} \frac{V_P^2}{d_{gap}^3} \qquad (3)$$

From (3), the electrode-to-resonator gap spacing clearly offers a very strong knob by which to tailor electromechanical coupling. Plotting (3) for the case of the 159-MHz disk used in this work versus electrode-to-resonator gap spacing alongside the range for k_t^2 of contour-mode AlN piezoelectric resonators [12][13] yields the curves of Figure 3. The right side of the curve, where electrode-to-resonator gap spacings are large, reveals why capacitive transducers were so maligned: the electromechanical coupling is orders of magnitude lower than that

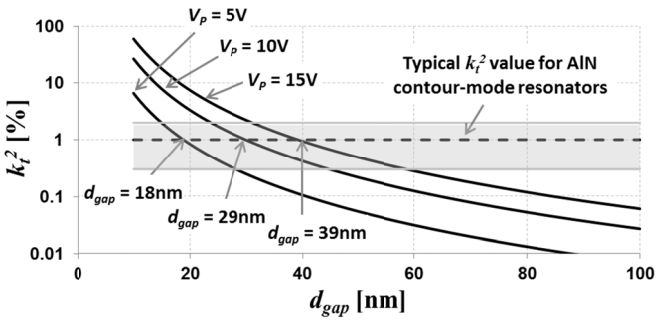


Figure 3: Simulated plot of k_t^2 for a polysilicon contour mode disk resonator operating at 159 MHz, plotted as a function of the electrode-to-resonator gap spacing for three different bias voltages. The shaded band indicates the k_t^2 range for AlN contour-mode resonators in the literature [12][13].

of AlN piezoelectric resonators. However, the left side, where the gap spacing shrinks to tens of nanometers, tells a completely different story. In particular, for the case of V_P =10V, when the gap spacing falls below about 29 nm, a complete reversal occurs, where the electromechanical coupling of the capacitive transducer begins to exceed the typical 1% of AlN contour-mode resonators. The increase in electromechanical coupling with decreasing gap spacing was recently verified in [7], where the electrode-to-resonator gap of a 61-MHz wine-glass disk was reduced to 37 nm from 92 nm to effect a change in motional resistance from 5,350Ω to 140Ω, all while still retaining a Q of 73,173—the only on-chip room temperature resonator in existence able to attain simultaneous high Q >70,000 and low impedance < 150Ω.

Unfortunately, however, the motional resistance of the resonator, as given by the equation in Figure 2, like its coupling coefficient, is inversely proportional to resonance frequency. In other words, the same 37 nm gap that allowed an R_x = 140Ω at 61 MHz in [7] would yield a much higher R_x of 2.3kΩ at 1 GHz. To counteract this increase, more aggressive gap scaling is needed. For instance, to achieve R_x = 50Ω for a polysilicon disk resonator operating at 1 GHz with an assumed Q of 30,000 and a dc bias voltage V_P =10V, the required gap is 9 nm. Needless to say, this is a very small gap that not only begs the question of whether or not such a small gap can be achieved with high yield, but that also raises nonlinearity concerns, since the 3rd-order intermodulation intercept point IIP_3 starts to drop quickly when the gap passes a certain threshold, effectively imposing a "gap scaling limit" [9].

When the gap scaling limit is met, but the impedance requirement still has not been, the next best course of action is to employ one of the other available knobs: either the dc-bias voltage V_P or the disk thickness h. If the supply voltage is limited, then thickness must be increased, which is unfortunate, since this would require a higher aspect ratio, so becomes increasingly difficult as lateral dimensions shrink to attain higher frequencies. Fortunately, the thickness knob is really a knob to change electrode-to-resonator overlap area A_o, which can be increased without increasing thickness by merely using many resonators so that their total combined overlap area increases. This, of course, is the main idea behind the mechanically coupled disk array-composites under study, here.

III. Mechanically Coupled Array-Composite

By summing the output currents of multiple resonators to

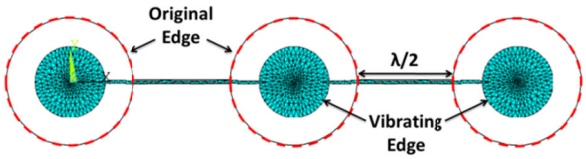


Figure 4: ANSYS modal analysis result for three identical disk resonators mechanically coupled with half-wavelength coupling beams.

attain a larger overlap area and in turn produce a larger total output current, an *N*-resonator array lowers the effective motional resistance linearly with the array size *N*, as described by

$$R_{x,Array} = \frac{v_i}{i_o} = \frac{v_i}{N \times i_x} = \frac{R_x}{N} \qquad (4)$$

where i_x and R_x are the motional current and resistance of a single resonator, respectively; and v_i is input voltage. Note that (4) is valid only when all the resonators generate identical output currents that are in phase and at precisely the same frequency. Given the very high *Q*'s >10,000 of the individual resonators, minute differences between the resonators likely to occur during any practical microfabrication process create slight frequency differences that prevent output currents from adding in phase. As a result, simply connecting the output electrodes of discrete resonators together is not an effective arraying method.

A much better approach uses mechanical coupling beams to physically link the arrayed resonators, as shown in Figure 1. A mechanically linked *N*-resonator array forms an *N* degree-of-freedom system with *N* modal frequencies, each associated with a specific mode shape [10]. This ensures that all resonators resonate at the exact same frequency for a given excited mode, very conveniently eliminating frequency mismatches, allowing same-frequency summing of resonator outputs. Still, most applications, e.g., oscillators, require a frequency response with only one output frequency, not *N* of them. So although its higher power output is welcome, the *N* peaks generated by this mechanically coupled array are not.

One way to eliminate the undesired modes is to use mechanical coupling beams that act as springs with infinite effective stiffness [3]. Specifically, since the coupling beams behave like mechanical transmission lines at this frequency, they can be given infinite effective stiffnesses by merely dimensioning them to correspond to a half-wavelength ($\lambda/2$) at the desired mode frequency. More specifically, the coupling beam lengths should be set equal to an odd multiple of the half-wavelength ($\lambda/2$) of the longitudinal waves travelling in the resonator medium with an acoustic velocity of $\sqrt{E/\rho}$:

$$L_{beam} = \frac{n}{2f_o}\sqrt{\frac{E}{\rho}},\ n = 1,3,5,\ldots \qquad (5)$$

where f_o is the resonance frequency . With effectively infinitely rigid couplers, the mechanically coupled resonators behave as a single array-composite resonator, selectively resonating at a single mode frequency that takes on a shape in which all the resonators move in phase, as shown in Figure 4. The remaining modes are (ideally) pushed to infinity.

In effect, an ideal half-wavelength coupled array of resonators, with its single mode frequency response, effectively realizes an array-composite resonator that behaves like a single of its constituent resonators, except with lower impedance, better linearity, and better power handling capability.

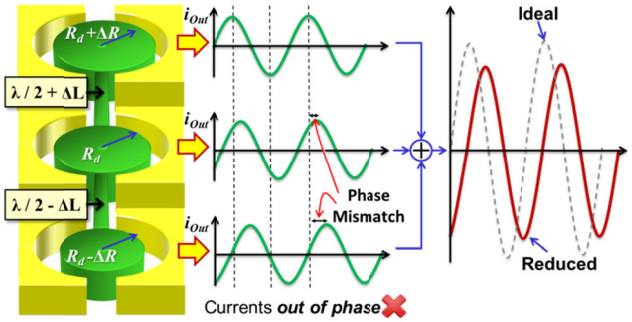


Figure 5: Schematic description of the effect of mismatches between resonators and coupling beams on array performance.

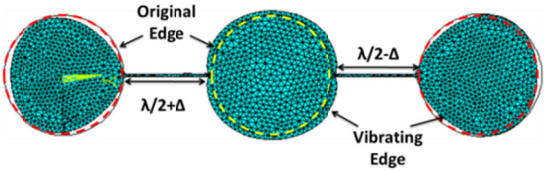


Figure 6: ANSYS simulation for a three-resonator array with 1% error in disk radii and coupling beam lengths in a worst case mismatch scenario.

IV. Effect of Non-Idealities on Arrays

The analysis in the previous section assumes an ideal case, where all the disks that form the array are identical and all coupling beam lengths exactly match the $\lambda/2$ condition of (5) at the common resonance frequency of the disks. In this perfect scenario, all disks resonate in phase and generate identical output currents that lead to the advertised *N*-times reduction in R_x. However, in practice, various process variations in lithography, etching, etc., lead to deviations from a perfect match between the array elements. As a result, the disk radii and coupling beam lengths in the array structure have a random distribution around the ideal design values, as illustrated in Figure 5. Hence, each coupled disk would have a slightly different resonance frequency were it to resonate by itself. Furthermore, variations in the coupling beam lengths create deviations from the half-wavelength condition, meaning the beams are no longer infinitely rigid couplers. Although the mechanical coupling beams still impose a common resonance frequency, they now do so with the following consequences:

1. Different disks reach their resonance displacement maxima at different times. Thus, the output currents generated by individual resonators are not precisely in phase.
2. Mismatches between coupled resonators and coupling beam lengths distort the mode shapes of constituent resonators, causing them to deviate from the ideal circular contour mode shape, as shown in Figure 6, which presents the mode shape predicted by ANSYS for a three-resonator array in a worst case mismatch scenario with 1% error in the disk radii and coupling beam lengths. The result is significant distortion of the mode shape compared to the ideal case of Figure 4. This mode shape distortion diminishes the resonance amplitude of the resonators, which then leads to reduced output current.
3. Spurious modes ensue that are no longer suppressed by infinite stiffness coupling beams.

As a result, even though all the resonators in the array generate an output current precisely at the same frequency—a

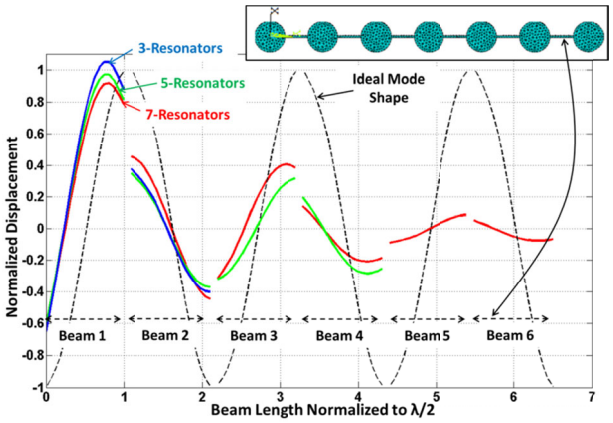


Figure 7: ANSYS harmonic analysis showing coupling beam mode shape displacements (normalized to the ideal mode shape) for mechanically non-ideal coupled disk resonator arrays with varying array sizes (3, 5, and 7 resonators, as marked) for a 1% worst-case dimensional mismatch between the array elements; compared to the ideal mode shape plotted as the dashed curve.

requirement imposed by mechanical coupling—the individual output currents are smaller and out-of-phase, so sum to a value lower than the theoretical maximum, as shown in Figure 5. This underperformance will likely get worse as frequencies rise, since smaller dimensions would likely amplify mismatches, and the larger array sizes needed to maintain low impedance and good linearity would further distort the individual element mode shapes, as verified by the FEA-generated plot of Figure 7. Indeed, any method for tuning out mismatches becomes more important at higher frequencies.

V. Voltage Controlled Electrical Stiffness Tuning of Resonator Array-Composites

Since frequency and stiffness deviations comprise the ultimate consequence of process-induced non-ideality, controlling stiffness is perhaps the best way to counteract such non-idealities. The capacitive transducers used in this work offer a very convenient method for controlling stiffness via mere adjustment of the voltages applied across their capacitive gaps, which changes the electrical stiffness across these gaps. Electrical stiffness [14] arises from the electric field force between the conductive resonator and its parallel-plate electrode that rises and falls as the resonator gets close and far, respectively, from the electrode. This force is in phase with resonator motion, so comprises an effective electrical stiffness k_e, given by

$$k_e = \frac{\varepsilon_o A_o}{d_{gap}^3} V_\Delta^2 \qquad (6)$$

where A_o is the electrode area, and V_Δ is the DC voltage difference between the electrode and the resonator. Note that this electrical stiffness acts to increase the resonator's displacement, so acts to oppose the mechanical stiffness of the resonator. Its effect on the resonance frequency of a single resonator can thus be modeled by

$$f_o = \frac{1}{2\pi}\sqrt{\frac{k_{re}}{m_{re}}} = \frac{1}{2\pi}\sqrt{\frac{k_m - k_e}{m_{re}}} \qquad (7)$$

where k_{re} is the total effective resonator stiffness at the location of the tuning electrode (i.e., usually at its midpoint), obtained by subtracting the electrical stiffness k_e from the mechanical stiffness at that location k_m; and m_{re} denotes the equivalent dynamic mass at that location. Note that for contour mode disks, location is not an issue, but it can be for more general mode shapes, e.g., wine-glass [3][6].

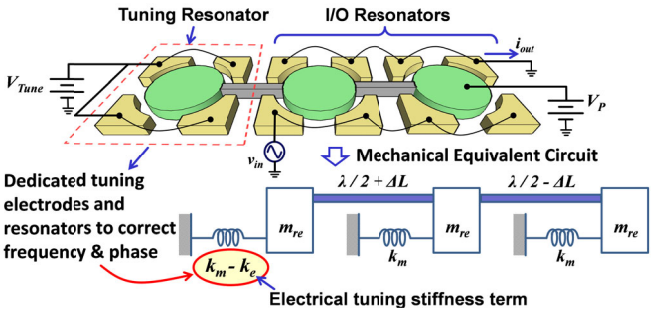


Figure 8: Schematic description of the voltage controlled electrical stiffness tuning method applied to selected tuning electrodes of an array-composite resonator to electrically tune out mechanical mismatches.

Equation (7) applies very well to individual resonators and has been used successfully to correct the passbands of small percent bandwidth filters, such as in [15], where constituent resonators are coupled by quarter-wavelength beams. For the case of the half-wavelength coupled array resonators of this work, however, (7) must be modified to account for the fact that a half-wavelength coupled array realizes an array-composite resonator with an effective combined stiffness equal to the sum of the stiffnesses of its constituents. Thus, a change in electrical stiffness applied to one resonator, such as shown in Figure 8, affects them all, i.e., affects the whole array-composite. For an N-resonator array-composite, where M resonators are used for input/output and (N-M) for frequency tuning, the expression for the array-composite frequency becomes

$$f_o = \frac{1}{2\pi}\sqrt{\frac{Nk_{m,s} - Mk_{e,V_P} - (N-M)k_{e,V_T}}{Nm_{re,s}}} \qquad (8)$$

$$k_{e,V_P} = \frac{\varepsilon_o\, 2\pi\, R_d\, h}{d_{gap}^3} V_P^2, \quad k_{e,V_T} = \frac{\varepsilon_o\, 2\pi\, R_d\, h}{d_{gap}^3}(V_{Tune} - V_P)^2 \qquad (9)$$

where $k_{m,s}$ and $m_{re,s}$ are the dynamic stiffness (with no bias voltage) and mass for a single resonator, respectively; while k_{e,V_T} and k_{e,V_P} are the electrical stiffness for a single resonator with and without the tuning voltage V_{Tune} applied to its electrodes, respectively.

Equations (8)-(9) indicate that the amount of tuning available can be tailored in numerous ways, from sizing of electrode-to-disk gap spacing, to specifying the number of devices in the array involved with tuning, to simple variation of voltages across selected electrode-to-resonator gaps. The voltage-control knob provides an especially convenient method to tune individual frequencies so as to minimize mode shape mismatches between the elements of a mechanically coupled resonator array. Note that a designer has a choice of whether to use a constituent resonator in an array composite as an input device, an output device, a frequency tuning device, or any combination of these, including all three.

It should also be noted that the arrays of Figure 1 and Figure 9 do not offer the ability to directly tune the coupling beams. Such tuning is possible with the right design, but most such designs become somewhat cumbersome, so are not implemented here. If coupler tuners are not available, then at least the frequency of the array-composite can be adjusted to minimize the average coupler deviation from half-wavelength at whatever their dimensions. In effect, pulling the resonance

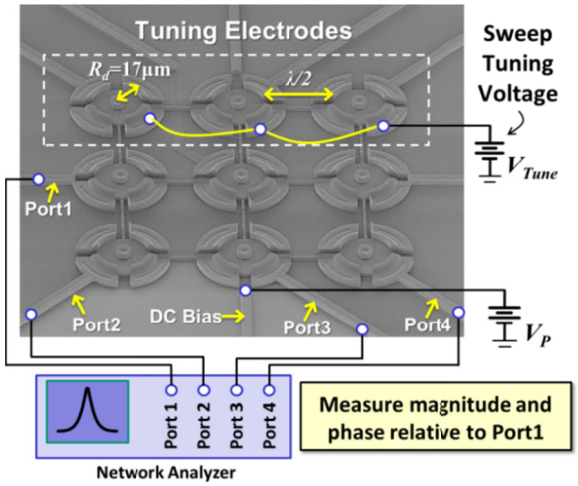


Figure 9: SEM of the fabricated and measured resonator array-composite tuning test structure and schematic of the experimental setup.

frequency closer to the optimum value that provides an average best fit match to the $\lambda/2$ coupling condition for the actual fabricated devices with non-idealities minimizes mode shape distortions and the resulting phase mismatches between individual disk output currents, hence maximizes array output.

VI. Experimental Results

To observe the extent of mismatches stemming from fabrication-induced non-idealities between resonators and the mechanical coupling beams that are otherwise matched in layout, disk array-composites were fabricated in polysilicon using a five-mask fabrication process similar to that of [16]. Figure 9 presents the scanning electron micrograph of a fabricated array structure that differs from previous ones [3] in that (1) it includes electrodes that can individually address each constituent resonator, so provides a mechanism for studying the degree to which an actual fabricated array-composite resonator deviates from the ideal, i.e., the degree to which the phases and resonance amplitudes of its constituent resonators are matched; and (2) it features electrode-resonator combinations specifically equipped to tune the effective electrical stiffness of the entire array structure via adjustment of tuning voltages. The structure consists of nine 17μm-radius contour mode disk resonators designed to operate at 159.8 MHz, with each disk mechanically coupled to its neighbors via half-wavelength beams of length L_{beam} = 26μm to form a 3×3 array-composite resonator.

Figure 9 also includes the test set-up used to gauge the impact of process non-idealities on the disk array-composite and evaluate the efficacy of the proposed corrective tuning method. Here, an Agilent E5071C four-port network analyzer drives the array-composite device into resonance from an input resonator, denoted as Port 1; and simultaneously records the phase and magnitude information of the output currents generated by three output resonators, denoted as Ports 2-3-4. A DC polarization voltage of V_P = 5V is applied common to all disk resonators, and three of the resonators in the array serve as dedicated electrical stiffness tuners, to which a DC tuning voltage separate from V_P is applied to effect selectable amounts of stiffness control. By sweeping the tuning voltage while recording with the network analyzer, the change in the

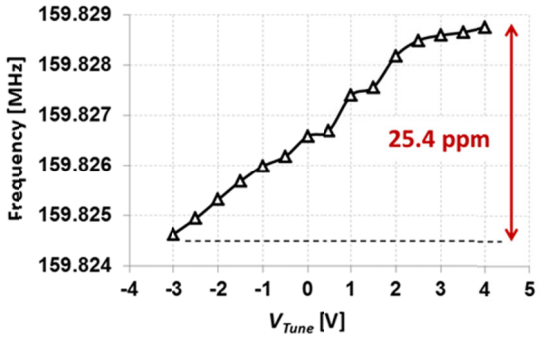


Figure 10: Measured plot of array-composite frequency vs. tuning voltage.

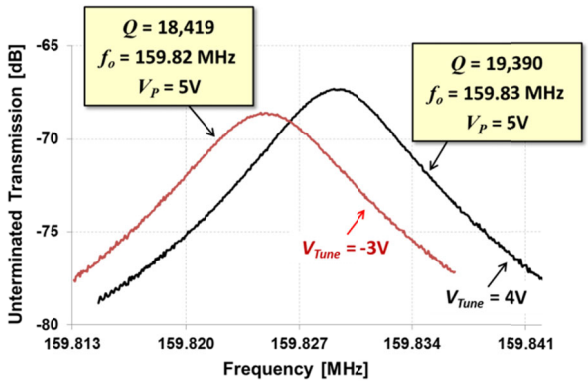


Figure 11: Frequency characteristics for the combined port 2-3-4 output of the array-composite measured at two different tuning voltages.

phase and magnitude of the individual resonators can be plotted as a function of the applied tuning voltage.

To first gauge the amount of tuning provided by the three tuning electrodes, Figure 10 presents a plot of array-composite frequency versus applied tuning voltage. Here, a tuning voltage excursion of 7V provides a 25.4 ppm change in frequency. Although small, this tuning range was still enough to affect noticeable changes in total array output current, as indicated by Figure 11, which plots measured frequency characteristics for the combined output currents from Ports 2-3-4 as a function of the tuning voltage. Here, a 1.3dB change in output is seen over a 7V excursion in tuning voltage.

To better elucidate some of the mechanisms behind tuning-based output improvement, Figure 12 presents measured curves of magnitude and phase of the individual currents generated by each output resonator (ports 2, 3, and 4) plotted as a function of the applied tuning voltage. There are two groups of curves presented in this plot: the solid curves indicate the phase mismatch of each output resonator current relative to the input resonator drive signal, with reference to the *y*-axis on the left; and the dashed curves present the output current amplitudes measured in dBm, with reference to the *y*-axis on the right. The results clearly demonstrate reduction of the phase mismatch between output resonators and input resonator, by more than 20° for the output resonator denoted Port2 as its phase mismatch reduces to almost zero, and by more than 10° for the other two output resonators.

For each output resonator, two maxima are observed: one where the phase of the resonator in question has its lowest phase mismatch with the input resonator; and one where all output resonators are matched in phase. Clearly, phase match-

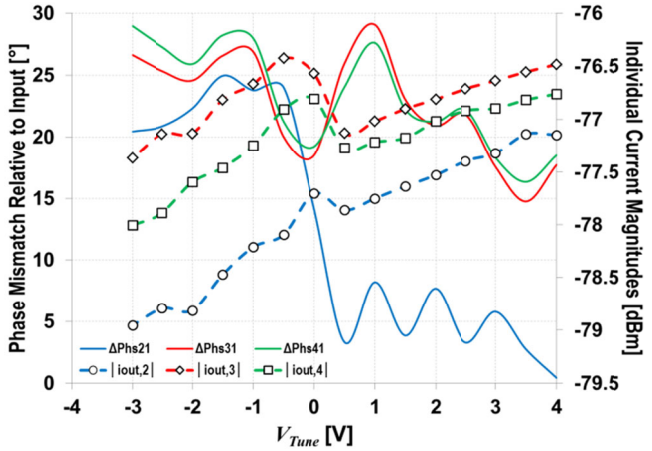


Figure 12: Measured individual output current phase mismatch relative to input (solid line curves – left *y*-axis) and amplitudes (dashed line curves – right *y*-axis) for each output resonator (ports 2, 3, and 4) vs. tuning voltage.

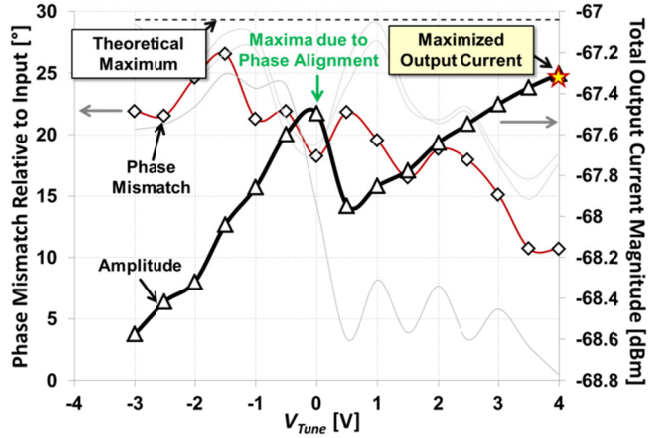


Figure 13: Measured curves of combined total output current phase mismatch relative to input (◇ data markers– left *y*-axis) and amplitude (△ data markers – right *y*-axis), plotted against the applied tuning voltage.

ing is an important consideration for optimized performance of even individual resonators making up an array composite.

Figure 13 plots the phase and amplitude of the total combined output current from the three output resonators against the applied tuning voltage and against a line indicating the maximum output current magnitude expected from an ideal array-composite with no mismatches. Interestingly, for the case of this nine-resonator array-composite, three of which are used for output, the output current is within 2 dB of the ideal over a 7V range of tuning voltage. This might raise questions on the need for tuning at all, but a quick glance at Figure 7 reminds one that a larger array would likely experience a much larger current drop for which correction by tuning would be important. In addition, where transmit power is concerned, every 0.1dB of loss is generally of great concern, especially in space applications. At any rate, Figure 13 attests that voltage-controlled electrical stiffness tuning is clearly an effective means for controlling and minimizing phase mismatches, and more importantly for maximizing the total output current amplitude.

VII. Conclusions

By restoring output power to levels closer to ideal, the described voltage-controlled electrical stiffness tuning method to correct phase and resonance amplitude mismatches between the constituent resonators in mechanically coupled array-composites allows such mechanical circuits to achieve their true potential, even in the midst of finite fabrication tolerances. Clearly, some more refinement in the tuning method, e.g., to allow more precise tuning of individual resonators, would be beneficial, and investigation into just how large an array can be tuned via the described methods is needed. Work towards these continues with the goal of someday making possible assembly of on-chip array-composites employing thousands of capacitively transduced micromechanical resonators to achieve motional resistances down to 1Ω, IIP_3's greater than 50 dBm, and power outputs in the several Watt range, all while retaining the high Q >30,000 typical of such devices.

References

[1] R. C. Ruby, P. Bradley, Y. Oshmyansky, A. Chien, and J. D. Larson III, "Thin film bulk wave acoustic resonators (FBAR) for wireless apps.," in IEEE Ultrasonics Symposium, 2001, vol. 1, p. 813–821.

[2] G. Piazza, "Integrated AlN piezoelectric microelectromechanical system for radio front ends," J. of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, vol. 27, no. 4, p. 776, 2009.

[3] Y. W. Lin, S. S. Li, Z. Ren, and C. T. C. Nguyen, "Low phase noise array-composite micromechanical wine-glass disk oscillator," in Electron Devices Meeting, 2005. IEDM Technical Digest, p. 4–281.

[4] Sheng-Shian Li, Yu-Wei Lin, Yuan Xie,Zeying Ren, C. T.-C. Nguyen, "Micromechanical 'hollow-disk' ring resonators", 17th IEEE Int. Conf. on MEMS, 2004, Tech. Digest, Maastricht, Netherlands, pp. 821-824.

[5] C. T.-C. Nguyen, "Integrated micromechanical RF circuits for software-defined cognitive radio", 26th Symposium on Sensors, Micromachines & App. Sys., Tokyo, Japan, Oct. 15-16, 2009, pp. 1-5.

[6] Y.-W. Lin, S. Lee, S.-S. Li, Y. Xie, Z. Ren, C. T.-C. Nguyen, "Series-resonant VHF micromechanical resonator reference oscillators," IEEE J. Solid-State Circuits, vol. 39, no. 12, pp. 2477-2491, Dec. 2004.

[7] Akgul M., Kim B., "Capacitively transduced micromechanical resonators w/ simultaneous low motional resistance and Q > 70,000", Solid-St. Sens., Actuators and Microsys. Workshop, Hilton Head 2010

[8] R. Navid, J. R. Clark, M. Demirci and C. T. -C. Nguyen, "Third-Order Intermodulation Distortion in Capacitively-Driven CC-Beam Micromechanical Resonators," IEEE MEMS Conf., 2001, Switzerland.

[9] Y. W. Lin, S. S. Li, Z. Ren, C. T. C. Nguyen, "Third-order intermodulation distortion in capacitively driven VHF micromechanical resonators," in IEEE Int. Ultrasonics Sym.,2005, p.18–21.

[10] M. U. Demirci, C. T.-C. Nguyen, "Mechanically Corner-Coupled Square Microresonator Array for Reduced Series Motional Resistance," J. of MEMS, vol. 15, no. 6, pp. 1419-1436, Dec. 2006.

[11] J. R. Clark, W.-T. Hsu, M. A. Abdelmoneum, C. T.-C. Nguyen, "High-Q UHF micromechanical radial-contour mode disk resonators", J. Microelectromech. Syst., vol. 14, no. 6, pp. 1298-1310, Dec. 2005.

[12] G. Piazza, P.J. Stephanou, A.P. Pisano, "One and two port piezoelectric higher order contour-mode MEMS resonators for mechanical signal processing," Solid-State Electronics, vol. 51, 2007, p. 1596–1608.

[13] M. Rinaldi, C. Zuniga, and G. Piazza, "5-10 GHz AlN contour-mode nanoelectromechanical resonators", IEEE 22nd International Conference on MEMS, 2009, p. 916–919.

[14] Harvey C. Nathanson, William E. Newell, Robert A. Wickstrom, and John R. Davis, Jr., "The resonant gate transistor", IEEE Trans. Electron Devices, vol. ED-14, p. 117, 1967.

[15] K. Wang and C. T.-C. Nguyen, "High-order medium frequency micromechanical electronic filters," IEEE/ASME J. Microelectromech. Syst., vol. 8, no. 4, pp. 534-557, Dec. 1999.

[16] J. Wang, Z. Ren, and C. T.-C. Nguyen, "1.156-GHz self-aligned vibrating micromechanical disk resonator," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 51, no. 12, pp. 1607-1628, Dec. 2004.

Acoustic Bragg Reflectors for *Q*-Enhancement of Unreleased MEMS Resonators

Wentao Wang, Dana Weinstein
HybridMEMS Lab, Massachusetts Institute of Technology
Cambridge, MA, USA
Email: wangwt@mit.edu

***Abstract*—This work presents the design of acoustic Bragg reflectors (ABRs) for unreleased MEMS resonators through analysis and simulation. Two of the greatest challenges to the successful implementation of MEMS are those of packaging and integration with integrated circuits. Development of unreleased RF MEMS resonators at the transistor level of the CMOS stack will enable direct integration into front-end-of-line (FEOL) processing, making these devices an attractive choice for on-chip signal generation and signal processing. The use of ABRs in unreleased resonators reduces spurious modes while maintaining high quality factors. Analysis on unreleased resonators using ABRs covers design principles, effects of fabrication variation, and comparison to released devices. Additionally, ABR-based unreleased resonators are compared with unreleased resonators enhanced using phononic crystals, showing order of magnitude higher quality factor (*Q*) for the ABR-based devices.**

I. Introduction

The need for monolithic integration of Micro Electromechanical (MEM) devices with CMOS is critical for successful implementation of high frequency active MEMS resonators. Monolithic integration of these devices can provide basic RF and mm-wave building blocks with high *Q*, small footprint, and low power for use in wireless communication, microprocessor clocking, navigation and sensing applications. Nevertheless, the majority of MEMS resonators require a release step to freely suspend the moving structures. This necessitates complex encapsulation and packaging, restricting fabrication to MEMS-last or Back-End-of-Line (BEOL) processing of large-scale devices [1].

Toward the goal of Front-End-of-Line (FEOL) CMOS integration, the authors have previously demonstrated the first fully unreleased MEMS resonator operating at 39 GHz with *Q* of 129 [2]. The Si bulk acoustic resonator, surrounded on all sides by SiO_2, demonstrates the feasibility of direct integration of resonator into the FEOL CMOS processing, offering on-chip low power clock generation and high-*Q* tank filters.

Whether unreleased RF MEMS resonators are realized in a CMOS or custom process, their implementation provides high yield, low cost, robustness in harsh environments, and minimal or no-packaging solutions.

This work is funded by the MARCO Center for Materials, Structures and Devices (MSD) and the DARPA Young Faculty Award.

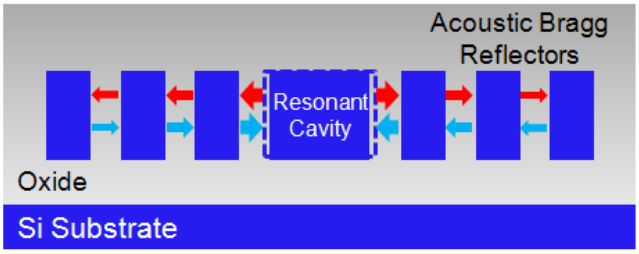


Figure 1. Sideview of unreleased resonator and ABRs lithographically defined in the same mask process. The resonant mode is longitudinal in the horizontal direction, exciting plane waves which are confined by the ABRs. Oxide cladding the resonator and ABRs illustrates the structures achievable in standard SOI CMOS processing.

As a result of the leakage of acoustic energy by wave paths into the surrounding medium, the performance of an unreleased resonator is degraded relative to its released counterpart [3]. This energy loss can be mitigated by adding acoustic isolation structures. In this work, two types of structures are proposed and investigated for the localization of acoustic vibrations in place of λ/4 suspension beams found in most released structures: the acoustic Bragg reflector (ABR) and the phononic crystal (PnC), with a focus on the former.

The acoustic Bragg reflector is named after its optical analogue, which is used as thin film optical mirrors to improve reflectivity at a designed wavelength. The ABR is composed of periodic layers of two materials of low and high acoustic impedance. At each layer interface of the ABR, a fraction of the acoustic wave energy is reflected. To form a coherent superposition of these reflected waves, these layers can be optimized at odd multiples of quarter wavelength, resulting in the overall reflectivity identical to that of free or rigid boundary conditions.

In 1965, W. E. Newell introduced the idea of employing ABRs to thickness mode piezoelectric resonators for wireless applications in high frequency integrated circuits [4]. This work was further developed in fully integrated surface micromachining technology for solidly mounted resonators (SMRs) [5]. These ABRs are composed of multiple depositions of alternating materials, resulting in acoustic isolation in one dimension at a single frequency per wafer. In-plane isolation can be achieved using lithographically defined

ABRs, as demonstrated in a suspended plate [6]. This configuration enables resonators of multiple frequencies to be fabricated side by side on the same chip.

PnCs provide an alternative solution to lithographically defined acoustic isolation. They have been recently explored by several groups for microscale applications, including acoustic mirrors for resonators in suspended plates [7], acoustic waveguides, and filters [8].

This paper presents the study of fully unreleased resonator surrounded by lithographically defined ABRs, embedded in a homogeneous cladding layer (Fig. 1). This one-mask design enables resonator banks of various frequencies on the same chip, providing multiple degrees of freedom in ABR design. With the goal of direct integration into FEOL CMOS processing, resonator performance is investigated for materials commonly found in the CMOS stack. The characteristics of these unreleased structures are compared with freely suspended resonators, released resonators isolated with lithographically defined ABRs, and PnC based unreleased resonators.

II. Bandgap of Acoustic Bragg Reflectors

Existing ABR analysis is based on a transmission line analogy, borrowing results from impedance transforming properties of cascaded transmission lines [4]. However, a more fundamental analysis can be performed conveniently from the wave propagation point of view through superposition, as shown in Fig. 2.

The acoustic impedance Z of an elastic medium for longitudinal wave is defined by material properties as

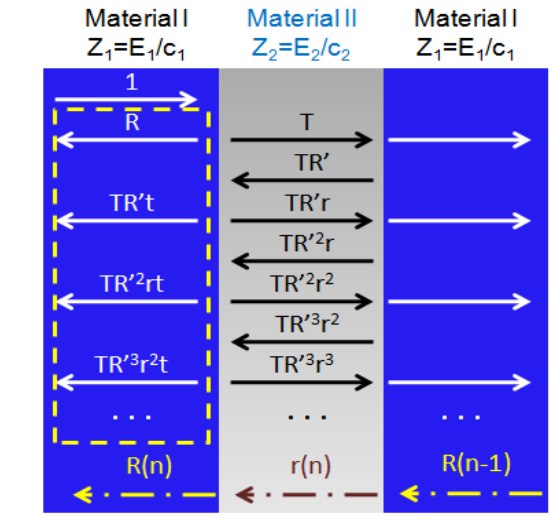
$$Z = \frac{E}{c} = \sqrt{\frac{(1+\nu)(1-2\nu)\rho E}{(1-\nu)}} \qquad (1)$$

where E, ρ, ν are the Young's modulus, density, and Poisson's ratio respectively.

In Fig. 2, an acoustic wave travelling to the right in material I is incident on an interface with material II, separated into reflected and transmitted components of R and T. At the next interface, the transmitted component T is reflected at an assumed lumped reflectivity of R' from the following stack of materials. The reflected wave now travelling to the left with amplitude TR' is then reflected (r) and transmitted (t) at the first interface. As a result, infinite reflections occur inside material II with amplitude decreasing in geometric progression. By summation of all the reflections, we acquire the recursion formula

$$R(n) = R + \frac{Ttr(n)e^{-i2\delta_2}}{1 - rr(n)e^{-i2\delta_2}} \qquad (2)$$

in which δ_2 is the phase change over the propagation length in material II, r(n) and R(n) are the lumped reflectivities of the nth ABR bi-layer as shown in Fig. 2. Similarly, the other recursion formula can be acquired:



$$R = \frac{Z_1 - Z_2}{Z_1 + Z_2} \quad T = \frac{2Z_1}{Z_1 + Z_2} \quad r = \frac{Z_2 - Z_1}{Z_1 + Z_2} \quad t = \frac{2Z_2}{Z_1 + Z_2}$$

$$R' = r(n) \qquad R(n) = R + \frac{Ttr(n)e^{-i2\delta}}{1 - rr(n)e^{-i2\delta}}$$

Figure 2. Derivation of recursion formula of reflectivity for ABRs through wave superposition. A summation of all the reflected components in the dashed line box generates relationship of lumped reflectivity R(n) and r(n) of the nth ABR bi-layer.

$$r(n) = r + \frac{tTR(n-1)e^{-i2\delta_1}}{1 - RR(n-1)e^{-i2\delta_1}} \qquad (3)$$

With these two relations, the properties of the ABR can be extracted through numerical calculation. Fig. 3 shows the result of ABR reflectivity of various bi-layer numbers, in which the ABR bi-layer consists of Si and SiO_2.

An ideal ABR composed of an infinite number of bi-layers is approximated numerically using 100 ABR pairs. As shown in Fig. 3, this ideal ABR provides total reflection over several frequency ranges called the bandgaps. The reflectivity in the bandgap is strongly dependent on the number of ABR pairs. However, numerical results show that only 7 bi-layers are enough to approximate the bandgap shape, providing 98.7% reflectivity at the center of bandgap.

The convergence rate toward total reflection differs based on the materials composing the ABR structure. Material pairs with higher acoustic impedance contrast ratio (Z_1/Z_2) form a wider bandgap, and require fewer ABR pairs for same level of reflectance. The performance of several quarter-wavelength material pairs available in CMOS are exhibited in Fig. 4. The sign of reflectivity is determined by the order of materials in the ABR bi-layer. If the first material in the bi-layer has a higher acoustic impedance, the overall ABR will approach a free boundary condition (positive reflectivity). On the other hand, if the first material has a lower acoustic impedance, it will approach fixed boundary condition (negative reflectivity).

Materials of large acoustic impedance contrast such as W/SiO_2 [9], AlN/SiO_2 [5], Mo/Al [10] have previously been used to form ABR pairs in SMRs. This large contrast

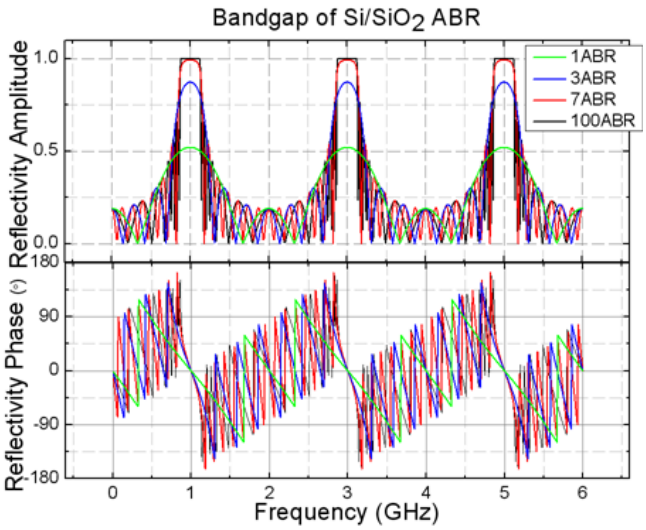


Figure 3. Frequency response of ABR amplitude (top) and phase (bottom). The bandgap approaches ideal shape with increasing number of ABRs. When the ABR pair number is small, the bandgap is attenuated and bandwidth increases.

minimizes the number of pairs required to achieve good acoustic isolation, reducing the number of deposition steps during fabrication. This will not complicate fabrication for unreleased resonator ABR design, since the entire ABR structure can be defined lithographically in one step. Si/SiO_2 ABRs are selected for the design and analysis in the paper, due to their FEOL CMOS compatibility, low internal loss, and small achievable footprint.

III. Acoustic Bragg Reflectors for Unreleased Resonators

The acoustic energy loss analyzed here in the unreleased resonators is analogous to anchor loss in freely suspended devices. In most released and unreleased resonators, this energy localization can be the limiting mechanism for Q. Accordingly, this analysis is focused on the mechanical domain, taking maximum or average strain as signal output. This mechanical-only analysis in turn allows the result to be applied for various sensing mechanisms that rely on the electrical signal modulated by strain, such as dielectric, piezoelectric, piezoresistive, or Field Effect Transistor sensing [11].

A. 1D Analysis of Unreleased Resonator with ABRs

Using the superposition method outlined in section II, a one dimensional unreleased resonator embedded in an Si/SiO_2 ABR stack can be analyzed numerically using the maximum strain as the output (Fig. 5). This is a best-case ideal resonator output; in practical design the mode cannot be perfectly one-dimensional. Finite width may introduce non-idealities such as plate modes and extra wave leakage, as discussed in the following sections.

In this analysis, the resonator is excited by a pair of equal and opposite forces on both edges, and after superposing wave components in both directions, strain is calculated at the center of the resonator. Resonant peaks appear periodically at odd harmonics, which agrees well with the shape of the ABR bandgap. The Q of the peak increases rapidly with the increase of ABR pairs as the reflectivity approaches 1. These properties confirm the ABR analysis results in Fig. 3.

Because both the lengths of the ABR layers and resonant cavity are lithographically defined, the dimensions are subject to fabrication variations. Over-etching may change the relative thickness of the Si layer to the SiO2 layer in one ABR periodicity, which subsequently shifts performance of the entire system including frequency, Q, and amplitude of vibrations. However, the finite bandgap generated by the ABR provides some level of tolerance to such variations. Analysis indicates that fabrication tolerances of 10-15% results in up to 10% degradation in targeted performance. The fabrication variation affects performance more as the ABR pair number increases, which is reasonable due to the narrowing of equivalent bandgap as shown in Fig. 3.

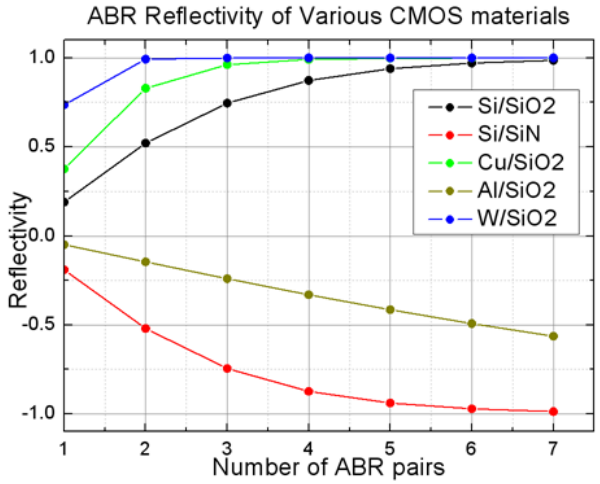


Figure 4. For finite number of ABR pairs, the reflectivity increases with the pair number and converges to 1 or -1 rapidly for materials commonly found in CMOS.

Figure 5. Numerical calculation results of maximum strain output of unreleased resonator embedded in ABR. The strain is normalized to the maximum strain at 7 ABR pairs.

B. *Comparison of Released and Unreleased Resonator with ABRs*

Two dimensional finite element simulations were performed in COMSOL Multiphysics 4.1 to demonstrate the contribution of ABRs to unreleased resonator performance. The frequency response of ABR-enhanced unreleased structures is compared with freely suspended resonators, simple unreleased resonators (no ABRs), and released resonators isolated with lithographically defined ABRs (Fig. 6). All resonators, driven by a pair of equal and opposite forces on both edges, are designed for 1st harmonic longitudinal vibrations at 1 GHz. The output signal is taken as the average strain across the entire resonant cavity. Perfectly matched layers (PMLs) are imposed at boundaries corresponding to energy sinks for waves propagating to infinity.

When defining PML parameters in COMSOL, it is recommended to match the PML to the desired acoustic wavelength. Although energy absorption is only optimized around the targeted wavelength, comparison of simulations with varying PML settings has shown robustness of PML performance over a broad range of frequencies for a single matched wavelength. For example, simulations of the resonator shown in Fig. 6(b, c) were performed with PMLs

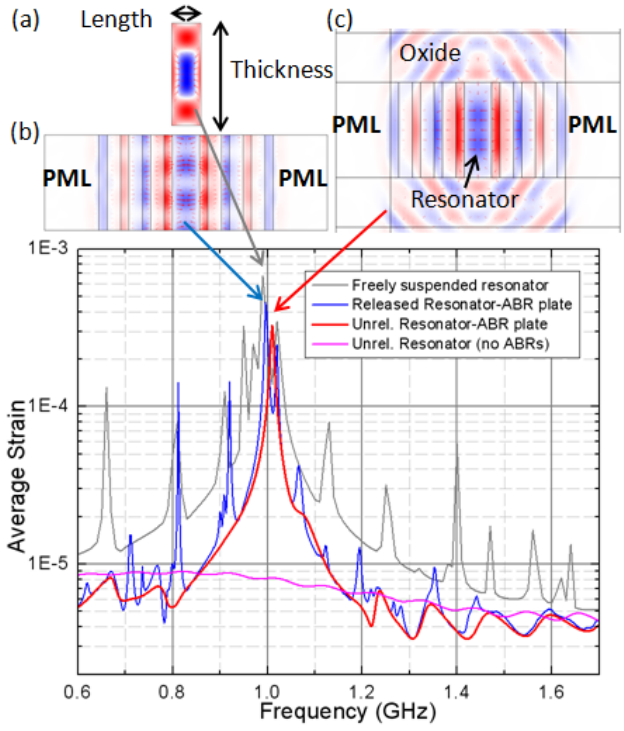


Figure 6. Frequency sweep in COMSOL for sideviews of free bar (a), released resonator-ABR plate (b), unreleased resonator-ABR structure embedded in oxide (c), and simple unreleased resonator embedded in oxide. Frequency response is for resonators of thickness-length ratio of 6.5 and ABR number of 7. Characteristic x-strain contour plots are given at an aspect ratio of 3.5 and ABR number of 3.

targeted at 1 GHz and 3 GHz wavelength, demonstrating maximum variation of less than 2% between the two simulations across a 3 GHz frequency sweep.

The released bar (Fig. 6(a)) provides the sharpest peak, which diverges for infinitesimal frequency steps due to the absence of internal damping in simulation. It is evident that due to the excitation of various plate modes, there are numerous spurious modes around the targeted frequency. At the other extreme, the unreleased resonator in oxide provides almost no peak at all, due to homogeneous energy dissipation into the surrounding medium. Comparison is made focusing on the released resonator-ABR plate (Fig. 6(b)) and the unreleased resonator-ABR structure (Fig. 6(c)). The output signal of the unreleased structure approaches that of the released one with increasing thickness-length aspect ratio. At an aspect ratio of 6.5, the unreleased resonator-ABR structure provides an output signal only 20% less than its released counterpart. The unreleased device also exhibits clear spurious mode suppression, providing a single peak at the desired frequency of 1 GHz.

Of all configurations, the unreleased resonator-ABR structure provides the purest mode due to damping of undesired plate modes in the non-resonant direction. Depending on the aspect ratio of the device, this out-of-plane damping may contribute to a reduced Q of the targeted mode, with a high aspect ratio being favorable for low loss. Nevertheless, it only requires an aspect ratio of approximately 7 for the output of unreleased resonator-ABR structure to match that of the released resonator-ABR plate. This effect results from the higher thickness uniformity of the mode present in the unreleased resonator (Fig. 6(c)), providing a larger effective area for sensing.

C. *In-Plane ABR Design*

To this point, analysis and optimization has been addressed for the sideview structure of both released and unreleased resonators. The lithographic definition of the ABRs under consideration affords a third degree of freedom in design to optimize for the quality factor and suppression of spurious modes. In the sideview case in the previous section, the resonator and ABRs are assumed to extend infinitely in the third dimension. In practice, the finite length of the structure introduces non-idealities and damping, which can be mitigated by in-plane ABR design.

Fig. 7 presents several in-plane ABR configurations for a 1D 2nd harmonic longitudinal-mode unreleased resonator. One half of the resonator is used for driving and the other half for sensing, providing maximum transduction area and selective excitation of the 2nd harmonic. As in the case of released resonator-ABR plate, the extra reflectors result in standing waves in the direction perpendicular to the longitudinal wave, causing signal cancellation and reduction in Q relative to the case of the infinite structure. Inclusion of corner ABRs also introduces spurious plate modes as exhibited in the inset in Fig. 7.

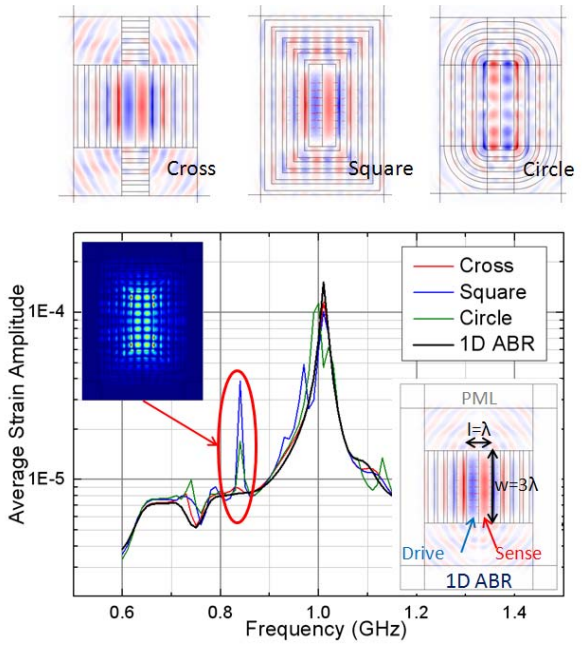


Figure 7. Frequency sweep around the 2nd harmonic in COMSOL on layout designs of cross, square, circular and 1D ABR structures (topview). Adding ABRs for longitudinal waves on the sides of the resonator does not enhance *Q*. Instead, it tends to introduce spurious modes. The inset presents the spurious mode for the square layout.

This result suggests that the 2D ABR design should serve to minimize any extra 2D in-plane vibrations and target a single mode. In the case of the longitudinal-mode resonator with 1D ABRs, in order to localize energy without mode distortion, side reflectors should coherently reflect only shear waves rather than longitudinal waves (Fig. 8). The shear reflectors do not act as standing wave boundary conditions to distort the 1D longitudinal mode, thereby providing higher *Q* and larger output signal.

This enhancement is more evident for higher harmonics. The reduced wavelength of the 6th harmonic in Fig. 8 results in a more uniform mode along the width of the resonator. In this case, the vibrations near the sides of the resonator are increased such that the contribution from the side ABRs is larger. As can be seen in Fig. 8, the 6th harmonic at 3 GHz exhibits a significantly larger amplitude and higher *Q* than the 2nd harmonic at 1 GHz due to the increased width to wavelength ratio. In this case, the shear ABRs also serve to enhance the *Q* by a factor of 2 relative to the 1D ABRs and by a factor of 4 relative to the cross configuration. This quality factor can be further enhanced by increasing the thickness-to-height aspect ratio of the resonator and surrounding ABRs.

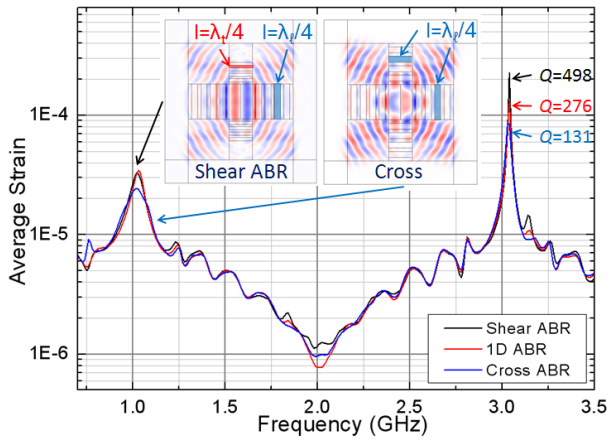


Figure 8. Frequency sweep in COMSOL including the 2nd and the 6th harmonic on unreleased resonator with side reflectors (topview). The width-length aspect ratio of the unreleased resonator is 1.5. Designing side reflectors at quarter shear wavelengths provides more uniform mode and larger signal output.

IV. Acoustic Bragg Reflectors vs. Phononic Crystals

The ABR can be understood as one dimensional case of the phononic crystal (PnC). The PnC is an acoustic wave analogue of photonic crystals, where a periodic array of scattering elements in a homogeneous matrix material causes complete rejection of acoustic waves at a band of frequencies. Much attention has been drawn to the implementation of PnCs for MEMS resonators. One widely used structure is an in-plane wave transmission line resonator analogous to the Fabry–Pérot interferometer [7]. Typically, air holes are etched as scattering elements in a silicon slab (or slab made of other materials) to form a PnC structure. There are also solid scattering approaches to fill these holes with materials with distinct acoustic impedance from the slab [12].

Solid scattering design can be utilized to localize vibrations in unreleased resonators. To compare the performance of a 1D ABR structure to that of a PnC, a silicon square lattice PnC with oxide scattering elements is considered. The fabrication process would be identical to that of the ABR, except for a change in layout, with holes in place

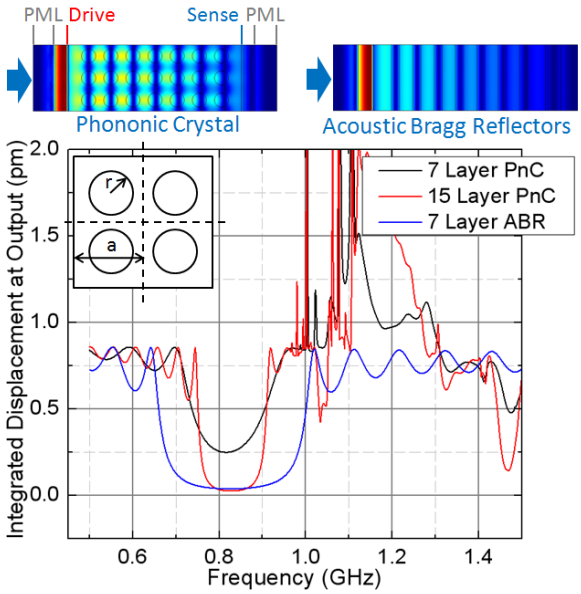


Figure 9. Acoustic bandgap comparison between phononic crystal and acoustic Bragg reflectors. The ABR provides a wider bandgap with no eigenmodes beyond the bandgap, and it requires a smaller footprint. The inset shows a 2x2 unit cell for the PnC used in this analysis. The contour plots show the topview of the transmission line structure, with amplitude of displacement decreasing with penetration depth.

of trenches defined by lithography. Similarly to the ABR, the central frequency of the bandgap for the PnC is dominantly dependent on periodicity, or lattice constant 'a', which approximately equals a half wavelength, while the width of bandgap is related to the filling ratio r/a (Fig. 9, inset). A COMSOL parameter sweep of the bandgap with respect to filling ratio shows the widest bandgap at filling ratio of ~0.3 for Si/SiO_2.

To compare the bandgap formed by the ABR and PnC, a transmission line configuration is investigated in COMSOL, with acoustic excitations generated on the left-hand side and sensing on the right, enclosed in PMLs on both ends (Fig. 9). Periodic boundary conditions are added on top and bottom so that the bandgap property is not distorted by the width dimension. The comparison result shows that the ABR has several advantages over the PnC for unreleased resonators:

- ABRs provide a much wider bandgap than that of PnCs for given impedance contrast of materials and for the same footprint.
- To achieve a near-perfect bandgap, it requires much fewer layers for the ABR compared to the PnC, requiring a smaller footprint.
- At frequencies beyond the bandgap, there exist spurious eigenmodes for the PnC, which results in undesired strong resonance. On the other hand, the ABR provides a perfect bandgap without introducing spurious modes.

As an outcome of these advantages, for the same footprint and material set, if we compare the PnC or ABR applied to an unreleased resonator driven on both ends, the ABR offers clear benefits, with a 9x higher Q, 20x larger signal output, and suppression of spurious modes at both low and high frequency (Fig. 10).

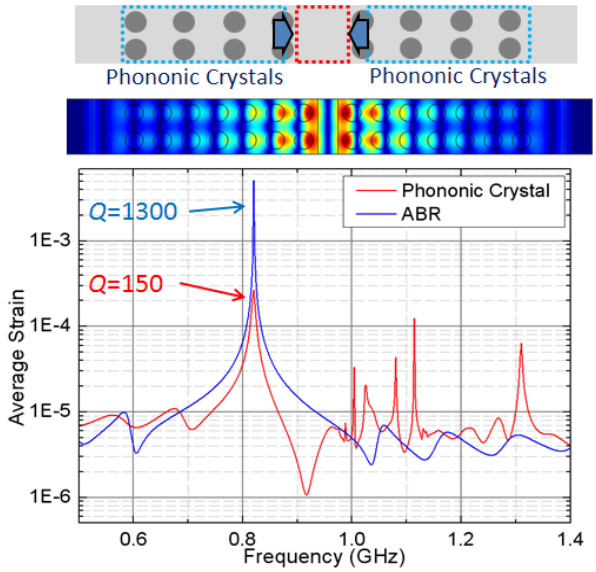


Figure 10. Comparison of performance of unreleased resonator embedded in the PnC and ABR. The contour plot is a topview simulation result in COMSOL with periodic boundary condition added on top and bottom. A pair of equal and opposite forces is applied on both edges of the resonator.

V. Conclusions

The design of the unreleased resonator with ABRs was optimized to suppress spurious modes and enhance the quality factor. With sufficient thickness-to-length aspect ratio, the unreleased design is able to provide signal level and Q that are comparable to its released counterparts. At high frequencies (>1 GHz) where resonator damping is dominated by anchor loss [13], the unreleased resonator with ABRs can outperform released devices.

Since the ABRs are lithographically defined in the same step as the resonator itself, their fabrication does not require any additional depositions or masks. In addition, it enables flexibility in ABR layout configuration, which can be implemented to optimize the ABR geometry for a single mode, resulting in high Q design with a finite footprint. Compared with the PnC, the ABR is more efficient in acoustic energy localization both in performance and footprint, and does not introduce any spurious modes.

The design of unreleased resonators enhanced with acoustic Bragg reflectors can provide basic building blocks for RF circuit designers with high-Q on-chip signal generation and processing that can be directly integrated into the FEOL in CMOS processing. Unreleased design also provides high yield, low cost, robustness in harsh environments, and minimal or no packaging for MEMS resonators.

References

[1] G.K. Fedder, R.T. Howe, T.-J. King Liu, E.P. Quevy, "Technologies for cofabricating MEMS and electronics," *Proceedings of the IEEE*, **96**(2), 306-322 (2008).

[2] W. Wang, L.C. Popa, R. Marathe, D. Weinstein, "An unreleased mm-wave resonant body transistor", *IEEE MEMS*, 1341-1344 (2011).

[3] D. Weinstein, S.A. Bhave, "Acoustic resonance in an independent-gate FinFET," *Hilton Head MEMS Workshop 2010*, 459-462 (2010).

[4] W.E. Newell, "Face-mounted piezoelectric resonators," *Proceedings of IEEE*, **53**(6), 575-581 (1965).

[5] K.M. Lakin, "Thin film resonators and filters," *IEEE Ultrasonics Symposium 1999*, 895-906 (1999).

[6] R.H. Olsson , J.G. Fleming, M.R. Tuck, "Contour mode resonators with acoustic reflectors," US patent 7385334 B1 (2008).

[7] S. Mohammadi, A.A. Eftekhar, W.D. Hunt, A. Adibi, "High-Q micromechanical resonators in a two-dimensional phononic crystal slab," *Applied Physics Letters*, **94**, 051906 (2009).

[8] R.H. Olsson, I. El-Kady, "Microfabricated phononic crystal devices and applications," *Measurement Science and Technology*, **20**, 012002 (2009).

[9] S. Marksteiner, J. Kaitila, G. Fattinger, R. Aigner, "Optimization of acoustic mirrors for solidly mounted BAW resonators," *IEEE Ultrasonic Symposium 2005*, 329–332 (2005).

[10] J. Enlund, D. Martin, V. Yantchev, I. Katardjiev, "Solidly mounted thin film electro-acoustic resonator utilizing a conductive Bragg reflector," *Sensors and Actuators A*, **141**, 598-602 (2008).

[11] D. Weinstein, S.A. Bhave, "The resonant body transistor," *Nano Letters* **10**(4), 1234-1237 (2010).

[12] R.H. Olsson III, J.G. Fleming, I.F. El-Kady, M.R. Tuck, F.B. McCormick, "Micromachined bulk wave acoustic bandgap devices," *IEEE Transducers 2007*, 317-321 (2007).

[13] R. Tabrizian, M. Rais-Zadeh, F. Ayazi, "Effect of phonon interactions on limiting the f.Q product of micromechanical resonators," *IEEE Transducers 2009*, 2131-2134 (2009).

Tunable Piezoelectric MEMS Resonators for Real-Time Clock

Diego Emilio Serrano, Roozbeh Tabrizian and Farrokh Ayazi
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA USA
diego.serrano@ece.gatech.edu; roozbeh.tabrizian@ece.gatech.edu; ayazi@ece.gatech.edu

***Abstract*—This paper reports on the design, simulation and characterization of small form factor, tunable piezoelectric MEMS resonators for real time clock applications (32.768 kHz). The structures were fabricated on a thin-film AlN-on-SOI substrate to enable piezoelectric actuation of an out-of-plane flexural mode, as well as electrostatic frequency tuning by utilizing the handle layer as a DC voltage electrode. Resonators of only a few hundred of µm in size exhibit greater than 3100 ppm of tuning using voltages no larger than 4 V; this tuning sufficiently compensates for frequency variations across temperature from -25 to 100 ºC. The devices exhibit low motional impedance that is completely independent of the tuning potential.**

I. INTRODUCTION

Advances in wireless communications and computer processing applications have led to the steady growth of the electronics timing market, which is currently at the $5 billion mark. This niche is mainly composed of resonators, oscillators and clock generators, whose implementation is dominated by quartz crystal technology. This has been partly due to historical reasons [1], but mainly because quartz exhibits excellent temperature-stable characteristics and reasonable quality factor values, which are desired for high-accuracy applications [2]. In recent years, silicon-based MEMS resonators are becoming strong contenders in the timing applications market as a result of their small form factors, high fQ products and the possibility of direct integration with interface circuits [3]. However, most research efforts in this field are presently being invested in high-frequency devices in the MHz to GHz range, leaving conventional tuning fork quartz resonators as the prevailing devices for real-time clock (RTC) applications. The low-frequency operation (32.768 kHz) and precise fabrication steps required for the implementation of quartz tuning forks make these structures large in size with surface areas in the millimeter range [4]. Additionally, to achieve vacuum levels necessary for high-Q performance, quartz resonators require metallic encapsulation, which further complicates their manufacturing and increases cost.

At such low frequencies, silicon microresonators operating in flexural modes can be utilized to overcome these complications; however, these devices exhibit large temperature coefficients of frequency (TCF) of about -30 ppm/ºC, so passive compensation techniques [5, 6] and frequency-tuning methods are required to counteract this effect [7].

Several techniques for passive temperature compensation of frequency in flexural silicon MEMS resonators have been reported. For example, thermally-dependent axial loads can be applied to a device to counteract the effects of the temperature coefficient of Young's modulus (TCE) [8]. However, this type of approach requires a significant amount of additional area to accommodate for the compensation structure. Another common technique is the use of composite resonators, where different materials with opposite values of TCE are stacked to cancel out or minimize the overall TCF [9]. Nevertheless, interfacial material losses tend to degrade the quality factor which compromises the device performance. Hence, in applications where minimal size and maximum performance are a necessity, electrical frequency tuning tends be a better option, as long as the device offers enough tuning range to enable compensation for both temperature and process variations.

As of today, very few 32 kHz MEMS resonators have been reported [10, 11]. These devices rely on electrostatic actuation and sensing; thus, they require a polarization voltage V_P to operate. To achieve low motional impedances, high aspect ratio capacitive gaps and/or large V_P values are necessary, both of which severely compromise power handling. Furthermore, electrostatic tuning is needed to compensate for TCF; hence, the polarization voltage of the device must be used to adjust the frequency utilizing a temperature compensation scheme. This in turn changes the motional impedance of the device, complicating the implementation of frequency-stable oscillators with a wide operating temperature range. In order to overcome these limitations, thin-film piezoelectric-on-SOI flexural resonators can be utilized to enable piezoelectric actuation as well as electrostatic frequency tuning [12]. Small form factor, highly-tunable low-frequency devices can be realized by carefully designing their geometry as well as the size and shape of their drive and sense electrodes. This paper reports on the optimization, implementation, and characterization of such structures utilizing a CMOS-compatible process flow that is also compatible with high frequency piezoelectric bulk acoustic resonators.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

II. TUNABLE PIEZOELECTRIC MEMS RESONATOR

A. Design and Operating Principle

Figure 1 shows a schematic diagram of the proposed 32 kHz piezoelectrically-actuated, capacitively-tunable resonator. This novel structure is composed of an external frame of four clamped-clamped beams anchored at the corners to the surrounding substrate. The beams are mechanically coupled to each other through a suspended rigid plate that provides sufficient structural support without compromising device compliance. More importantly, the plate contributes additional mass to achieve low-frequency operation and a large capacitive electrode area for increased electrostatic tuning.

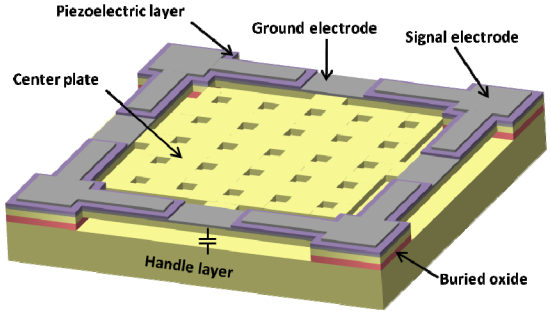


Figure 1. Small form factor piezoelectrically-transduced and electrostatically tunable 32 kHz resonator.

The resonator was implemented on an AlN-on-SOI substrate to allow piezoelectric transduction. A pair of drive and sense signal electrodes were carefully defined on top of each tether by etching the AlN and top molybdenum layers of the piezoelectric stack. The bottom molybdenum, which is in electrical contact with the single crystal silicon (SCS) device layer, serves as the ground electrode.

By applying an AC voltage between each pair of drive and ground electrodes, shear stress is generated over the beams by transverse piezoelectric coupling, which is determined by the d_{31} coefficient of AlN. This stress produces a bending moment that excites the structure into its first out-of-plane flexural mode (Fig. 2). The displacement, amplified by the quality factor of the structure, can be sensed by collecting the charge generated at the opposite pair of signal electrodes. Exciting a vertical mode rather than a lateral vibration is advantageous because the frequency value of a flexural structure is highly dependent on the dimensions along the displacement axis. The displacement magnitude of in-plane modes is determined by lithography, making the resonance frequency highly prone to process variations. In the case of out-of-plane modes, displacement occurs along the device thickness, which can be more precisely controlled through deposition steps.

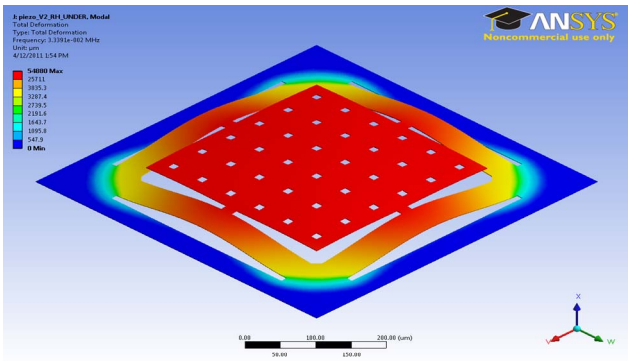


Figure 2. ANSYS® modal simulation results of fundamental out-of-plane flexural mode of piezoelectric 32 kHz resonator.

Although the structure does not require a polarization voltage to resonate, electrical access to the handle layer enables it to be used as a DC tuning electrode. Because the device layer is directly connected to ground, any potential applied to the handle layer will cause an electrostatic spring softening effect that can be utilized to tune the central operating frequency to compensate for temperature and process variations.

B. Device Optimization

Different design techniques were utilized to optimize the device motional impedance, frequency-tuning range, size and manufacturability. Single clamped-clamped beams can be used to implement a simpler version of the resonators proposed in this work; however, their length has to be greatly increased in order to achieve low-frequency operation, which significantly reduces the structural stiffness. This results in devices with large surface-area-to-volume ratios that are highly prone to stiction. To overcome this issue, the presented design takes advantage of mass loading, provided by a rigid plate rather than a decrease in stiffness, to achieve the target resonance frequency. This makes the structure more robust without compromising the overall area.

In piezoelectrically-transduced resonators, the motional impedance is inversely proportional to the square of the electromechanical coupling. To maximize this value and achieve low impedance, the signal electrodes should be carefully sized and shaped to guarantee that, at no point of the covered area, the induced stress experiences a change in sign; this avoids charge cancellation and maximizes coupling. It can be shown that for clamped-clamped beams operating in their fundamental mode, the electrodes should cover up to $1/4^{th}$ of the beam length in order to achieve maximum electromechanical transduction [13]. For the presented design, FEM simulations were used to sweep the electrode length and verify these results. As shown in Figure 3, the optimal point occurs when the electrode is 22.5% of the beam length rather than 25%. This is due to the undercut of the device layer, which changes the effective beam anchor location and hence its maximum and minimum stress points.

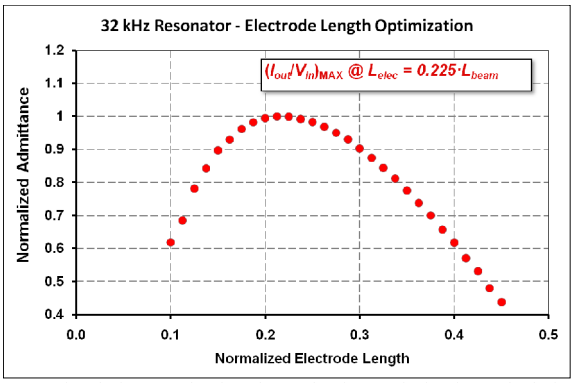
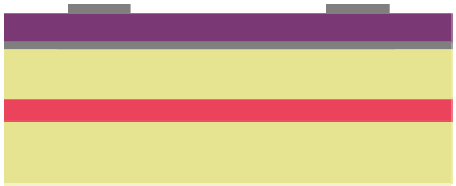


Figure 3. Simulation results for electrode size optimization. Ideal size is slightly smaller than theoretical value due to undercut of device layer.

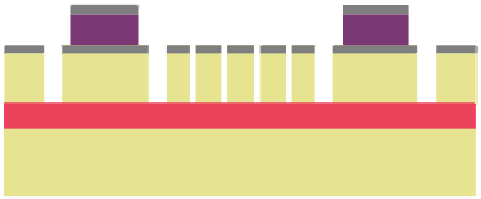
III. Fabrication Process

For compatibility purposes, these low-frequency structures were implemented with a similar process flow utilized for high-frequency thin-film piezo-on-substrate (TPoS) bulk-acoustic wave resonators [14]. SIMOX SOI wafers with device layer of 1.5 µm and buried oxide (BOX) of 1 µm, were used as the starting substrate. A stack of Mo/AlN/Mo was deposited on the silicon substrate with thicknesses of 0.1/0.5/0.1 µm, respectively. The top Mo layer was patterned to define the top signal electrodes (Fig. 4a), followed by the etching of the AlN layer to provide access to the bottom ground electrode. AlN is also removed from the center portion of the beams and the rigid plate to reduce Q damping (Fig. 4b). Lateral trenches and release holes were then patterned on the Si device layer to delimit the structure geometry (Fig. 4c). Access to the handle layer to define tuning electrodes was also attained in this step. Finally, the device was released in hydrofluoric acid (HF), leaving a capacitive gap between the structure and the handle layer (Fig. 4d).

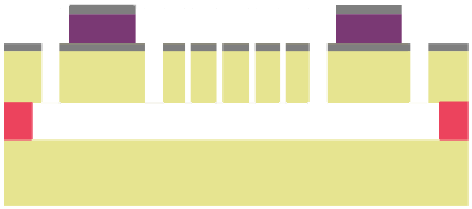


(a) Deposit Mo/AlN/Mo stack and pattern top electrode

(b) Etch AlN to grant access to ground electrode



(c) Pattern Si trenches and release holes to define geometry



(d) Device release by etching BOX layer in HF

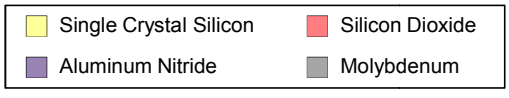


Figure 4. Fabrication process flow diagram of AlN-on-SOI resonator.

Figure 5 shows an SEM view of a fabricated 32 kHz resonators with a center plate area of only 250x250 µm^2. Structures with different plate areas where implemented to analyze the trade-offs between performance and area; however, the results presented in this work are only related to the aforementioned structure.

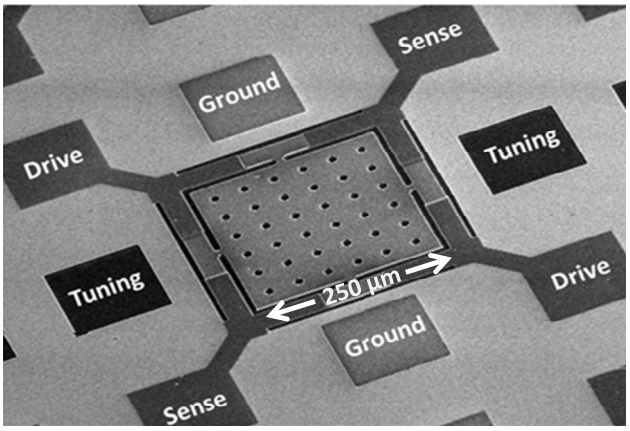


Figure 5. SEM view of 32 kHz tunable piezoelectric MEMS resonator.

IV. Measurement Results

The devices were characterized in an open-loop configuration. Figure 6 shows the frequency response of a resonator after three different tuning potentials were applied. It can be seen that both the motional impedance and the Q remain constant even when the DC voltage changes. This is a clear advantage over most capacitively-transduced resonators since it simplifies the design of the interface circuit required to build up a temperature stable oscillator.

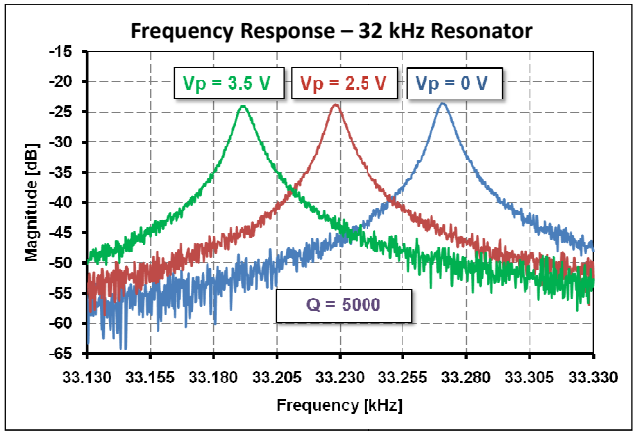


Figure 6. Resonator frequency response. Insertion loss and quality factor are independent of the applied tuning voltage.

To compensate for frequency changes with temperature, the device requires a wide frequency tuning range. Figure 7 shows that by applying voltages no larger than 4 V, this particular design can achieve more than 3100 ppm of tuning. The pull-in voltage of this structure is close to 6 V, meaning that up to 6500 ppm is theoretically achievable if needed. This extra margin can be use to counteract frequency deviations due to process variations.

A linear TCF of -27.8 ppm/ºC was measured for the same device at a constant DC potential. The tuning voltage was then adjusted at each measured temperature point to bring the frequency back to its original value. Full TCF compensation was achieved for temperatures ranging from -25 to 100 ºC. Figure 8 shows the temperature characteristic of the device before and after calibration.

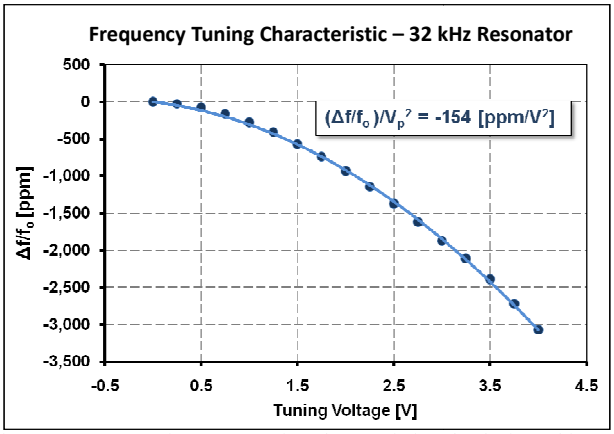


Figure 7. Resonator frequency tuning characteristic. The device achieves 3100 ppm of frequency tuning for voltages no larger than 4 V.

CONCLUSION

This work demonstrates the implementation of tunable piezoelectric MEMS resonators with form factors of at least one order of magnitude smaller than current quartz crystals. The devices operate in an out-of-plane flexural mode, allowing for accurate frequency control during fabrication. 3100 ppm of frequency tuning was achieved for full TCF compensation from -25 to 100 ºC utilizing no more than 4 V. The quality factor and motional impedance of the devices remain constant regardless of the tuning potential, facilitating the implementation of temperature stable oscillators for real-time clock applications.

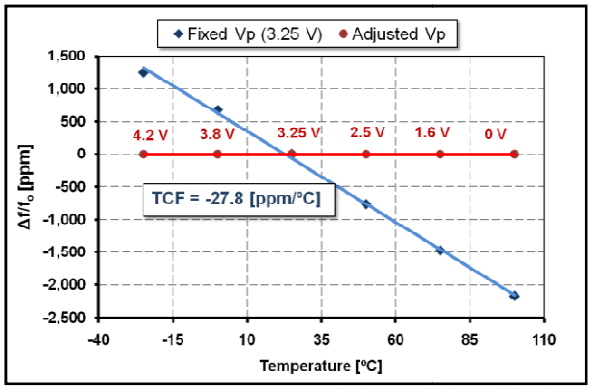


Figure 8. Resonator temperature response before and after calibration. Tuning range is sufficient for full TCF compensation from -25 to 100 ºC.

ACKNOWLEDGMENT

This work was supported by Integrated Device Technology (IDT), Inc. The authors would like to thank the staff at the Nanotechnology Research Center (NRC) at the Georgia Tech for fabrication assistance.

REFERENCES

[1] P. Rako, "Making Oscillator Selection Crystal Clear," *EDN,* pp. 28-37, Feb. 19, 2009.

[2] C. S. Lam, "A review of the recent development of MEMS and crystal oscillators and their impacts on the frequency control products industry," in *IEEE International Ultrasonics Symposium*, 2008, pp. 694-704.

[3] F. Ayazi, "MEMS for integrated timing and spectral processing," in *Proc. IEEE Custom Integrated Circuits Conference*, 2009, pp. 65-72.

[4] FC-125 - Low Profile SMD kHz Range Crystal Unit. EPSON Toyocom [Online]. Available: http://www.epsontoyocom.co.jp/

[5] A. K. Samarao, G. Casinovi, and F. Ayazi, "Passive TCF Compensation in High Q Silicon Micromechanical Resonators," in *IEEE International Conference on Micro Electro Mechanical Systems*, Hong Kong, 2010, pp. 116-119.

[6] R. Tabrizian, G. Casinovi, and F. Ayazi, "Temperature-Stable High-Q AlN-on-Silicon Resonators with Embedded Array of Oxide Pillars," in *Solid-State Sensors, Actuators, and Microsystems Workshop*, Hilton Head, 2010, pp. 100-101.

[7] K. Sundaresan, G. Ho, S. Pourkamali, and F. Ayazi, "Electronically Temperature Compensated Silicon Bulk Acoustic Resonator Reference Oscillators," *IEEE Journal of Solid-State Circuits,* vol. 42, pp. 1425-1434, 2007.

[8] H. Wan-Thai, J. R. Clark, and C. T. C. Nguyen, "Mechanically temperature-compensated flexural-mode micromechanical resonators," in *Tech. Dig. IEEE International Electron Devices Meeting*, 2000, pp. 399-402.

[9] R. Melamud, S. A. Chandorkar, K. Bongsang, L. Hyung Kyu, J. C. Salvia, G. Bahl, M. A. Hopcroft, and T. W. Kenny, "Temperature-Insensitive Composite Micromechanical Resonators," *IEEE Journal of Microelectromechanical Systems,* vol. 18, pp. 1409-1419, 2009.

[10] K. R. Cioffi and H. Wan-Thai, "32KHz MEMS-based oscillator for low-power applications," in *Proc. IEEE International Frequency Control Symposium and Exposition*, 2005, pp. 551-558.

[11] G. Ho and F. Ayazi, "Low Frequency Process-Variation-Insensitive Temperature-Stable Micromechanical Resonators," US Patent 7,859,365, Dec. 28, 2010.

[12] G. Piazza, R. Abdolvand, G. Ho, and F. Ayazi, "Piezoelectrically-Transduced, Capacitively-Tuned, High-Q Single-Crystal Silicon Micromechanical Resonators on SOI Wafers," *Sensors and Actuators A: Physical,* vol. 111, pp. 71-78, 2004.

[13] D. L. DeVoe, "Piezoelectric thin film micromechanical beam resonators," *Sensors and Actuators A: Physical,* vol. 88, pp. 263-272, 2001.

[14] W. Pan and F. Ayazi, "Thin-Film Piezoelectric-on-Substrate Resonators with Q Enhancement and TCF Reduction," in *IEEE International Conference on Micro Electro Mechanical Systems*, 2010, pp. 104-107.

Linear Acoustic Bandgap Arrays for Spurious Mode Suppression in Piezoelectric MEMS Resonators

Logan Sorenson, Jenna Fu, and Farrokh Ayazi
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia, USA
logan.sorenson@gatech.edu

***Abstract*— In this work, we present a technique that utilizes arrays of coupled-ring linear acoustic bandgap (LAB) structures, which exhibit characteristic frequency stop bands, to reduce spurious modes in the wide-band response of AlN-on-Si lateral mode resonators. Aggressive LAB tethering (one support per finger) of high-order modes yields an increase in atmospheric quality factor (*Q*) from 1860 to 2890 for the designed peak at 209 MHz, compared to resonators with traditional tethers. Spurious modes from 0 to 360 MHz are suppressed by at least 30 dB relative to the main peak. The LAB structures presented achieve this goal through a combination of complete acoustic bandgaps and deaf bands.**

I. Introduction

Today's mobile devices are becoming increasingly powerful as a result of the incorporation of multiple wireless frequency standards into a single handheld unit. Applications such as wireless internet, GSM/CDMA voice networks, GPS navigation, and BlueTooth/ZigBee interdevice networking are already included in many mobile phones. Future applications will be enabled by emerging frequency standards (e.g., IEEE 802.15 wireless personal area networks [1]). Mobile phones may soon feature personal health monitoring through body area networks (BANs) [2].

To support integration of these standards into consumer products, microelectromechanical systems (MEMS) resonators, which offer several advantages for RF applications including small size, high mechanical quality factor (*Q*), and low loss, can be configured in the local oscillators of a wireless transceiver. Furthermore, lithographically-defined frequencies of certain MEMS resonators enable multiple frequency standards to be implemented on the same substrate. A typical MEMS oscillator consists of a MEMS resonator connected in the feedback path of a transimpedance amplifier (TIA) (Fig. 1a).

While piezoelectric MEMS resonators can be easily interfaced with low-power electronics, the appearance of sufficiently strong low-*Q* spurious modes near the designed operation frequency can cause degraded oscillator performance. Oscillator phase and gain conditions may cause the spurious mode to become preferred over the desired mode during start-up, causing the oscillator to lock into the spurious mode. Additionally, feedthrough and coupling between the desired and the spurious modes can degrade phase noise

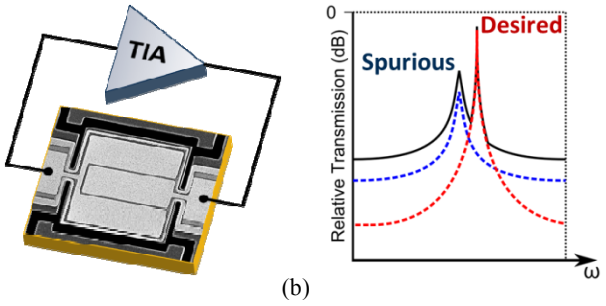


Figure 1. (a) Schematic of MEMS oscillator with resonator in feedback with CMOS TIA; (b) Transmission plot representation showing negative feedthrough and coupling effects of a spurious mode occurring close to the desired resonance peak.

performance and increase frequency instability, depending on the proximity of the spur (Fig. 1b) [3].

Therefore, it is essential to eliminate wide-band spurious modes in piezoelectric micromechanical resonators for use in oscillators [4]. The linear acoustic bandgap (LAB) structures presented in this work may be used to achieve this goal by creating support structures that utilize a combination of complete acoustic bandgaps [5] and deaf bands [6] to suppress the spurious modes without lowering the *Q* of the desired mode.

II. Linear Acoustic Bandgap Structures

LAB structures are one-dimensional phononic crystals (PCs) with small cross-sectional dimensions perpendicular to the axis of the 1D lattice. The coupled-ring LAB structure presented in this work is a linear chain of rings connected by narrow beams, repeated at a characteristic distance *a* known as the lattice constant (Fig. 2). This periodically-repeated section is termed the basis of the 1D PC. The ring inner and outer radii (r_i , r_o), coupling beam width *w*, and layer thicknesses d_i parameterize the basis dimensions.

The periodic nature of the 1D PC combined with large acoustic mismatch at the solid/air interface creates acoustic bandgaps, or frequencies of no allowed acoustic propagation (shaded blue in Fig. 3). In addition to complete acoustic bandgaps, "deaf bands" [6] can also be found in the coupled-ring LAB structure dispersion characteristic. These bands

978-1-61284-110-6/11/$26.00 ©2011 IEEE

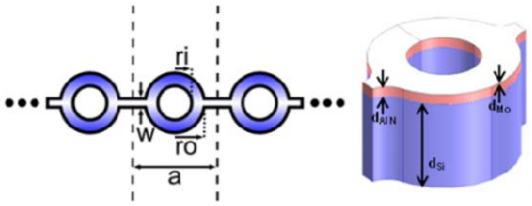


Figure 2. Schematic and 3D view of coupled-ring LAB structure with AlN-on-Si stack (a = 20 µm, r_i = 4 µm, r_o = 8 µm, w = 2 µm, d_{Si} = 10 µm, d_{AlN} = 1 µm, d_{Mo} = 100 nm).

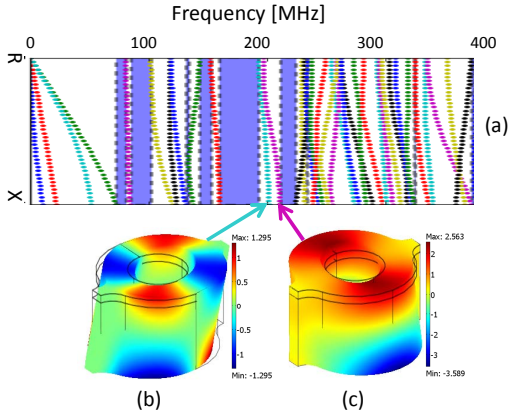


Figure 3. (a) Acoustic dispersion curve for AlN-on-Si coupled-ring linear acoustic bandgap (LAB) structure (bandgaps shaded blue); (b, c) mode shapes of indicated bands showing displacement polarization orthogonal to the axis of the LAB structure.

demonstrate minimal coupling to the LAB structure when waves from the acoustic source exhibit orthogonal polarization to the mode of the LAB structure at the boundary between the two components. The mode shapes of two such deaf bands are depicted in Fig. 3, where the acoustic source is a longitudinal acoustic wave propagating orthogonally to the axis of the coupled-ring LAB structure. For these modes, displacement corresponding to the polarization of the longitudinal wave is effectively zero on the exposed faces of the coupling beam, meaning these bands of the LAB structure are deaf to the longitudinal waves propagating in this direction. Thus, LAB structures and integrated devices must be carefully designed to maximize the attenuation effects provided by complete acoustic bandgaps as well as deaf bands.

III. Device Design and Fabrication

LAB structures can serve as Q-enhancing support elements for MEMS resonators by tailoring bandgaps and deaf bands to confine acoustic energy within the device at the desired operation frequency. This confinement of energy enables an aggressive tethering approach using LAB structures, where one pair of tethers is used for each finger of the interdigitated transducer, to simultaneously improve the support loss component of Q ($Q_{support}$) and provide significant attenuation of spurious modes which lie outside of the designed bandgaps.

A. $Q_{support}$ Enhancement via Aggressive LAB Tethering

The perfectly matched layer (PML) method was employed in COMSOL finite element software to study the effect of aggressive tethering on resonator support loss [7]. PMLs are

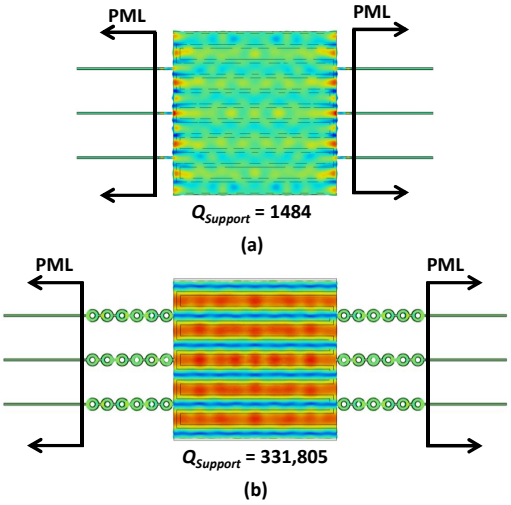


Figure 4. COMSOL PML simulation for a device with three support pairs; (a) simple tethers; (b) LAB tethers.

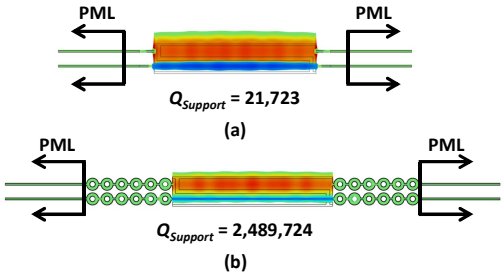


Figure 5. COMSOL PML simulation for an infinite order device; periodic boundary conditions are applied to the appropriate faces of the device to emulate infinite extent; (a) simple tethers; (b) LAB tethers.

additional artificial domains attached to the device which absorb incoming acoustic waves and emulate loss of acoustic energy into the substrate surrounding the device. In this way, PMLs give an estimation of the theoretical lower bound on $Q_{support}$.

Figure 4 shows the results of applying PMLs to an 11th order, 20 µm pitch, 220 µm wide AlN-on-Si resonator, comparing both traditional simple beam support tethers and six periods of coupled-ring LAB supports. For both designs, three pairs of support tethers are used to minimize the number of parallel acoustic leakage paths while maintaining sufficient structural support.

To avoid acoustic mismatch, the PML domains were created by extending the cross-section of the support tethers by a certain distance known as the PML length. In addition to this length, there are several parameters in the PML implementation of COMSOL which must be swept to obtain accurate values for the support loss. Additionally, a sufficiently fine mesh is needed to ensure convergence of the support loss value. Sweeping these parameters enables exploration of the PML design space to determine the lower bound of $Q_{support}$. Because reflections and confinement of acoustic energy due to material mismatch are neglected in simulation, the PML result underestimates the value of $Q_{support}$ (i.e., the values obtained from the PML method are lower than those obtained in reality).

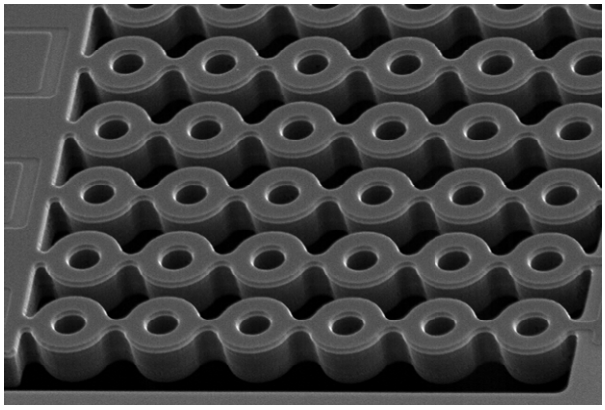


Figure 6. SEM image of fabricated LAB ring tether array.

With these considerations in mind, the $Q_{support}$ for the 11th-order resonator with three pairs of simple tethers is estimated to be 1484 (Fig. 4a), while a $Q_{support}$ of 331,805 is obtained for the same device with LAB tethers, showing over 200-fold improvement of support loss (Fig. 4b). Note that the mode shape in Fig. 4a shows deterioration, and a wave-like pattern develops along the supported edge of the resonator. Comparison of these results with other sets of PML parameters shows that as $Q_{support}$ is increased, the mode shape of the design in Fig. 4a approaches that of Fig. 4b, where the rippling along the supported edge is less apparent, suggesting that the support loss can have localized impact on the resonator mode shape.

Figure 5 studies a device that uses one pair of support tethers per finger of the interdigitated transducer structure. In addition to the PML method, periodic boundary conditions are employed to reduce the number of fingers to two. This effectively models an infinite order resonator and greatly reduces required computation time to calculate $Q_{support}$ (each support contributes a large number of elements). The high number of support tethers confines the mode more closely to the resonator body and nearly eliminates the waves along the supported edge. The confinement also reduces the number of possible resonance modes in the vicinity of the desired mode, which translates into higher $Q_{support}$, even in the traditional tether case. Over 100-fold improvement is obtained in simulation by employing aggressive LAB tethers, boosting the estimated $Q_{support}$ from 21,723 to 2,489,724 (Fig. 5).

B. Fabrication Process Flow

Devices were fabricated on a silicon-on-insulator (SOI) substrate with 10-µm device layer thickness (d_{Si}). The piezoelectric stack, which consists of 100-nm Mo electrodes and 1-µm thick AlN, was deposited directly on the silicon device layer and subsequently patterned to define the top Mo and provide access to the bottom Mo for grounding. The resonator shape, including LAB ring tethers, was defined by RIE etching the piezoelectric stack and Bosch etching the device layer to maintain straight sidewalls. The devices were released by etching the handle silicon layer and removing the buried oxide (BOX) layer. A close-up SEM image of an aggressive LAB tether array is shown in Fig. 6. SEM images of fabricated devices with simple tethers and LAB tether arrays are shown in Fig. 7. The LAB-supported resonators have six periods per support tether. Because the LAB support array provides Q enhancement of insertion loss (IL), the width of the resonator body was decreased.

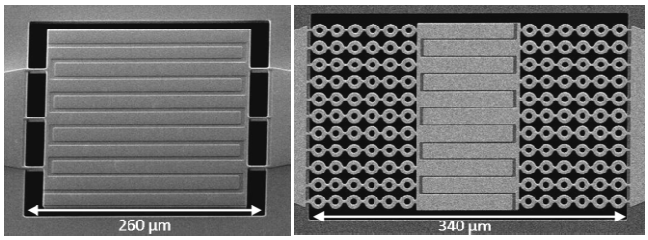


Figure 7. SEM images of 11th-order, 20 µm pitch AlN-on-Si resonators fabricated on 10 µm SOI wafers: (left) 3 pairs of traditional beam supports (20 µm x 2 µm); (right) 11 pairs of 6 period coupled-ring LAB supports.

IV. Experimental Results

To quantify the impact of the LAB support array, the devices of Fig. 7 were measured on a Cascade Microtech Summit 12000 M probe station at atmospheric pressure and room temperature. The data were taken with an Agilent N5241a PNA-X network analyzer in two-port configuration.

A 1 GHz span was measured to observe the wide-spectrum response of both devices (Fig. 8). All spurious modes below 360 MHz are suppressed by at least 30 dB relative to the desired mode, which lies near 209.5 MHz. A few spurs are amplified in the range 150-250 MHz as a result of acoustic confinement introduced by the LAB support array; however, they are not large enough to impact oscillator operation at the desired peak.

In addition to spurious mode suppression in the vicinity of the desired peak, at least 25 dB of suppression is obtained at higher frequencies between 360 MHz and 680 MHz. The cluster of spurious modes near 370 MHz is well-suppressed in IL compared to the reference device. There is also an acoustically quiet section of the frequency spectrum from 380 MHz to 480 MHz in the LAB-supported resonator, when compared to several spurious modes in the reference response. A spurious mode is found at 680 MHz in the LAB-supported case. It is not clear if a relationship exists between this mode and the reference spurious modes in that vicinity; however, it may indicate a bandgap or deaf band introduced by LAB structures at that frequency.

Figure 9 shows a 50 MHz span centered about the desired mode. The LAB array has completely removed a spur on the high-frequency side of the reference device peak. Underlying modes are boosted by the LAB tethers but remain near the noise floor and will have minimal effect on oscillator performance. Finally, Figure 10 gives a narrow-span (2 MHz) measurement of the main peaks for both reference and LAB array devices. The LAB array is found to improve the device Q by almost 60% from 1860 to 2890.

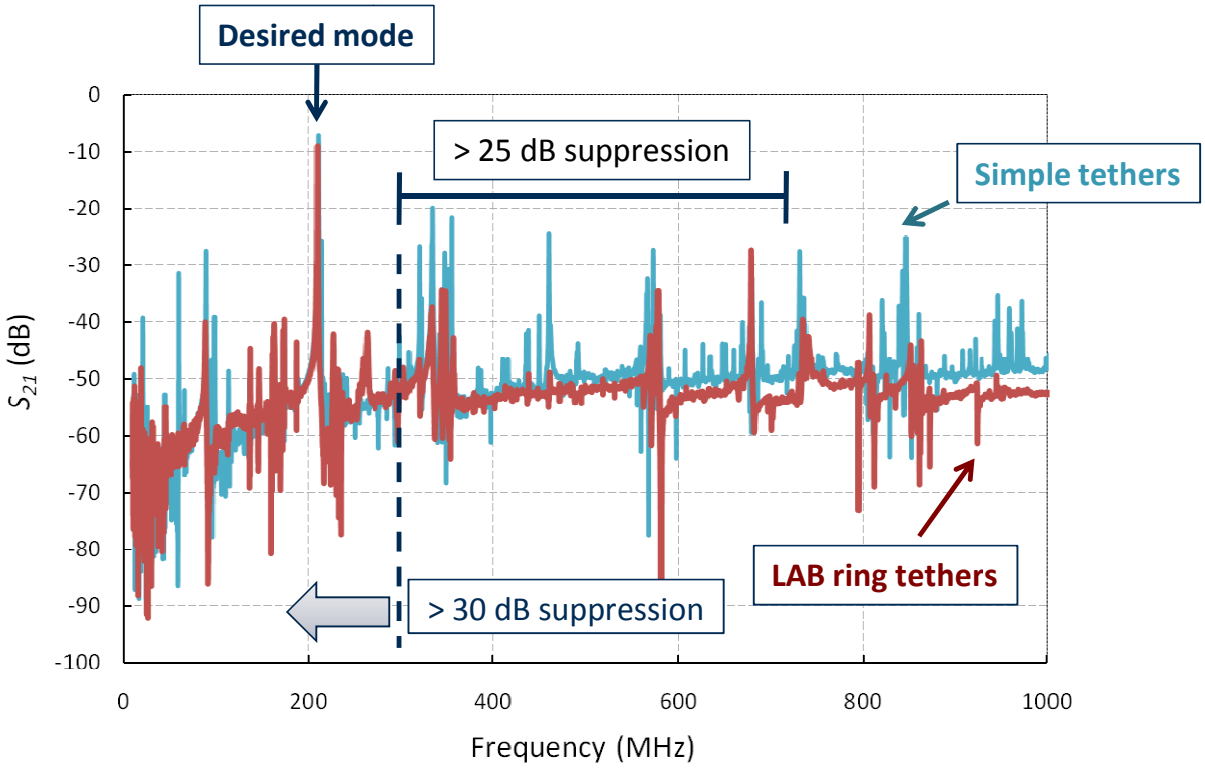


Figure 8. Measured S_{21} response showing wide-band spurious suppression. Below 360 MHz, spurious modes are suppressed at least 30 dB relative to the desired mode. Between 360 MHz and 680 MHz, spurious modes are suppressed at least 25 dB.

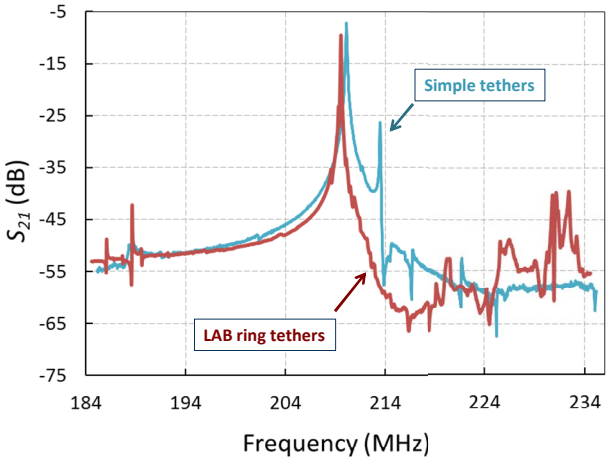


Figure 9. Medium-span (50 MHz) response for both devices at 210 MHz.

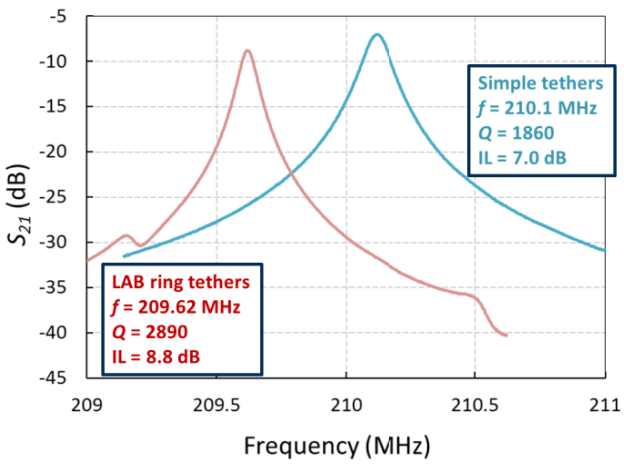


Figure 10. Narrow-span (2 MHz) response for both devices at 210 MHz.

V. Discussion

Comparison of the measured Q value with the $Q_{Support}$ value predicted by COMSOL finite element PML simulations for the simple tether case (Fig. 10 vs. Fig. 4a) yields the conclusion that support loss simulations are accurate to about 20%. However, this particular case ignores reflections back into the resonator body from material and geometry mismatches in the support structure. In that context, COMSOL PML simulations predict the lower bound of the Q value, which is accurate for support loss limited resonators.

However, the predicted $Q_{Support}$ for the LAB-enhanced resonator was in the millions, whereas the measured value was still limited to a few thousand (Fig. 10 vs. Fig. 5b). In high-order piezoelectric MEMS resonators, a large component of the dissipation can be attributed to the piezoelectric transduction stack [8]. LAB array supports nearly eliminate support loss, which causes other factors to become dominant in the expression of the overall Q:

$$Q = \left[\frac{1}{Q_{Support}} + \frac{1}{Q_{Other}}\right]^{-1} \quad (1)$$

Finally, the enhanced Q of the LAB supported device also improves the device IL, enabling a reduced transduction area for similar levels of IL. Thus, LAB structures can be incorporated into MEMS resonators without significantly increasing device area.

VI. Conclusion

In this work, we have presented aggressive tethering of MEMS resonators with LAB arrays to control and suppress spurious modes which negatively impact MEMS oscillator performance. We have also used the PML method to predict the lower bound of support loss contribution to overall resonator Q, and demonstrated that this approach agrees well with measured experimental values when the resonator is support-loss limited.

Acknowledgment

This work was supported by Integrated Device Technology, Inc. The authors would like to thank the OEM Group for AlN deposition, the staff at the Georgia Tech NRC for fabrication support, and Dr. Saeed Mohammadi for helpful discussions.

References

[1] IEEE Std 802.15.1™-2005, "IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements; Part 15.1: Wireless medium access control (MAC) and physical layer (PHY) specifications for wireless personal area networks (WPANs)," ed, 2005.

[2] T.G. Zimmerman, "Personal area networks: near-field intrabody communication," *IBM Systems Journal,* vol. 35, 1996, pp. 609-617.

[3] M.E. Tobar, "Effects of spurious modes in resonant cavities," *Journal of Physics D: Applied Physics,* vol. 26, 1993, p. 2022.

[4] R. Abdolvand, H. Lavasani, G. Ho, and F. Ayazi, "Thin-film piezoelectric-on-silicon resonators for high-frequency reference oscillator applications," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control,* vol. 55, 2008, pp. 2596-2606.

[5] S. Mohammadi, AA Eftekhar, A. Khelif, H. Moubchir, R. Westafer, WD Hunt, and A. Adibi, "Complete phononic bandgaps and bandgap maps in two-dimensional silicon phononic crystal plates," *Electronics Letters,* vol. 43, 2007, pp. 898-899.

[6] Fu-Li Hsiao, Abdelkrim Khelif, Hanane Moubchir, Abdelkrim Choujaa, Chii-Chang Chen, and Vincent Laude, "Complete band gaps and deaf bands of triangular and honeycomb water-steel phononic crystals," *Journal of Applied Physics,* vol. 101, 2007, pp. 044903-5.

[7] W. Chu and Q. Liu, "Perfectly matched layers for elastodynamics: A new absorbing boundary condition," *J. Comp. Acoustics,* vol. 4, 1996, pp. 341–359.

[8] Farrokh Ayazi, Logan Sorenson, and Roozbeh Tabrizian, "Energy Dissipation in Micromechanical Resonators," in *Proc. SPIE*, 2011, pp. 1-13.

Reduced Kalman Filters for Clock Ensembles

Charles A. Greenhall
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, CA 91109 USA
E-mail: cgreenhall@jpl.nasa.gov

***Abstract*— This paper summarizes the author's work on timescales based on Kalman filters that act upon the clock comparisons. The natural Kalman timescale algorithm tends to optimize long-term timescale stability at the expense of short-term stability. By subjecting each post-measurement error covariance matrix to a non-transparent reduction operation, one obtains corrected clocks with improved short-term stability and little sacrifice of long-term stability.**

I. Free-Running Timescales

This paper is a summary of work on the formation of timescales, or composite clocks, that the author has carried out over the last ten years, with the goal of surpassing the performance of a previous timescale algorithm [1] that uses a Kalman filter to estimate what the clocks are actually doing.

Let us describe briefly the notion of a timescale derived from an "ensemble" of clocks. One has a set of n free-running clocks with "phase" (i.e, time) residuals $x_1^1(t), \ldots, x_1^n(t)$ at the nominal time t. One cannot read the values $x_1^i(t)$ directly (to the desired precision) but can only measure the phase differences $x_1^{ij}(t) := x_1^i(t) - x_1^j(t)$ at a set of measurement times t. A timescale algorithm, which uses these data and a mathematical model for the clocks, calculates numerical phase estimates $\hat{x}_1^i(t)$ to be subtracted from the clock phases, whether logically or physically, to give corrected clock phases

$$\tilde{x}_1^i(t) = x_1^i(t) - \hat{x}_1^i(t). \tag{1}$$

One hopes to produce corrected clocks with better stability[1] than the physical clocks over a range of averaging times. Since a motion common to all the clock phases is transparent to their differences, one cannot expect the corrected clocks to stay close to a physical clock. Improved timekeeping accuracy is not to be expected from combining clock readings in this way.

Section II describes the most common clock model. Section III sets up an ensemble of these clocks and describes the natural Kalman filter algorithm for calculating the phase estimates. The existing results motivate the author's work to reduce the short-term noise of the corrected clocks. Sections IV and V describe the author's improved Kalman filter algorithms for the case of noiseless measurements, as for a local ensemble. Section VI describes Davis's improvements. Section VII extends the method to remote sets of clocks with significant time-transfer noise. Section VIII gives a recent result on the reduced Kalman covariance matrices of a clock ensemble.

[1]Stability of a phase $x_1(t)$ for an averaging time τ is defined by the Allan deviation $\sigma_A(\tau)$ or Hadamard deviation $\sigma_H(\tau)$, where $\sigma_A^2(\tau) = (2\tau^2)^{-1} \mathrm{E}\left[\Delta_\tau^2 x_1(t)\right]^2$, $\sigma_H^2(\tau) = (6\tau^2)^{-1} \mathrm{E}\left[\Delta_\tau^3 x_1(t)\right]^2$.

II. A Clock Model

Much work in this area has been based on a clock model [2][3] with three independent process-noise components, called white FM (WFM), random-walk FM (RWFM), and random-run FM (RRFM), with phase spectra $S(f)$ proportional to f^{-2}, f^{-4}, and f^{-6}, and Hadamard deviations $\sigma_H(\tau)$ proportional to $\tau^{-1/2}$, $\tau^{1/2}$, and $\tau^{3/2}$, respectively. Missing are the odd-powered models, in particular flicker FM (FFM), for which $S(f) \propto f^{-3}$, $\sigma_H(\tau) = \text{const}$.

The state of one clock at time t is represented as a column vector, $x(t) = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) \end{bmatrix}^{\mathrm{T}}$, whose components are called phase, frequency state, and drift state. The state evolves according to the stochastic differential equations

$$\frac{dx_1}{dt} = x_2 + w_1, \quad \frac{dx_2}{dt} = x_3 + w_2, \quad \frac{dx_3}{dt} = w_3 \tag{2}$$

where $w_i(t)$ are independent mean-zero white noises with known two-sided spectral densities q_i. One may set $q_3 = 0$, for example, to simulate a constant drift $x_3(t) = c$, in which case $x_1(t)$ contains a quadratic component $\frac{1}{2}ct^2$.

III. Natural Kalman Timescale

The state of an ensemble of n clocks is a $3n$-vector $X(t)$, a stack of n clock vectors $x^i(t)$. The measurements performed on the state consist of noisy phase comparisons, $x_1^{ij}(t) + v_{ij}(t)$, at a sequence of times $t = t_k$, where the measurement noise variances, $\mathrm{E}\, v_{ij}^2(t) = r_{ij}$, are known. Even in the absence of noise, the measurements do not determine the clock states, since all clock phases could conceivably move in the same way together.

To set up a Kalman filter to estimate the state from the measurements, the differential equations (2) for each clock are integrated between two measurement times, call them $t-\delta$ and t, to give a stochastic difference equation. For one clock this takes the form

$$x(t) = \phi(\delta)\, x(t-\delta) + w(t,\delta) \tag{3}$$

where

$$\phi(\delta) = \begin{bmatrix} 1 & \delta & \delta^2/2 \\ 0 & 1 & \delta \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

978-1-61284-110-6/11/$26.00 ©2011 IEEE

and the process noise $w(t,\delta)$ is a mean-zero 3-vector with covariance matrix

$$\mathrm{E}\, ww^{\mathrm{T}} = q(\delta) = \begin{bmatrix} q_1\delta + q_2\frac{\delta^3}{3} + q_3\frac{\delta^5}{20} & * & * \\ q_2\frac{\delta^2}{2} + q_3\frac{\delta^4}{8} & q_2\delta + q_3\frac{\delta^3}{3} & * \\ q_3\frac{\delta^3}{6} & q_3\frac{\delta^2}{2} & q_3\delta \end{bmatrix} \tag{5}$$

where the $*$ entries are determined by symmetry. The difference equation for the whole ensemble is written

$$X(t) = \Phi(\delta) X(t-\delta) + W(t,\delta)$$

where Φ has n copies of ϕ down the diagonal and W is a stack of n uncorrelated noises w^i; consequently $\mathrm{E}\, WW^{\mathrm{T}} = Q(\delta)$, which has n 3×3 matrices $q^i(\delta)$ down the diagonal.

The phase comparison measurements at time t take the form $y(t) = H(t) X(t) + v(t)$, where H is a matrix of ± 1's and 0's such that each row takes the difference of two phase components of X. The mean-zero measurement noise v satisfies $\mathrm{E}\, vv^{\mathrm{T}} = R(t)$, a known matrix, usually diagonal. For a local ensemble one sets $R = 0$.

Let $\hat{X}(t)$ be an unbiased estimate of $X(t)$; its quality is specified by an error covariance matrix

$$C(t) = \mathrm{E}\left[X(t) - \hat{X}(t)\right]\left[X(t) - \hat{X}(t)\right]^{\mathrm{T}}.$$

Given $\hat{X}(t-\delta)$, $C(t-\delta)$, and $y(t)$, the Kalman filter equations propagate $\hat{X}(t)$ and $C(t)$ as follows. For clarity, some dependencies on δ and t are suppressed.

$$\begin{aligned} \hat{X}^-(t) &= \Phi\hat{X}(t-\delta) \quad \text{(prediction of } X(t)\text{)} \\ C^- &= \Phi C(t-\delta)\Phi^{\mathrm{T}} + Q \quad \text{(error cov. of } \hat{X}^-(t)\text{)} \\ K &= C^- H^{\mathrm{T}}\left(HC^- H^{\mathrm{T}} + R\right)^{-1} \quad \text{(Kalman gain)} \\ \hat{X}(t) &= \hat{X}^-(t) + K\left[y(t) - H\hat{X}^-(t)\right] \\ C(t) &= (I - KH)C^- \end{aligned}$$

Evidently $\hat{X}(t)$ is a linear function of $\hat{X}(t-\delta)$ and $y(t)$. In fact, $\hat{X}(t)$ is unbiased for $X(t)$ and has the minimal error covariance matrix $C(t)$ of any unbiased linear function in the sense that $C_1 - C(t)$ is non-negative definite if C_1 is the error covariance matrix of any other unbiased linear function. Initialization of this scheme is not treated here.

The natural Kalman corrected clock states are defined by $\tilde{X}(t) = X(t) - \hat{X}(t)$. The corrected phase $\tilde{x}_1^i(t)$ represents the timescale at clock i. The corrected phases need not coincide for remote clocks with noisy phase comparisons, but they do coincide for a set of local clocks being compared with one another with negligible measurement noises ($R = 0$). In this case the common corrected phase constitutes the unique timescale, which may be called $\tilde{x}_1(t)$.

A natural Kalman timescale was implemented at NIST during the 1980's as the local timescale TA(NIST). The GPS Kalman filter [4] also incorporates a natural Kalman timescale. When applied to a simulated ensemble of dissimilar clocks, the NIST timescale was observed to be quiet in long term but noisier in short term than some of the clocks in the ensemble. Weiss and Weissert [5] suggested that the Kalman filter is optimizing accuracy at the expense of stability. This author confirmed their observations with his own simulations and set out to find an improved method of using a Kalman filter to make a timescale that is more stable than any clock in its ensemble over a wide range of averaging times. Stein [6][7] had already accomplished this goal with the KAS-2 algorithm and had described some of its principles, but a full description remains proprietary.

In searching for improved Kalman timescales, the author has held to one fixed principle: the natural Kalman filter does a good job of estimating the underlying frequency and drift states, which, in the author's opinion, reflect the physics of the frequency standards driving the clocks. The author's filter modifications preserve the frequency and drift state estimates but modify the phase estimates to improve the stability of the corrected phases.

IV. Kalman Plus Weights

The author's first successful timescale algorithm [8][9] applies to a local ensemble with negligible measurement noise. The natural Kalman phase estimates are not used at all, but the Kalman frequency and drift estimates, $\hat{x}_2^i$ and $\hat{x}_3^i$, are used in a version of the "basic timescale" iteration. The propagation of the timescale, call it x_1^e, from $t-\delta$ to t can be expressed by

$$\begin{aligned} x_1^e(t) - x_1^e(t-\delta) &= \sum_{i=1}^{n} \lambda_i(t) \\ \cdot \left[x_1^i(t) - x_1^i(t-\delta) - \delta\hat{x}_2^i(t-\delta) - \tfrac{1}{2}\delta^2\hat{x}_3^i(t-\delta)\right], \end{aligned} \tag{6}$$

which says that the increment of x_1^e is a weighted average of the true phases $x_1^i(t)$ minus their predicted values. The weights λ_i are to be chosen subject to $\sum\lambda_i = 1$. The true phases are unknown, of course. To turn this into something that depends only on known quantities, subtract $x_1^j(t) - x_1^j(t-\delta)$ from both sides to give

$$\begin{aligned} x_1^{ej}(t) - x_1^{ej}(t-\delta) &= \sum_{i=1}^{n} \lambda_i(t) \\ \cdot \left[x_1^{ij}(t) - x_1^{ij}(t-\delta) - \delta\hat{x}_2^i(t-\delta) - \tfrac{1}{2}\delta^2\hat{x}_3^i(t-\delta)\right], \end{aligned} \tag{7}$$

which is now a computable iteration for $x_1^{ej}(t) := x_1^e(t) - x_1^j(t)$, the offset of the timescale from physical clock j.

To set the weights λ_i one tries to minimize the variance of $x_1^e(t) - x_1^e(t-\delta)$ in (7). From (3) one can argue for the approximation that $1/\lambda_i$ should be made proportional to the matrix element $q_{11}^i(\delta)$ of clock i (see (5)). As a further approximation, one can make the weights inversely proportional to the white FM noise levels q_1^i of the clocks.

Figure 1 [9] shows the performance of the natural Kalman and KPW timescales on a simulated ensemble of 8 clocks. The odd-numbered clocks have a combination of the three noise types WFM, RWFM, and RRFM. The KPW scale is much quieter than the natural Kalman scale at short averaging times; the natural Kalman scale is slightly quieter around $\tau = 10^6$ s.

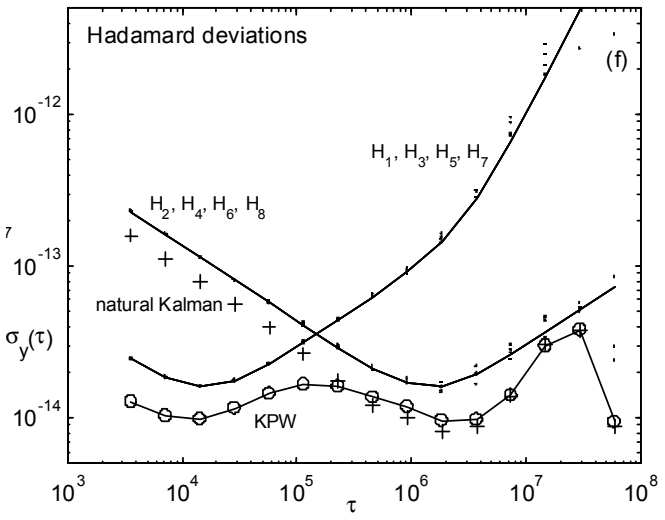


Fig. 1. Hadamard deviation of 8 simulated clocks and two timescales.

V. Reduced Kalman Timescale

For the KPW scale the Kalman phase estimates are irrelevant, but those state components are still carried along by the Kalman filter. Because only the phase differences are measured, the phase rows and columns of the covariance matrix $C(t)$ grow fast. After a processing a set of noiseless measurements, C is singular, however; in fact, the phase rows and columns are all the same vector. The author proved that setting all these matrix elements to zero after each measurement leaves unchanged the future Kalman frequency and drift estimates and their covariances, thus giving a more stable calculation. The author was fortunate to discover, just before giving the paper [9] at BIPM, that this crude operation of covariance reduction, called the "guillotine", also causes the ignored phase estimates $\hat{x}_1^i(t)$, when used as in (1), to yield automatically a different "reduced" timescale $\tilde{x}_1(t)$ that equals a KPW timescale $x_1^e(t)$ with the exact weights λ_i that minimize the mean squared value of the timescale increment $x_1^e(t) - x_1^e(t-\delta)$ [10]. These implicit optimal weights, which can easily be obtained from the Kalman gain matrix, are usually not far from the explicit weights chosen previously for KPW, and the two timescales differ little in practice. The natural Kalman timescale is also a KPW scale with implicit weights [11], but these can differ greatly from the optimal weights.

A Kalman filter that modifies its future estimates by using a reduced covariance matrix is called a reduced Kalman filter. The covariance matrix becomes an operational tool instead of a measure of the actual uncertainty of the phase estimates.

VI. Work with J. Davis

During 2003–5 John Davis of NPL and the author worked together on the reduced Kalman timescale of a local ensemble. Two of his improvements of the art [12] are summarized here.

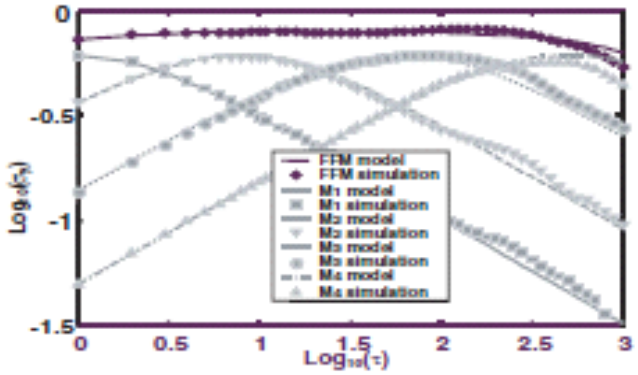


Fig. 2. Approximating flicker FM noise by a sum of independent Markov frequency processes.

First, an approximation of FFM noise can be added to the 3-state clock model by including a sum of independent Markov frequency processes (white noise through a single-pole lowpass filter). One such process is described by

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} + Rx_2 = w$$

where x_1 is phase, x_2 is frequency, and w is a white noise with spectral density S_w. Integrating this from $t-\delta$ to t gives the stochastic difference equation

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{R}\left(1-e^{-R\delta}\right) \\ 0 & e^{-R\delta} \end{bmatrix} \begin{bmatrix} x_1(t-\delta) \\ x_2(t-\delta) \end{bmatrix} + \begin{bmatrix} w_1(t,\delta) \\ w_2(t,\delta) \end{bmatrix}$$

where the covariance matrix of the w_i depends on S_w, R, and δ. One can include several such components in each clock with values of S_w and R chosen so that their sum approximates FFM noise over a range of time scales, as shown in figure 2, which plots the Allan deviation of the components and their sum. One can vary the parameters of the Markov components to produce a more complex frequency stability profile between regions of short and long averaging times.

Second, Davis devised a numerical calculation of the theoretical stability of the reduced Kalman timescale. This is done by calculating the true covariance matrix in a separate iteration that runs alongside the iteration for the operational reduced covariance matrix. One can also calculate the theoretical stability of each clock phase minus the timescale, and the results can be compared with the measured stability of the estimated phases $\hat{x}_1^i(t)$ calculated by the operational filter.

VII. Time-Transfer Noise and IEM

When noise is admitted to the clock phase comparisons, the corrected clocks need not coincide, but the ith corrected phase $\bar{x}_1^i$ can be said to represent the timescale at clock i. In his implementation of the natural Kalman timescale, Brown [4] devised a so-called implicit ensemble mean (IEM), a 3-vector $\bar{x}^B = \begin{bmatrix} \bar{x}_1^B & \bar{x}_2^B & \bar{x}_3^B \end{bmatrix}^{\mathrm{T}}$. This is a weighted average $W_{IE}\tilde{X}$ of the n corrected clock states $\tilde{x}^i$, using a $3\times 3n$ weight matrix W_{IE} that is calculated by solving a least-squares problem

whose noise covariance matrix is just the Kalman filter matrix C. The corrected states tend to cluster about the IEM, and the covariance matrix C_B of the n "representation errors" $\tilde{x}^i - \bar{x}^B$ is Brown's reduced form of C. The representation errors tend to be smaller than the corrected clock states, and C_B grows more slowly than C does. The use of C_B in the Kalman filter is transparent to the future state estimates, and the $\tilde{x}^i$ still represent the natural Kalman timescale. For the case of noiseless measurements, when all the corrected phases coincide, Galleani and Tavella [13] showed that the IEM phase component $\bar{x}_1^B$ does not reduce to the common timescale phase $\tilde{x}_1$. This happens because $\bar{x}_1^B$ depends on the corrected frequency and drift states [15].

By modifying Brown's ideas, the author was able to devise an extension of the reduced Kalman timescale that works well in the presence of measurement noise [14]. The IEM is now a one-dimensional phase $\bar{x}_1^G$, a weighted average $w_{IE}\tilde{X}_1$ of the n corrected clock phases $\tilde{x}_1^i$, where w_{IE} is a row of weights calculated from a 1-parameter least-squares problem with noise covariance matrix C_{11}, the phase-phase submatrix of C. After processing the measurements, one replaces the error covariance matrix C by the reduced version $C_G := SCS^{\mathrm{T}}$, where

$$S = \begin{bmatrix} I_n - 1_n w_{IE} & 0 \\ 0 & I_{2n} \end{bmatrix}$$

and 1_n is a column vector of n 1's. (The state vector has been rearranged so that all n phases come first.) It can be shown that C_G is the true covariance matrix of the corrected states with each $\tilde{x}_1^i$ replaced by its representation error $\tilde{x}_1^i - \bar{x}_1^G$.

The new covariance reduction is not transparent to the future phase estimates. When it is carried out after each set of measurements, the mean squared values of the following increments are minimized:

- $\tilde{x}_1^i(t) - \bar{x}_1^G(t-\delta)$, $i = 1$ to n, the corrected phases minus the previous IEM;
- $\bar{x}_1^G(t) - \bar{x}_1^G(t-\delta)$, the IEM increment.

In the absence of measurement noise, the covariance reduction reduces to the simple guillotine reduction of section V, the IEM $\bar{x}_1^G$ and all $\tilde{x}_1^i$ reduce to $\tilde{x}_1$, and the optimized increments reduce to $\tilde{x}_1(t) - \tilde{x}_1(t-\delta)$.

Figure 3 shows how this works for a simulated two-clock ensemble with measurement noise. For the natural Kalman scale, the corrected phases and Brown's IEM phase component all have the stability of the green clock, which is noisy in short term and quiet in long term. For the reduced Kalman scale, the IEM and corrected blue clock have essentially the same stabilities, while the corrected green clock overcomes the noise and joins the IEM for $\tau \geq 8$.

VIII. A Recent Result on Covariance Reduction

In a further comparison of the natural and reduced Kalman timescales [16], the author's colleagues M. Süß and D. Matsakis discovered that the covariance reductions of Brown (B) and the author (G) can be concatenated in either order to give a further reduced covariance matrix. Afterwards the author was

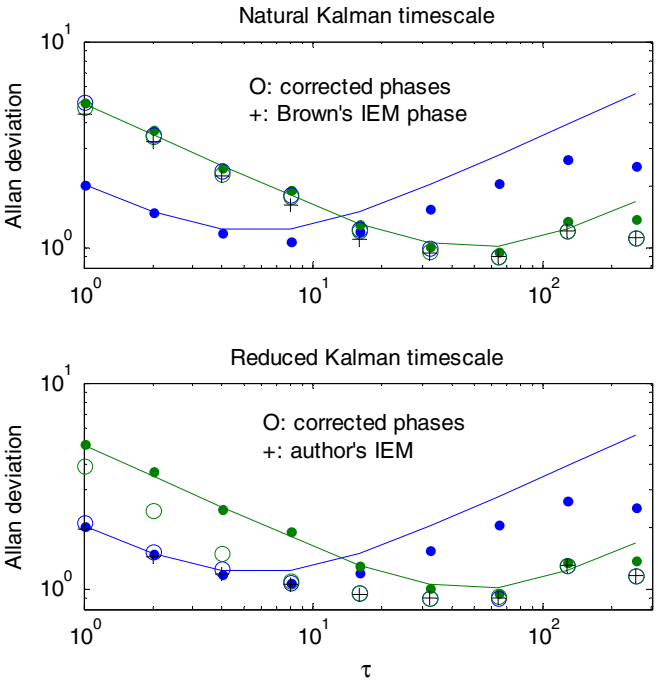


Fig. 3. A two-clock ensemble with measurement noise, showing the stability of the clocks, corrected clocks, Brown's IEM (phase component) and the author's IEM.

able to prove this result mathematically. The reductions are illustrated in fig. 4, in which the matrices are rearranged so that all the phase components come first. For a starting covariance matrix C, the final matrix, $B(G(C)) = G(B(C))$ has the phase border of $G(C)$ and the frequency-drift submatrix of $B(C)$. The final matrix can be used in place of $G(C)$ for operating the reduced Kalman timescale.

IX. Concluding Remarks

By running simulated clock models, the author has found that that the natural method of constructing a timescale from a Kalman filter can be improved by a suitable reduction of the Kalman error covariance matrix, different from the transparent reduction of Brown. The natural Kalman timescale, when applied to dissimilar clocks, tends to minimize the long-term stability of the corrected clocks without regard to their short-term stability. The reduced Kalman timescale, which is designed to minimize the mean squared increments of the corrected clocks instead of their values at one time, seems to produce good results for stability over all available averaging times. So far, though, this is only an empirical observation from experiments in simulation playpens. One needs to gain experience with these algorithms on real clock ensembles while solving practical problems such as noise parameter estimation, clock dropouts, and changes of clock behavior.

The author thanks P. Tavella for inviting him to give this paper. This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract

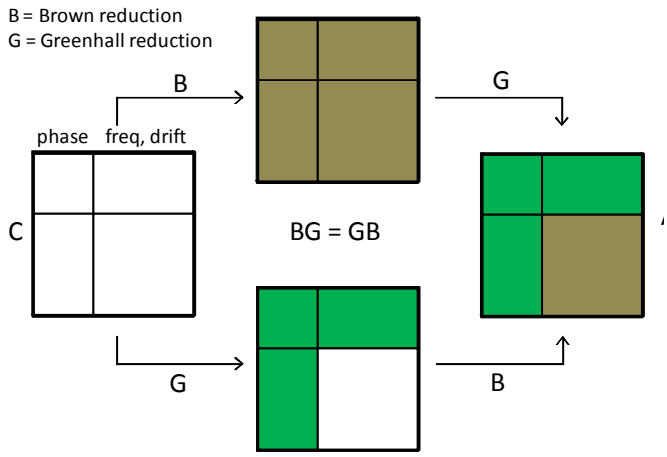


Fig. 4. Two commuting reduction operations, B and G, applied to a covariance matrix C.

with the National Aeronautics and Space Administration.

REFERENCES

[1] R. H. Jones and P. V. Tryon, "Continuous time series models for unequally spaced data applied to modeling atomic clocks", SIAM J. Sci. Stat. Comput., vol. 8, pp. 71–81, 1987.

[2] J. W. Chaffee, "Relating the Allan variance to the diffusion coefficients of a linear stochastic differential equation model for precision oscillators", IEEE Trans. UFFC, vol. 34, pp. 655-658, Nov. 1987.

[3] C. Zucca and P. Tavella, "The clock model and its relationship with the Allan and related variances", IEEE Trans. UFFC, vol. 52, no. 2, pp. 289–296, Feb.. 2005.

[4] K. R. Brown Jr., "The theory of the GPS composite clock", Proc. ION GPS-91 Meeting, pp. 223–241, 1991.

[5] M. Weiss and T. Weissert, "AT2, a new time scale algorithm: AT1 plus frequency variance", Metrologia, vol. 28, pp. 65–74, 1991.

[6] S. R. Stein, "Advances in time-scale algorithms", Proc. 24th PTTI Meeting, pp. 289–302, 1992.

[7] S. R. Stein, "Time scales demystified", Proc. Joint IEEE IFCS and EFTF Meeting, pp. 223-227, 2003.

[8] C. A. Greenhall, "Kalman plus weights: a time scale algorithm", Proc. 33rd PTTI Meeting, pp. 445–454, 2001.

[9] C. A. Greenhall, "Addition of random run FM noise to the KPW time scale algorithm", IV Time Scale Algorithms Symposium, Sèvres, 2002.

[10] C. A. Greenhall, "Forming stable timescales from the Jones–Tryon Kalman filter", Metrologia, vol. 40, pp. S335–S341, 2003.

[11] M. A. Weiss, D. W. Allan, and T. K. Peppler, "A study of the NBS time scale algorithm", IEEE Trans. Instrum. Meas., vol. 38, pp. 631-635, 1989.

[12] J. A. Davis, C. A. Greenhall, and P. W. Stacey, "A Kalman filter clock algorithm for use in the presence of flicker frequency modulation noise", Metrologia, Vol. 42, pp. 1–10, 2005

[13] L. Galleani and P. Tavella, "On the use of the Kalman filter in timescales", Metrologia, vol. 40, pp. S326–S334, 2003.

[14] C. A. Greenhall, "A Kalman filter clock ensemble algorithm that admits measurement noise", Metrologia, vol. 43, pp. S311–S321, 2006.

[15] C. A. Greenhall, "A Kalman filter clock ensemble algorithm that admits measurement noise: corrections and update", Metrologia, vol. 44, pp. 491–494, 2007.

[16] M. Suess, D. Matsakis, and C. Greenhall, "Simulating future GPS clock scenarios with two composite clock algorithms", Proc. 42nd PTTI Meeting (to appear), 2010.

A Kalman filter UTC(k) prediction and steering algorithm

J. A. Davis, S. L. Shemar and P. B Whibberley
National Physical Laboratory,
Hampton Road,
Teddington, Middlesex, TW11 0LW
john.davis@npl.co.uk

***Abstract*— A clock prediction and steering algorithm has been developed to produce a UTC(k) timescale that may be maintained close to UTC. The algorithm is based on predicting individual (UTC – C_l) clock offsets which are then combined to form a physical UTC(k) timescale. The algorithm is designed to use data from all the available laboratory clocks without the use of a clock ensemble algorithm. The performance of the underlying Kalman filter predictor is examined in depth, along with initial results using three years of historical NPL clock and UTC data combined with a simulated steering process.**

I. INTRODUCTION

The uncertainty in synchronising a UTC(k) timescale to UTC, and the resulting stability of the UTC(k) timescale, are limited by the latency in computation and publication of Circular T by BIPM. This latency is normally 40 days but may be as large as 45 days. To maintain a highly stable UTC(k) timescale, several laboratories have used clock ensemble algorithms to generate a highly predictable free running ensemble timescale. The prediction properties of this ensemble timescale are used to steer a physical timescale UTC(k) which is then maintained close to UTC.

In this paper an alternative and potentially simpler approach is used to generate a simulated UTC(k) timescale that may be maintained close to UTC, without the requirement to compute an ensemble timescale. An algorithm has been designed that is based on the use of individual (UTC – C_l) clock predictions. A Kalman filter-based clock predictor forms the core of this algorithm. The design and performance of the predictor is presented, using both simulated and real NPL clock and UTC data. Initial results from the generation of a simulated UTC(k) timescale are presented using three years of historical NPL clock and Circular T data and simulating the physical clock steering process.

II. OVERVIEW OF THE ALGORITHM

The UTC(k) predicting and steering algorithm consists of several components. The core of the algorithm is a (UTC(k) – C_l) Kalman filter clock predictor, which is designed to provide close to optimal predictions in the presence of standard clock noise types, including Flicker Frequency Modulation (FFM), and also in the presence of linear frequency drift. An algorithm has been developed to weight the individual clock predictions to provide regular (UTC – UTC(k)) estimates. Finally, a steering component is required in order to maintain the UTC(k) timescale close to UTC using the (UTC – UTC(k)) estimates.

III. EVALUATING THE (UTC – C_l) KALMAN FILTER CLOCK PREDICTOR PERFORMANCE

The Kalman filter clock predictor is an extension of the clock predictor described in [1]. The main components of the algorithm are given below.

The state vector at iteration n is given by:

$$x_n = \left(\xi_n,\ \eta_n,\ \zeta_n,\ m_{1n}\ \cdots\ m_{jn},\ p_{1n}\ \cdots\ p_{kn}\right)^T \qquad (1)$$

where ξ_n, η_n and ζ_n represent the (UTC – C_l) time, normalized frequency and linear frequency drift offsets, respectively; $p_{1n} \cdots p_{kn}$ are a group of integrated Markov noise parameters, typically five, used to model FFM noise processes; $m_{1n} \cdots m_{jn}$ are a second group of Markov noise parameters, typically five, used to model non-white measurement noise. The measurement noise is modeled as elements within the state vector, and the measurement covariance matrix R is not used. The components of the parameter covariance matrix P, process covariance matrix Q, and state propagation matrix Φ are described in depth in [1] and [2], as is the initialization of the Kalman filter.

To test the correct operation of the Kalman filter under close to ideal conditions, a data set was simulated using the same constructions of state vector, process covariance matrix and parameter covariance matrix as are used in the filter. The simulated state estimates x_{Sn} evolve from iteration n–1 to n as:

$$x_{Sn} = \Phi(\tau) x_{S(n-1)} + L_Q(\tau) e_n \qquad (2)$$

where $Q(\tau) = L_Q(\tau) L_Q^T(\tau)$, τ is the data spacing, and e_n is a vector of uncorrelated Gaussian noise values of unit variance. The covariance matrix of the simulated data sets at iteration n is given by P_{Sn} where:

$$P_{Sn} = \Phi(\tau) P_{S(n-1)} \Phi^T(\tau) + Q(\tau)\ . \qquad (3)$$

Figure 1 shows the characterization of a simulated data set in terms of Allan deviation (ADEV), modified Allan deviation (MDEV) and Hadamard deviation (HDEV). The simulation plots are obtained by applying ADEV, MDEV and TDEV statistics to the simulated data sets. The theory plots are obtained by computing the ADEV, MDEV and TDEV expectation values directly from the elements of P_{Sn} using a method similar to that described in section 5 of [2]. The two sets of plots agree well.

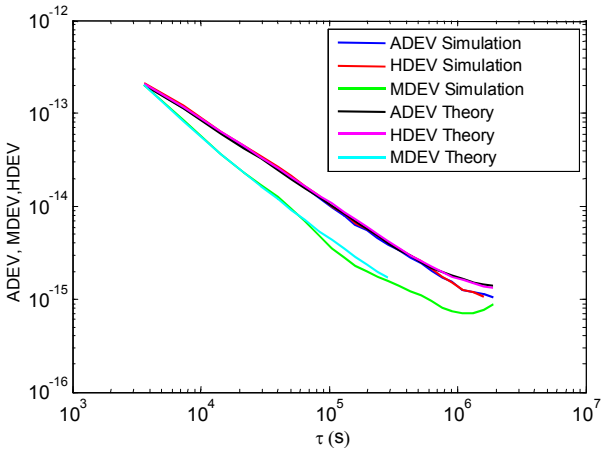


Figure 1. Plots of ADEV, MDEV and HDEV against τ obtained from simulated clock data

The noise process shown in figure 1 is part way between White Phase Modulation (WPM) and Flicker Phase Modulation (FPM), and has characteristics similar to typical time transfer noise.

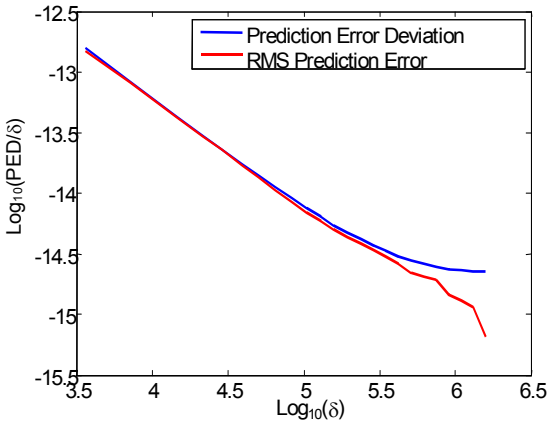


Figure 2. Plots of $Log_{10}(PED/\delta)$ against $Log_{10}(\delta)$ obtained from the Kalman filter predictor operated using the simulated clock data and δ represents the prediction length

In the presence of White Frequency Modulation (WFM) and Random Walk Frequency Modulation (RWFM) the Prediction Error Deviation (PED) is related to ADEV by [3, 4]:

$$PED = ADEV\tau \tag{4}$$

where τ represents the Allan deviation averaging time and is equal to the prediction length δ. This formula is a reasonable first approximation for other power law noise processes. In figure 2 the Root Mean Square (RMS) prediction error has been determined from operating the Kalman filter predictor on the simulated data set used to produce figure 1. The prediction errors are computed from the output data sets, and RMS statistics are then applied. The prediction error deviation is obtained from the post prediction Kalman filter parameter covariance matrix. The plots of $Log_{10}(PED/\delta)$ against $Log_{10}(\delta)$ in figure 2 agree well with the plots of ADEV against τ in figure 1.

The optimal Prediction Error Deviation (PED) obtained in the presence of a FFM noise process is given by [3, 4]:

$$PED = \frac{ADEV\tau}{(\ln(2))^{0.5}} \tag{5}$$

where τ again represents the Allan deviation averaging time and is equal to the prediction length δ.

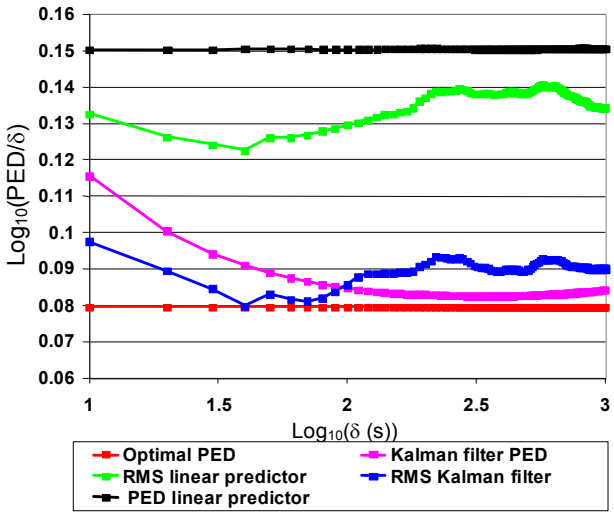


Figure 3. (PED/ δ) estimates obtained from using both the Kalman filter predictor and simple linear predictor on a FFM noise data set

In figure 3 the performance of the Kalman filter predictor is compared against that of a simple linear predictor [1], using a new simulated data set that was dominated by FFM noise. The red line shows the optimal (PED/δ) values obtained from equation 5. Equation 5 assumes a data length and data density prior to prediction that are both tending towards infinity. The pink curves give the theoretical values of (PED/δ) obtained from the Kalman filter covariance matrix. The pink curve deviates from the red line at short prediction lengths when the density of data points relative to prediction length is low, and at large prediction lengths where the total length of the data set is at its smallest multiple of the prediction length. In these two conditions equation 5 is likely to be too optimistic. The black line shows the theoretical performance limit of the simple linear predictor. The blue and green curves are the RMS prediction errors obtained from operating the Kalman filter and simple linear predictor respectively using the same

simulated data set, and then computing the prediction errors from the output data sets. It is concluded that the Kalman filter predictor performs very close to optimal in the presence of FFM noise. There is a significant improvement over using a simple linear predictor.

Figure 4 shows the results of using the Kalman filter clock predictor to predict the time offsets between a pair of NPL's active hydrogen masers. To facilitate direct comparison with the use of a simple linear predictor, the linear frequency drift between the pair of masers was removed before undertaking the analysis.

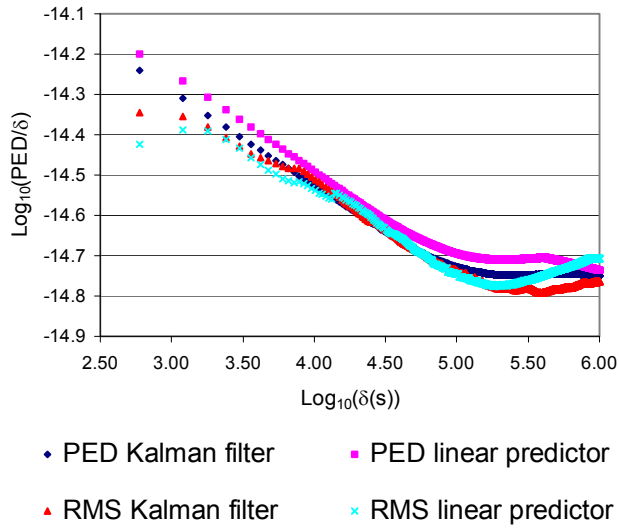


Figure 4. Plots of $Log_{10}(PED/\delta)$ against $Log_{10}(\delta)$ obtained from two of NPLs active hydrogen masers, HM1 and HM2, using both the Kalman filter predictor and a simple linear predictor.

The RMS curves obtained from error statistics agree well with the PED curved obtained from theory, suggesting that the noise models used in the Kalman filter match reality reasonably well. The results obtained from the Kalman filter predictor are better than those from the simple linear predictor, particularly in the long-term where FFM noise dominates.

Active hydrogen masers display significant linear frequency drift. A key advantage of the Kalman filter predictor is that it should be almost insensitive to the presence of linear frequency drift. To demonstrate this characteristic a relatively large linear frequency drift was added to the simulated data sets. The resulting plots of ADEV, MDEV and HDEV against τ are shown in figure 5. The linear frequency drift is clearly identified in the ADEV and MDEV but not HDEV curves obtained from the simulated data sets. The theory values only compute the stochastic component of ADEV, MDEV and HDEV and so are insensitive to linear frequency drift

The resulting plots of $Log_{10}(PED/\delta)$ against $Log_{10}(\delta)$ are shown in figure 6. Comparing figure 6 with figure 2 shows that the theoretical PED values obtained from the Kalman filter remain unchanged and that the linear frequency drift has had only a very slight impact on the RMS prediction errors at averaging times of around one day, the minimum averaging time at which the linear frequency drift dominates the statistics. The ability of the Kalman filter to give close to optimal performance in the presence of linear frequency drift is its main advantage over using a simple linear predictor.

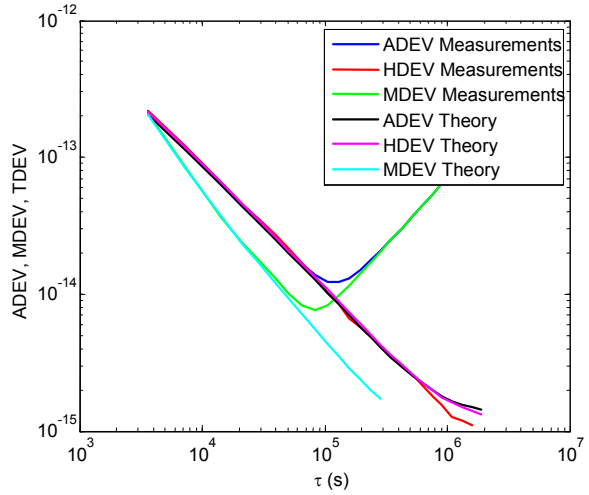


Figure 5. Plots of ADEV, MDEV and HDEV against τ obtained from a simulated data set in the presence of significant linear frequency drift

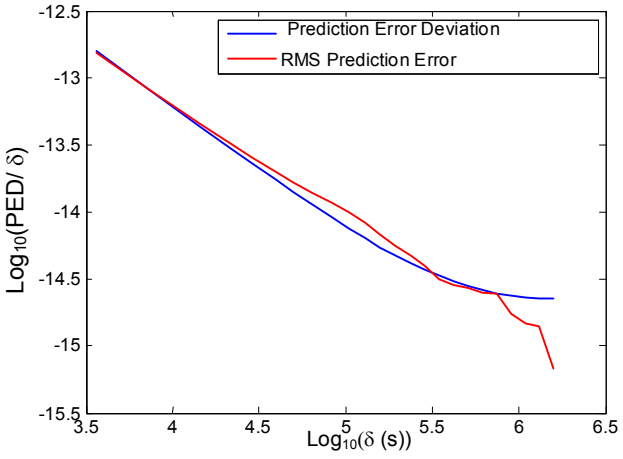


Figure 6. Plots of $Log_{10}(PED/\delta)$ against $Log_{10}(\delta)$ obtained from the Kalman filter predictor operated using the simulated clock data in the presence of strong linear frequency drift

IV. ESTIMATING THE (UTC – UTC(K)) OFFSET

A physical realization of a UTC(k) timescale may be obtained by taking the output signal from an active hydrogen maser and steering the frequency of the output signal using a phase micro-stepper. A counter timer or phase comparator logging system may then be used to make measurements of $(UTC(k) - C_I)$ for each individual free-running laboratory clock C_I.

Circular T data from BIPM contains historical estimates of (UTC – C_l) for each free-running laboratory clock C_l with five-day data spacing. The (UTC – UTC(k)) timescale offsets may then be estimated on an hourly basis using :

$$UTC - UTC(k) = \sum_{l=1}^{r} w_l[(UTC - C_l)_P + (C_l - UTC(k))] \quad (6)$$

where w_l is the weight of individual clock C_l, and r is the number of local free-running laboratory clocks.

Using the Kalman filter clock predictor we are able to predict $(UTC - C_l)_P$ up to 45 or more days ahead and with a much smaller data spacing, e.g. hourly. On receiving the monthly update of (UTC – C_l) the next set of $(UTC - C_l)_P$ predictions will be computed using a whole months additional (UTC – C_l) data. This may result in a discrete change in the estimate of (UTC – UTC(k)). To prevent the algorithm from applying large steers to the UTC(k) timescale to correct these steps, a paper timescale $UTC(k)_S$ is constructed. This timescale has the properties that there is no "step" in $UTC(k)_S$ following the monthly update of BIPM data and that the $(UTC\text{-}UTC(k)_S)$ offset decays exponentially between successive monthly updates. This is achieved by computing the $(UTC\text{-}UTC(k)_S)$ offset up to 50 days ahead at the time the monthly (UTC – C_l) data is being incorporated into the algorithm, and including the latest published value (UTC – UTC(k)) by BIPM. Equation 6 becomes:

$$UTC(k)_S - UTC(k) = \sum_{l=1}^{r} w_l[(UTC(k)_s - C_l)_P + (C_l - UTC(k))] \quad (7).$$

The $(UTC(k)_S - C_l)_P$ individual clock predictions are computed from the $(UTC - C_l)_P$ output of the Kalman filter predictor using:

$$(UTC(k)_S - C_l)_P = (UTC - C_l)_P - (UTC - UTC(k)_S) \quad (8)$$

In the short term e.g. hourly, the physical timescale UTC(k) may be regularly steered to minimize the $(UTC(k)_S - UTC(k))$ offset. Making the assumption that $UTC(k)_S$ may be maintained close to UTC using regular steering it is possible to use historical UTC(k) clock data to determine how well UTC(k) would have been maintained close to UTC assuming that the above weighting algorithm had been applied.

The weights w_l are determined by using the historical prediction RMS prediction errors obtained at the prediction length currently under use. Because the prediction length of the $(UTC - C_l)_P$ predictions will change during each monthly cycle of the algorithm, so do the corresponding clock weights. This is because the optimal clock weight is a function of the prediction length.

V. Results using historical NPL UTC(k) data and simulated rapid steering

Assuming that the physical realization of the UTC(k) timescale may be maintained close to $UTC(k)_S$ then we may use historical UTC(k) data to determine the effectiveness of this UTC(k) prediction and steering algorithm. The errors in the $(UTC - C_l)_P$ individual clock predictions for individual clocks are shown in figures 7 and 8.

Figures 7 and 8 show very sharp changes in the magnitude of the prediction errors on the inclusion of the next month's Circular T data. The prediction errors of active hydrogen masers HM1 and HM2 are generally higher during the first half of the graph. This was a period when NPL was not operating its Two-Way Satellite Time and Frequency Transfer (TWSTFT) earth station. The degraded time transfer using single-frequency GPS measurements resulted in an increased prediction error. The RMS prediction error of the second half of the graph is around 5 ns (1 σ). At a mean prediction length of 30 days this would result in a corresponding approximate ADEV of 2×10^{-15} at an averaging time of 30 days, based on the use of equation 1. This indicates that the Kalman filter predictor is working well and that the two masers are intrinsically predictable in the long term. The results from NPL's high performance (CsHp) and standard tube (CsStd) caesium clocks were as expected. However the prediction errors observed using active hydrogen maser HM4 were far greater than expected and heavily biased, with the sign of the prediction errors changing part way through the run. These effects were caused by a relatively rapid change, followed by a discrete change, in the underlying linear frequency drift of the maser. Such behavior was not accounted for in the clock model.

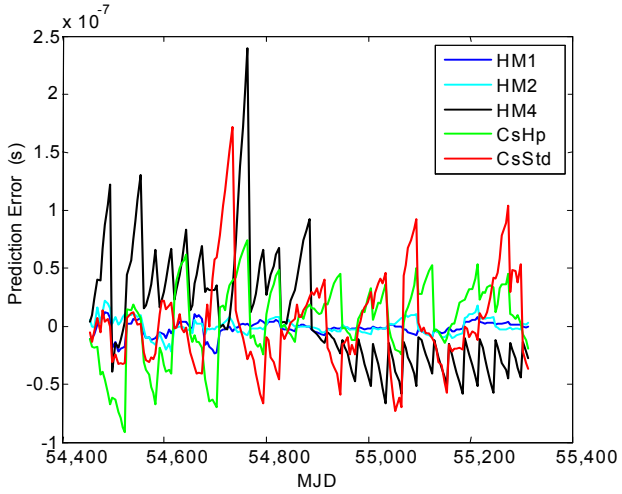


Figure 7. $(UTC - C_l)_P$ prediction errors of each of NPL's clocks

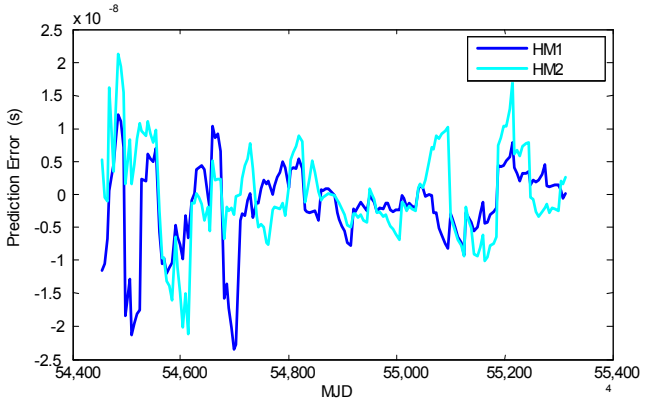


Figure 8. $(UTC - C_l)_P$ prediction errors of NPL's active hydrogen masers HM1 and HM2

Figures 9 and 10 show the resulting (UTC – UTC(k)) timescale offsets obtained using three years of historical UTC clock data. Each offset is computed using a single clock. Again the results obtained from HM1 and HM2 are encouraging, with the offsets being significantly lower during the second half of the period when NPL's TWSTFT system was operational. The (UTC- UTC(k)) offsets obtained using each of the clocks mirrored the prediction errors closely.

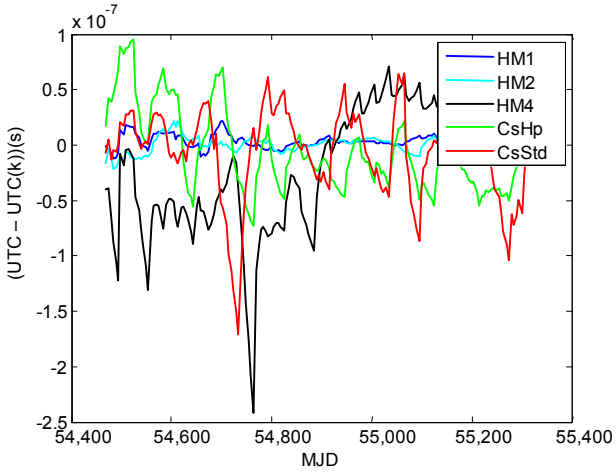


Figure 9. (UTC – UTC(k)) time scale offsets calculated using only one of NPL's clocks

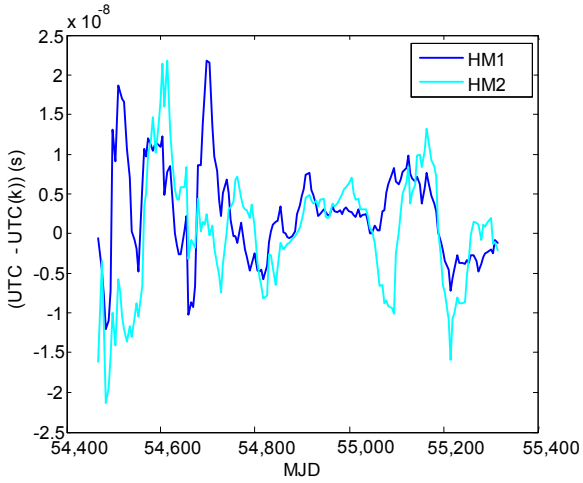


Figure 10. (UTC – UTC(k)) time scale offsets calculated using only NPL's active hydrogen masers HM1 and HM2

Figure 11 shows the (UTC- UTC(k)) timescale offsets again obtained using three years of historical data but computed from a weighted mean of all the (UTC – C_I) clock offset predictions computed as discussed in section IV. The RMS of the (UTC- UTC(k)) offset is significantly lower than was obtained by using any single clock, although the improvement is not dramatic. This is explained by the weighting that attributes well over half of the clock weights to HM2 and the vast majority of the remainder to HM1. This however is a reasonable reflection of the current performance of NPL's commercial atomic clocks. There is also significant GPS and TWSTFT time transfer noise that will be common to all the individual clock measurements. The results are significantly better than is currently achieved using relatively infrequent manual steering of a single active hydrogen maser. Much work is still required in constructing a physical implementation for this algorithm, which has so far only been demonstrated through the post processing of historical NPL clock data and Circular T results.

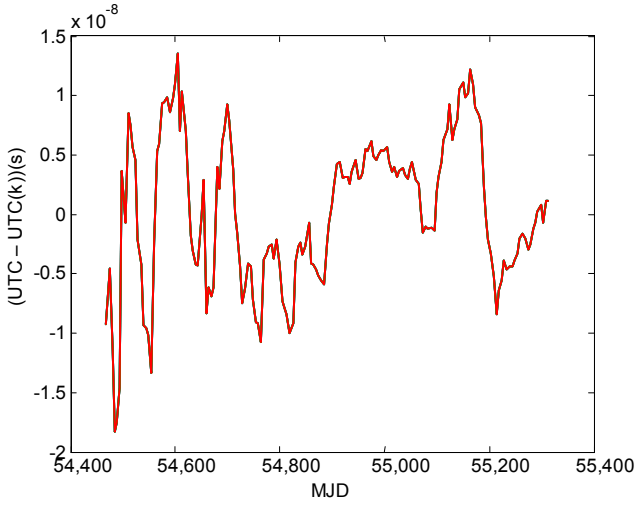


Figure 11. (UTC – UTC(k)) time scale offsets calculated using a weighted mean of all of NPL's clocks

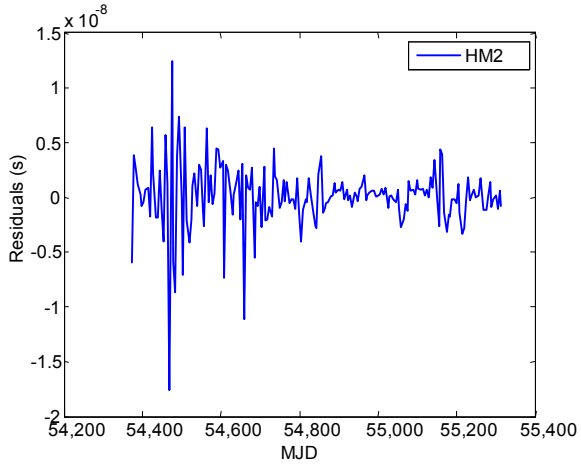


Figure 12. (UTC –C_I) Kalman filter residuals obtained from NPL's active hydrogen maser HM2

Examination of the Kalman filter residuals often provides useful information on the quality of the clock and time transfer noise models used in the Kalman filter. Figures 12 and 13 show the residuals obtained from HM2 and HM4 respectively. In both plots the magnitude of the residuals is lower in the second half of the plots when TWSTFT time transfer was operational. During this period the magnitude of the residuals is very similar despite the resulting prediction errors being very different. In both plots the residuals are close to being white; however, some residual structure may be observed. The key difference is that in the case of HM2 the residuals are

unbiased but in the case of HM4 there is a significant bias that reverses part way through the plot. This is due to a mis-modeling of the linear frequency drift changes occurring with HM4. Either an improved model needs to be developed or a mechanism for detecting significant linear frequency drift changes needs to be incorporated within the Kalman filter.

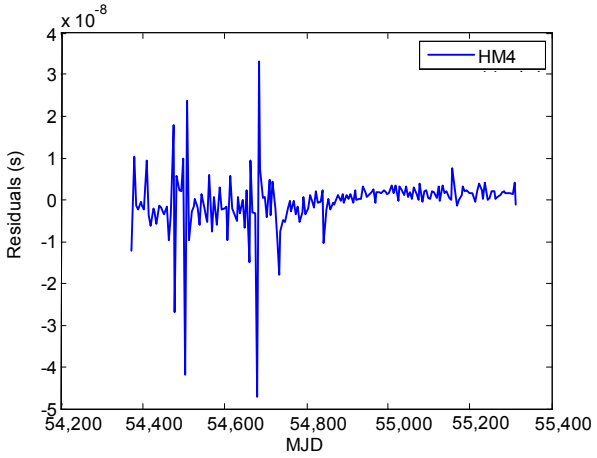


Figure 13. (UTC –C_l) Kalman Filter residuals obtained from NPL's active hydrogen maser HM4

VI. CONCLUSIONS

NPL has developed a prediction and steering algorithm that may be used to maintain a UTC(k) timescale close to UTC. The algorithm operates by weighting individual (UTC – C_l) clock predictions and avoids the requirement to use a clock ensemble algorithm. The core of the algorithm consists of a Kalman filter clock predictor that is shown to produce close to optimal predictions in the presence of FFM clock noise and to be almost insensitive to linear frequency drift. The performance of the algorithm has been demonstrated through the post-processing of three years of historical NPL clock and Circular T data, where the implementation of a rapid steering process has been assumed. Initial results are encouraging. The use of multiple clocks in the algorithm is demonstrated to improve the quality of the resulting UTC(k) timescale and is consistent with the relative weightings given to individual clocks.

Significant work is still required to set up a real-time rapid steering mechanism based on the use of a phase micro-stepper and a suitable algorithm. The prediction algorithm also needs to be improved, to either model better changes in linear frequency drift or detect rapidly the resulting anomalies.

VII. ACKNOWLEDEMENTS

This work is supported by the UK Department for Business, Innovation and Skills as part of the NMS Physical Metrology Programme.

REFERENCES

[1] J A Davis, C A Greenhall, and R Boudjemaa "The development of a Kalman filter clock predictor, Proc 19th EFTF, pp. 90-95, 2005.

[2] J A Davis, C A Greenhall, and P W Stacey, "A Kalman filter clock algorithm for use in the presence of flicker frequency modulation noise", Metrologia **42,** 1-10, 2005.

[3] L-G Bernier, "Use of the Allan deviation and linear prediction for the determination of the uncertainty on time calibrations against predicted timescales". IEEE Trans. Instrumentation & Measurement, **52**, no.2, 2003, p.483-486.

[4] P Tavella and M Gotta, "Uncertainty and prediction of clock errors in space and ground applications". Proceedings of the 14th EFTF, 2000, pp 77-81 Turin.

Synchronizing Computer Clocks Using Kalman Filters

Judah Levine
Time and Frequency Division and JILA
National Institute of Standards and Technology and the University of Colorado
Boulder, Colorado 80305
jlevine@boulder.nist.gov

Abstract -- I have used the Kalman Filter algorithm to improve the link between the Internet Time Servers operated by the National Institute of Standards and Technology (NIST) and the primary atomic clock ensemble in Boulder, which realizes UTC(NIST) and which is used as the reference for the time servers. The Kalman algorithm is better able to separate the contributions of multiple noise sources such as the fluctuations in the asymmetry of the channel delay and the statistical fluctuations in the clock used as the internal reference for the time server. This improved separation has made it possible to compensate to some extent for the lower-quality telephone circuits that are often used as the links for synchronizing the time servers to the atomic clock ensemble.

INTRODUCTION

The National Institute of Standards and Technology (NIST) currently operates 35 public network time servers that are located at 21 different sites in the United States. The servers provide time over the public Internet in a number of different formats. All of the servers are synchronized to UTC(NIST). The time servers are linked to the atomic clock ensemble in Boulder, Colorado by the use of a hard-wired connection for the systems that are located at the NIST Boulder laboratories and by dial-up telephone lines, which implement the ACTS protocol [1] which will be described below, for the other systems that are located at remote locations.

THE ORIGINAL ACTS SYSTEM

The ACTS system transmits time using standard dial-up telephone lines and modems. It is a two-way protocol, in which the one-way transmission delay between the server and the user is modeled as one-half of the round-trip value. The message from the server to the client contains a time stamp and an on-time marker character, which the client echoes back to the server with as little delay as possible. The internal system latency is on the order of microseconds, whereas the channel delay is typically tens of milliseconds, so that the system latency is small enough to be ignored. (Both the server and the client have operating system software that has been modified to guarantee that the system latency will be not larger than 15 microseconds even when the client is handling many requests for time from the network.) The round-trip delay is measured by the server as the time that has elapsed from when it transmitted the on-time marker to when it receives the echo from the remote system. The one-way delay is estimated as one-half of this value, and the next on-time marker is advanced using this one-way delay estimate so that it will arrive at the client on-time. The server inserts this one-way delay estimate into each message, so that the client can see the advance that was used. When the client receives the on-time marker, it uses the time stamp of that message to compute the difference between its clock and the clock on the server. This process is repeated every second until the telephone connection is broken. The client can break the connection at any time, and the server will automatically break the connection after 40 s if the client has not done so before this.

There are two important aspects to the design of the ACTS system. The first is that the advance that is applied to any on-time marker is derived from the previous round-trip measurement, so that fluctuations in the delay with a period close to the one-second interval between messages will not be handled correctly. The second aspect, which has more subtle consequences, is that the delay is measured by the server and not by the client. Neither of these was important when ACTS was first designed in 1988, but changes in the telephone system since that time have made both of them significant.

The advance used by the ACTS servers will be accurate and unbiased on the average if the asymmetry in the delay is close to 0 (so that the one-way delay is accurately estimated as one-half of the round-trip value) and if both the delay itself and its asymmetry are well characterized as white noise processes, so that the advance that is used by the server is an accurate estimate of the one-way delay on the average. When these conditions are satisfied, an average of the time differences between the client and the server measured using consecutive ACTS transmissions converges to the true time difference between the two systems. The standard-deviation of the mean is improved by the square root of the number of measurements that are used to compute it, and it was relatively easy to realize sub-millisecond timing accuracy using telephone calls that lasted 15 or 20 seconds.

THE EFFECT OF NEWER TELEPHONE CIRCUITS

The original ACTS servers had no way of evaluating the accuracy of the one-way delay estimate, since they could not estimate the asymmetry of the delay. If we define a symmetry parameter, k, where k= 0.5 indicates that the inbound and outbound delays are exactly equal, then the error in the time-difference measurement, Δt, resulting from an asymmetry in the measured round-trip delay, D, is given by,

$$\Delta t = (k - 0.5)D. \quad (1)$$

The limiting values k= 0 and k= 1 indicate that the inbound or outbound delays (with respect to the server), respectively, are negligibly small compared to the delay in the opposite direction. That is, when k= 0 the transit time from the server to the client dominates the round trip value, so that the clock on the client appears slow.

The measured round-trip delay on analog voice circuits was typically about 0.08 s and the delay asymmetry was typically less than 1%, so that the timing accuracy was of order 0.001 s. These values were realized using the same brand of modem on both ends of the connection and by using a signaling speed of 9600 baud, which has the minimum asymmetry for the modems that we used.[2]

Both the asymmetry and the delay are larger on newer telephone circuits, which are often implemented as a combination of conventional analog and digital packet-switched technologies. This combination is increasingly common even on local subscriber loops – the circuit that links the end user to the local telephone exchange. These circuits often have round-trip delays as larger as 0.25 s, with a varying delay asymmetry that can reach 5% or even more. From eq. 1, the timing error in this case exceeds 0.01 s – about a factor of 10 poorer than the original design. As a practical matter, many users of the Internet Time Services do not need even this level of accuracy and the degraded service may still be adequate for their needs. Nevertheless, I have designed software to try and ameliorate this degradation as much as possible, and I describe the initial tests of these improvements in this paper.

A NEW ACTS CLIENT

In the original ACTS design, the client only had to echo the on-time marker back to the server with negligible delay – all the real work of the protocol was done by the server. However, the client can measure the apparent time difference between the time of the local clock and the time stamp in the ACTS message every second – information that is not available to the server. In order to make use of this information, we must construct a model of the clock in the client system. We use the typical iterative model, in which the time difference at the current epoch, t, is estimated based on a clock model computed at a previous time t-τ. The parameters of the model are the time difference as a function of epoch, x, measured in s and the dimensionless frequency offset, y. The parameters ξ and η are the stochastic noise contributions to the measurement process and the frequency of the local clock, respectively. We assume that the measurement noise, ξ, is approximately stationary white phase noise, so that it does not depend on t or τ. The frequency noise, η, is typically estimated using the Allan deviation for an averaging time of τ. We assume that the Allan deviation is stationary as so does not depend on t.

$$x(t) = x(t-\tau) + y(t-\tau)\tau + \xi \quad (2)$$

$$y(t) = y(t-\tau) + \eta(\tau) \quad (3)$$

Although clock models often include a frequency aging parameter, *d*, whose units are s^{-1}, this parameter is not very useful in this application for several reasons. The frequency aging parameter would add terms $0.5d\tau^2$ to eq. 2 and $d\tau$ to eq. 3. In both cases, the contribution of the aging parameter is masked by the stochastic contributions to the corresponding estimates, so that it is difficult to compute a robust estimate of the aging parameter unless relatively large values of τ are used. Unfortunately, the aging parameter is usually not a constant over these longer averaging times, since the frequency of the oscillator is usually affected in a quasi-random manner by temperature fluctuations and other local environmental perturbations.

The quartz-crystal oscillators used in computer systems typically have y= 2×10^{-5}, η(1)= 10^{-7}, and η(30)= 1.8×10^{-8}. (These parameters characterize the oscillator as seen through the operating system software, and therefore include system latency and jitter. It is almost always impossible to measure the characteristics of the actual "bare" oscillator itself, which is probably considerably more stable than these values of η for almost all averaging times.) If we take 30 s as the maximum duration of a typical connection between a client system and the ACTS server, the stochastic frequency variations contribute about 0.1 μs to the estimates of consecutive time differences spaced 1 s apart, and about 0.5 μs to the dispersion of the time differences over the entire connection. In contrast, the measurement noise, ξ, is typically at least 80 μs even on a very good telephone connection, so that the contribution of the stochastic frequency fluctuations plays no role over the duration of a telephone call. The measurement noise is larger than the system latency because of the quasi-synchronous operation of the message interchange over the telephone line between the two modems. Therefore, the measured time dispersion of the clock in the client system over the duration of the telephone connection can be modeled as white noise superimposed on a simple linear variation. The linear variation can be ignored relative to the measurement noise for short telephone connections lasting less than about 4 s, but must be considered for the 30 s connections that we use for the synchronization of the time servers.

DETAILS OF THE MEASUREMENT ALGORITHM

The client system connects to the ACTS server and receives *N* time messages. Each message contains a time stamp and the advance applied to the on-time marker of that transmission. As described above, the advance parameter was actually computed by the server using the previous measurement that it made of the round-trip delay. The client measures the time difference between its clock and the time stamp of each of the messages, and also records the associated advance parameter. These values are T_i and A_i, respectively.

The algorithm models each of the advance values as the sum of a constant value that is a characteristic of that telephone connection and an additional stochastic value that has some unknown statistical distribution. We explicitly *do not* assume that the distribution of the advance values about the constant value has any particular form. However, we do assume that the variations in the advance have a mean of 0. That is, the mean of the advance values over the entire telephone connection is an unbiased estimate of the true transmission delay. (This assumption is basic to all two-way methods. There is no way of detecting a static bias in the asymmetry.) Thus, the algorithm assumes that the true advance for this telephone connection is given by

$$\overline{A} = \frac{\sum_{i=1}^{N} A_i}{N} \tag{4}$$

and that the additional stochastic contribution to the advance in each message is given by

$$a_i = A_i - \overline{A} \quad . \tag{5}$$

We model the measured time differences (by the use of eq. 2) with a constant time offset, a linear time variation due to a deterministic, constant frequency offset and a measurement noise parameter, which we assume to be stationary. We assume that the measurement noise has a random distribution with a mean of 0 that can be characterized as white phase noise. Therefore, the measured time differences should be accurately modeled using a least-squares straight line, and the measurements should scatter about that line with a random distribution that is characterized by the measurement noise of the process. We ascribe statistically significant deviations from the straight line assumption (relative to the estimate of the measurement noise) as due to an error in the advance parameter for that measurement. The error can be due either to a change in the asymmetry of the telephone connection or to a rapid change in the actual round-trip delay between one-second measurements. Our experience is that rapid changes in the round-trip delay are much less likely after the first few seconds of the connection, and changes in the asymmetry are much more common. See fig. 2 of [2], which shows that the advance varies by less than 1 ms P-P after the first few seconds of the connection.

I can illustrate the basis of the method with a simple example. The client system connects to the ACTS server and receives *N* messages, each of which has a time tag, an advance parameter and an on-time marker. When the on-time marker is received, the client measures the time difference between the time tag and the time of its clock. It also records the advance value that was used.

The first step of the analysis is to fit the time differences to a least-squares line and examine the residuals. In the simplest case, the RMS magnitude of the residuals is of order 1 ms, and the maximum deviation is not greater than 3 ms. The time difference is taken as the intercept of the fitted line (with the appropriate time tag), the frequency is the slope of the line, and the algorithm continues with the next step as described below.

If any residual exceeds the threshold of 3× the RMS value, the algorithm examines the advance values to see if the residual could be due to an error in the advance calculation. The algorithm calculates the mean advance and the deviation of each individual advance from that mean as in eq. 4 and 5. The time difference associated with each message is corrected by the difference between the advance used for that message and the average of all of the advances. This operation models statistically large time differences as due to errors in the advance calculation by the server – most often a result of a change in the asymmetry of the connection, which the server cannot detect. The sign of the correction depends on the direction of the asymmetry – whether the inbound or outbound paths were affected, as shown in eq. 1. Therefore, the algorithm must investigate both possibilities and choose the one that reduces the magnitude of the residuals. In most cases, this calculation reduces the scatter in the time differences, confirming the assumption that fluctuations in the advance are really the cause of the apparent fluctuations in the time differences. Any time difference that still exceeds 3× the standard deviation of the residuals of a least squares straight-line fit to the modified time differences is taken to be an error due to some other unknown reason. It is dropped from the estimate and the least-squares line is re-calculated. When no further outliers are detected, the time difference and frequency offset are estimated as the intercept and slope, respectively, of the fitted line, and the algorithm continues with the next step. If more than 5 points are discarded in this process, the estimate does not proceed. Either the RMS deviation of the model is too optimistic for the real data or the system has failed. The algorithm waits a short time and tries again; the time server is declared unhealthy as a precaution, since the time difference calculation did not complete successfully. The time server will continue to be set to an unhealthy state as long as the problem continues. This condition will trigger an operator alarm.

In the more general case, we use the standard Kalman formalism to estimate average values of each of the parameters over the duration of the telephone connection. In other words, the Kalman algorithm partitions the variance of the time differences measured during a single telephone connection as

due to fluctuations in the advance parameter, which have some unknown statistical distribution with a mean of 0 and a linear time variation with a slope and intercept that are constant during the connection. This calculation is considerably more complicated than the simpler algorithm described above, but it is more general and able to handle more complex noise types. For example, we have experimented with adding a diurnal term to the Kalman model, which could be helpful in estimating the effect of the nearly-diurnal fluctuations in ambient temperature on the frequency of the oscillator in the client. This term must be incorporated into the covariance matrix, since we do not know the amplitude or phase of the admittance to the temperature. This process is often described as "state vector augmentation." [3]

When the computation is completed, we have an estimate of the time offset of the system clock, the average frequency over the time of the connection, and the measurement noise parameter, ξ, which is essentially the RMS value of the residuals of the computation.

We cannot determine the stochastic frequency variation, η, in this way, since its contribution to the time differences is too small over the duration of the telephone call. For averaging times less than a few hours, the frequency fluctuations are well modeled as white frequency modulation. Therefore, we estimate the frequency of the oscillator by

$$y(t) = \frac{y(t-\tau) + T_y \dfrac{x(t) - x(t-\tau)}{\tau}}{1 + T_y} \quad . \tag{6}$$

The first term in the numerator on the right side is the estimate of the frequency on the previous measurement cycle, and the second term is the current average frequency estimated as the evolution of the time difference over the interval since the last measurement cycle. When the time server is operating in steady state, the control software drives the time difference to zero on every measurement cycle, so that $x(t-\tau)= 0$ in this mode. The estimate of the current frequency is then simply the time difference that has accumulated divided by the interval between measurments. The weighting factor T_y implements an exponential filter with a time constant derived from the assumption that the frequency variations can be characterized as white frequency modulation. For most of the computers that we have tested, this assumption is valid for averaging times up to about 12 000 s, so that

$$T_y = \frac{\tau}{12000} \quad . \tag{7}$$

For averaging times shorter than τ, the value of η, the stochastic variation in the frequency of the local clock oscillator is measured using special-purpose hardware that connects directly to the computer bus.

The value of τ, the time interval between measurement cycles is determined as a compromise among several competing considerations. If the interval between measurements is too short, it is difficult to estimate the deterministic frequency, y, in the presence of the measurement noise, ξ and the stochastic fluctuations in the estimate of the delay. Using a conservative value of 1 ms for the noise term, the interval between measurements should be at least

$$\tau \geq \frac{0.001}{2 \times 10^{-5}} = 50\,s \quad . \tag{7}$$

In the limit of a very short measurement interval, the measurement noise dominates the calculation, and the performance is dominated by the characteristics of the noisy channel. The local clock is more stable than the remote clock seen through the noisy channel, and the inherent frequency stability of the local clock oscillator is not being exploited in an optimum way. On the other hand, the interval cannot be made too long without violating the assumption that the evolution of the time difference over the measurement interval satisfies the model equations. The algorithm must also consider the possibility of non-statistical glitches in the clock oscillator, and decreasing the interval between measurements would decrease the impact of these glitches because they would be detected (and removed) more rapidly. Finally, the cost of the telephone connections varies as $1/\tau$. Using the interval between calibrations based on the minimum value in eq. 7 would be too expensive and therefore impractical for this reason.

The value of τ was initially about 3 000 s – 4 000 s, but this value is generally too large for some of the poorer telephone connections, and a value of 2 000 s is often needed to realize an acceptable level of timing accuracy at the server of no worse than 5 ms P-P.

DISCUSSION AND CONCLUSIONS

The algorithm I have developed has made it possible to continue to operate the NIST Internet time servers with the telephone circuits that are becoming increasingly common both for the local connection to the telephone central office and for long-distance links. Maintaining a network of time servers synchronized in this way is important for two reasons. In the first place, the time service provided by these systems is independent of satellite signals in general and the global positioning satellites in particular, so that the systems are not compromised by a denial of service attack on satellite signals. In addition, the servers do not need external antennas and can be installed at any location that has telephone connections. This is a significant advantage, since many of our systems are installed in underground or windowless, secure locations with no access to the outside. On the hand, the servers depend on periodic long-distance connections to the ACTS serves in the NIST Boulder laboratories, and this is an expense that would

not be needed for a system that was synchronized using a satellite timing receiver.

As with all two-way methods, the algorithm cannot detect a static asymmetry in the measurement of the round-trip delay of the telephone circuit. We can estimate the magnitude of this effect by comparing the times of different servers to each other, but these comparisons introduce possible asymmetries in the delay of the Internet path between them. As a practical matter, the asymmetries of the Internet paths are typically at least as large or even larger than the values for the telephone system, so that the results of these comparisons are not definitive.

The method I have described here cannot cope with every type of asymmetry. At one of our sites, the asymmetry had a bimodal character. That is, the asymmetry was constant for some period and then abruptly changed to a different value. The change in the asymmetry resulted in time steps of order 20 ms with an irregular period of a few hours. I had no basis for eliminating these steps, since I didn't know the actual time offset of the client system. Both of the time offsets satisfied the usual statistical tests within a single telephone connection, and the asymmetry was constant over several telephone connections, suggesting that the routing of the telephone calls changed periodically in an unpredictable manner.

The problem of a bi-model asymmetry is not very common on telephone connections – we have only one example on 35 connections. However, a slowly varying asymmetry is common, and it also arises quite often in time synchronization applications that use the Internet as the communications medium. A nearly bimodal asymmetry on the Internet can arise for a system, such as a web server, that often has very different inbound and outbound network activity. The "Huff-'n-Puff" [4] filter is an attempt to deal with both of these cases. The procedure assumes that the minimum delay will also have the minimum asymmetry, which is often true for web servers and similar systems because the asymmetry is a strong function of the load on the client or on the server and not a fundamental characteristic of the intervening path itself, which is assumed to have better symmetry. We do not have a good idea of the source of the varying asymmetry in our configuration, and it is less clear that assuming that the minimum delay also has the minimum asymmetry is appropriate. Nevertheless, there could be an advantage to using only the time differences with the minimum delay in the computation of the estimate of the time of the client. From eq. 1, the time error due to any asymmetry is bounded by one-half of the round-trip delay, so that a smaller delay will have a smaller time error, even if the asymmetry fraction is the same. However, even the shortest delay is many tens of milliseconds, so that this limit is not very useful, since the goal of the time service is to synchronize the servers with an uncertainty of 1 ms RMS. Unfortunately, we also do not have an independent estimate of the true time difference between the two systems, so that we have no basis for choosing one time difference over the other one.

One way to address this asymmetry problem would be to improve the frequency stability of the oscillator in the time server so that much longer averaging times could be supported. The easiest way to do this would be to add an external atomic frequency standard or even a high-quality quartz oscillator and then lock the computer oscillator to the external device. Some of our time servers have external rubidium oscillators interfaced to the system through an interrupt line on the serial port for this purpose. Both the frequency offset and the frequency stability of the clock oscillator are improved by several orders of magnitude using this technique.

When an external oscillator is present, the time-varying asymmetry (and similar slowly-varying perturbations) can be averaged for much longer periods – at least 1 day and probably longer than this, since the frequency stability of the enhanced local oscillator will now support these longer intervals between measurements. Since the variations in the asymmetry are bounded, it can always be divided into two components: a static value and a bounded variation about that value that has a mean of 0 for sufficiently long averaging times. A long averaging time would attenuate the time variations in the asymmetry and convert a static asymmetry into a time offset. This time offset would have a slow variation in time if the static value of the asymmetry was not stationary. (This would be true if the asymmetry had a random-walk spectrum, for example, since such a spectrum has no robust mean value.) The time offset could be estimated through comparisons with the NIST time servers that are located at NIST facilities and are synchronized by direct connection to the NIST clock ensemble. These directly-connected time servers would serve as the constraints on the ensemble of network servers, so that the ensemble would provide a stable source of time messages that would be less affected by the asymmetry of any single network path. To the extent that the network paths are disjoint, an ensemble of network paths might also have an overall asymmetry smaller than the asymmetry of any one of them.

REFERENCES

[1] J. Levine, M. Weiss, D. D. Davis, D. W. Allan and D. B. Sullivan, "The NIST Automated Computer Time Service," J. Res. NIST, vol. 94, pages 311-321, 1989.

[2] Judah Levine, "Improvements to the NIST network time servers," Metrologia, vol. 45, pages S12-S22, 2008.

[3] Arthur Gelb, Editor, "Applied Optimal Estimation," 1974, Cambridge, Massachusetts, The M.I.T. Press. See pages 78-82.

[4] David L. Mills, “Computer Network Time Synchronization,” Boca Raton, Florida, CRC Press, 2006. See pages 54-55 and 103-104.

Using the Kalman filter to detect frequency jumps in atomic clocks

Lorenzo Galleani
Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino – Italy
Email: galleani@polito.it

Patrizia Tavella
INRIM
Strada della Cacce 91
10135 Torino – Italy
Email: tavella@inrim.it

***Abstract*— We propose an algorithm for the detection of frequency jumps in atomic frequency standards. The algorithm is based on the innovation generated by the Kalman filter. When the magnitude of the innovation is larger than a given threshold, an anomaly is detected. We characterize the performance of our method for the case of a Rubidium clock by estimating, through numerical simulations, the probability of false alarm, the probability of detection, and the latency time on a large number of numerical simulations.**

I. Introduction

The fast detection of anomalous behaviors of atomic frequency standards is becoming a very urgent issue in the most demanding technological applications, such as navigation systems. The capability to identify an anomalous behavior of a frequency standard is a need in timekeeping and in several other metrological applications, but the peculiarity of navigation systems is that any anomaly which can impact the user positioning has to be detected and signaled in short times and with very low probability of false alarm. These requirements ask for an extremely fast and efficient algorithm, whose design is challenging. Several algorithms have been proposed in the last years [1]-[6], and the research is still very active. In [3] anomalies are detected by analyzing the variations of the clock stability with time, represented by the dynamic Allan variance [7], [8].

In this article we propose a detection algorithm based on the Kalman filter. The Kalman filter [9] estimates the state by a two-step procedure: first, the current state is predicted, based on the previous estimate; then, when the new measurement is available, the prediction and the measurement are properly averaged by taking into account their variance. Therefore, the Kalman filter can be seen as a weighted average of the prediction (based on the knowledge of the dynamical evolution of the system) and of the new measurement. We now note that the prediction should be close, on the average, to the new measurement. A large deviation of the difference between the new measurement and the prediction, known as innovation, can hence be used as an anomaly detector. This idea is the key point of our method and, even though this use of the innovation is already known [10], to the best of our knowledge it has never been applied to atomic clocks.

First, the paper describes the clock model and the proposed implementation of the Kalman filter. Then, the evaluation of the innovation as a test on possible anomalies is introduced. Finally, the detector is characterized for the case of a space Rubidium Atomic Frequency Standard, and the probability of false alarm, the probability of detection, as well as the latency time are estimated by numerical simulations.

II. Kalman Filter Detector

We first define the model of the clock noise and then we describe the Kalman filter detector.

A. The clock noise model

We assume that the frequency deviation $y(t)$ is made by the sum of a white frequency noise (WFN) $\xi_{WFN}(t)$, of a random walk frequency noise (RWFN) $W_{RWFN}(t)$, and of a slowly varying deterministic drift $d(t)$

$$y(t) = \xi_{WFN}(t) + W_{RWFN}(t) + d(t) \tag{1}$$

The autocorrelation function of $\xi_{WFN}(t)$ is given by

$$R_{WFN}(t_1, t_2) = E\left[y(t_1)y(t_2)\right] = \sigma_1^2 \delta(t_1 - t_2) \tag{2}$$

where E is the expected value and $\delta(t)$ is the Dirac delta function. The RWFN noise component is a Wiener process defined as

$$W_{RWFN}(t) = \int_0^t \xi_{RWFN}(t')dt' \tag{3}$$

where

$$R_{RWFN}(t_1, t_2) = \sigma_2^2 \delta(t_1 - t_2) \tag{4}$$

In discrete-time, we model RWFN and the drift as [11], [12]

$$\mathbf{x}[n] = \mathbf{\Phi}\mathbf{x}[n-1] + \eta[n-1] + d[n] \tag{5}$$

where $n = t/T_s$ is the (dimensionless) discrete time, T_s is the sampling time, $\mathbf{x}[n]$ is the two-component vector

$$\mathbf{x}[n] = \begin{bmatrix} x_1[n] \\ x_2[n] \end{bmatrix}, \tag{6}$$

$\mathbf{\Phi}$ is the transition matrix given by

$$\Phi = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}, \tag{7}$$

978-1-61284-110-6/11/$26.00 ©2011 IEEE

$\eta[n-1]$ is a two-component vector of Gaussian random variables with zero mean and covariance

$$Q = \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & 0 \end{bmatrix}, \tag{8}$$

and $d[n] = d(nT_s)$ is the deterministic drift. We take the initial conditions

$$\mathbf{x}[n] = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{9}$$

We now add WFN by making believe that it is an additive measurement noise $v[n]$, obtaining

$$z[n] = H\mathbf{x}[n] + v[n] \tag{10}$$

where

$$H = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{11}$$

is the measurement matrix, $v[n]$ is a Gaussian random variable with zero mean and variance $R = \sigma_1^2$. Therefore, $z[n]$ represents the discrete-time average frequency deviation corresponding to the frequency deviation $y(t)$ of Eq. (1), and it is given by

$$z[n] = \frac{1}{T_s} \int_{(n-1)T_s}^{nT_s} y(t')dt' \tag{12}$$

B. Kalman filter detector

The system model equation of the Kalman filter detector is given by

$$\mathbf{x}[\mathbf{n}] = \mathbf{\Phi}\mathbf{x}[n-1] + \bar{\eta}[n-1] \tag{13}$$

where $\bar{\eta}[n-1]$ is a two-component vector of Gaussian random variables with zero mean and covariance matrix

$$\bar{Q} = \begin{bmatrix} \bar{\sigma}_2^2 & 0 \\ 0 & 0 \end{bmatrix} \tag{14}$$

We take $\bar{\sigma}_2$ to be larger than σ_2 to accelerate the transient phase that arises when a frequency jump occurs. The measurement equation of the Kalman filter is identical to Eq. (10).

At time n, we first obtain the extrapolated version of the previous estimate as

$$\hat{\mathbf{x}}[n|n-1] = \mathbf{\Phi}\hat{\mathbf{x}}[n-1|n-1] \tag{15}$$

Then, we build the innovation

$$\nu[n] = z[n] - H\hat{\mathbf{x}}[n|n-1] \tag{16}$$

whose variance is given by

$$\sigma_\nu[n] = HP[n|n-1]H^T + R \tag{17}$$

where $P[n|n-1]$ is the extrapolated error covariance matrix. Finally, we detect an anomaly when

$$|\nu[n]| > k\sigma_\nu[n] \tag{18}$$

where k is a parameter which influences the performance of the detector.

III. EXAMPLE

In Fig. 1 we show the simulated average frequency deviation (blue line) of a Rubidium clock experiencing a frequency jump, and the estimate $\hat{x}_1[n]$ generated by the Kalman filter (black line). The red vertical bars are located at the time instants corresponding to a detection. We see that the detector quickly reveals the frequency jump. After the frequency jump, the estimate $\hat{x}_1[n]$ experiences a transient phase and then adapts to the time-varying drift. The Rubidium clock is simulated by using the following values of the parameters

$$\sigma_1 = 5 \times 10^{-12} \tag{19}$$
$$\sigma_2 = 5.2 \times 10^{-16} \tag{20}$$
$$d = 1 \times 10^{-12}/\text{day} \tag{21}$$
$$T_s = 1\text{s} \tag{22}$$

The detection threshold is obtained by using $k = 5$.

Performance analysis. We describe a preliminary statistical characterization of the detector obtained by Monte Carlo simulations. When $k = 5$, the probability of false alarm (PFA) estimated on 10^8 realizations is given by

$$PFA = 5 \times 10^{-7} \tag{23}$$

The probability of detection (PD) estimated on 10^4 realizations of the frequency jump with size

$$\Delta_f = 2.5 \times 10^{-11} \tag{24}$$

is given by

$$PD = 0.967, \tag{25}$$

and the corresponding latency time is

$$\Delta n = 1.31 \text{ s} \tag{26}$$

Smaller jumps can be detected by increasing the latency time. This problem is currently under study.

IV. CONCLUSIONS

An anomaly detector based on the evaluation of the innovation, which is a quantity embedded in the Kalman filter, is presented and applied to atomic clocks. The developed detector seems to have several advantages. First, it is fast, in that it can detect anomalies with a very short latency time. Second, due to the recursive nature of the Kalman filter, the algorithm has low computational cost and requires the storage of a small and fixed amount of data. Therefore, being fast and light, the detector can be used in real time applications and onboard satellites, where low computational power is available.

We are currently working on several improvements to our method, such as the detection of slow frequency jumps.

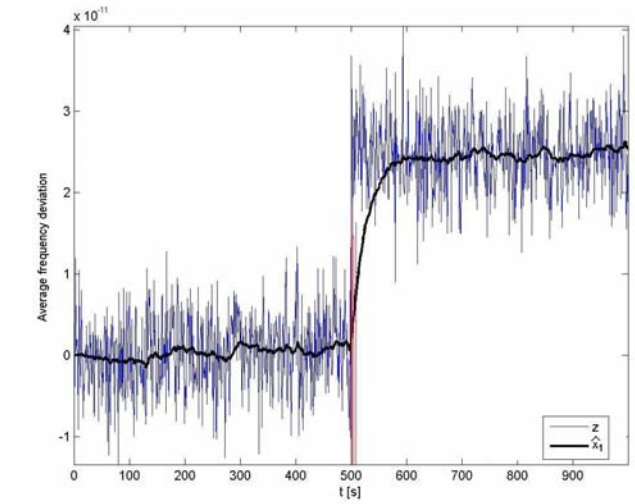


Fig. 1. Simulated average frequency deviation of a Rubidium clock (blue line) and estimated component $\hat{x}_1[n]$ (black line). The vertical bars are located at the detection time instants.

REFERENCES

[1] W. J. Riley, "Algorithms for frequency jump detection," *Metrologia*, vol. 45, pp. S154-S161, 2008.

[2] L. Galleani and P. Tavella, "An algorithm for the detection of frequency jumps in space clocks," *42^{th} Annual PTTI Meeting*, 15-18 November 2010, Reston, Virginia, USA.

[3] L. Galleani and P. Tavella, "Statistical characterization of clock anomalies in the Dynamic Allan variance domain," *Metrologia*, vol. 45, pp. S127-S133, 2008.

[4] M. Weiss, P. Shome, and R. Beard, "GPS Signal Integrity Dependencies on Atomic Clocks," 34^{th} PTTI, 3-5 December 2002, Reston, Virginia, USA.

[5] E. Nunzi, L. Galleani, P. Tavella, and P. Carbone, "Detection of anomalies in the behavior of atomic clocks," *IEEE Trans. on Instrumentation and Measurement*, vol. 56, no. 2, pp. 523-528, 2007.

[6] I. Rodriquez Perez, C. Garcia Serrano, C. Catalan Catalan, A. Mozo Garcia, P. Tavella, L. Galleani, and F. Amarillo, "Inter-Satellite Links for Satellite Autonomous Integrity Monitoring Advances in Space Research," *Advances in Space Research*, pp. 197-212, 2011. Available online at http://dx.doi.org/10.1016/j.asr.2010.07.019

[7] L. Galleani and P. Tavella, "The Dynamic Allan Variance," *IEEE Trans. Ultra. Ferro. Freq. Contr.*, vol. 56, no. 3, pp. 450-464, 2009.

[8] L. Galleani, "The Dynamic Allan Variance II: A Fast Computational Algorithm," *IEEE Trans. Ultra. Ferro. Freq. Contr.*, vol. 57, no. 1, pp. 182-188, 2010.

[9] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.

[10] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global Positioning Systems, Inertial Navigation, And Integration*, John Wiley, New York, 2001.

[11] L. Galleani, "A tutorial on the two-state model of the atomic clock noise," *Metrologia*, vol. 45, pp. 175-182, 2008.

[12] L. Galleani, L. Sacerdote, P. Tavella, and C. Zucca, "A mathematical model for the atomic clock error," *Metrologia*, vol. 3, pp. 257–264, 2003.

Dark Line Resonances in Cs-Ne Vapor Microcells for Chip Scale Atomic Clocks

Rodolphe Boudot, Piotr Dziuban, Xiaochi Liu, Madoka Hasegawa, Ravinder Chutani, Serge Galliou, Vincent Giordano and Christophe Gorecki
FEMTO–ST Institute
UMR 6174 CNRS - ENSMM
32 av. de l'Observatoire, 25044 Besançon Cedex, France
Email: rodolphe.boudot@femto-st.fr

***Abstract*— We report the characterization of Coherent Population Trapping (CPT) resonances in Cs vapor microcells filled with Neon (Ne) buffer gas. The impact on the atomic hyperfine resonance of some external parameters such as laser intensity and cell temperature is studied. We show the suppression of the first-order light shift by proper choice of the microwave power. The temperature dependence of the Cs ground state hyperfine resonance frequency is shown to be canceled in the 77-80°C range for various Ne buffer gas pressures. We preliminary demonstrate a 852 nm VCSEL-modulated based CPT atomic clock exhibiting a short term fractional frequency instability $\sigma_y(\tau) = 1.5 \times 10^{-10}\ \tau^{-1/2}$ until 200 s.**

I. Introduction

Coherent Population Trapping physics [1] combined with the progress of micro-fabrication techniques and semiconductor laser technologies allows the realization of highly miniaturized battery-powered chip scale atomic clocks (CSAC). These references can be used in a wide-range of electronics-based systems such as global positioning, synchronization of telecommunication networks or military applications. The first microfabricated CPT clock was demonstrated at NIST in 2004 [2] while such sources now become commercially available [3]. Nevertheless, leading research remain to be done in this field to further improve the performances of these clocks in terms of frequency stability, power consumption or size or to propose simpler and original architectures.

A CSAC is based on the combination of a high-frequency modulated Vertical-Cavity Surface Emitted Light (VCSEL) laser, micro-optics to shape and circularly polarize the laser beam, a mm-scale alkali (Cs or Rb) vapor cell in which CPT interaction takes place, a photodiode to detect transmitted light through the cell and packaged electronics including a local oscillator (LO) and various low-frequency servo loop electronics to stabilize experimental parameters. The microcell is currently filled with a buffer gas or mixture of buffer gases in order to operate in the Dicke regime [4]. A buffer gas mixture is often prefered to cancel the temperature-dependent collisional shift [5], [6] at a pre-determined inversion temperature [7]. However, the partial and total buffer gas pressure is very difficult to control in microfabricated cells because of high temperatures required to operate the anodic bonding process. In the frame of the challenging european MAC-TFC project, this paper reports a detailed characterization of CPT resonances measured in Cs-Ne microcells. These novel-technology cells are fabricated by following the technology process described in [8], [9], [10]. We recently reported the observation of a quadratic temperature dependence of the Cs 0-0 ground state hyperfine resonance frequency in such cells [11], [12]. The inversion temperature, expected to be independent of the buffer gas pressure, has been measured to be in the 77-80°C range. This paper aims to further prove the potential of our cells for microclocks applications.

Section II describes the experimental set-up. Section III is devoted to present the characterization of CPT resonances obtained in our Cs-Ne microcells. Section IV presents the dependence on laser intensity and cell temperature of the CPT resonance frequency. The last section V shows preliminary frequency stability measurements of a Cs-Ne microcell-based CPT clock.

II. EXPERIMENTAL SET-UP

Figure 1(a) describes the experimental set-up of the Cs CPT clock.

The Cs-Ne vapor cell is 2-mm in diameter and 1.4 mm-long. The microcell is filled with a Ne buffer gas pressure using an original two-step anodic bonding procedure recently described in [10]. The microcell temperature is stabilized to within 1 mK. A static magnetic field is applied around the cell in order to split the Zeeman components and isolate the $|F=3, mF=0\rangle \rightarrow |F=4, mF=0\rangle$ hyperfine clock transition. The ensemble is inserted into a cylindrical mu-metal magnetic shield in order to reduce spurious environmental magnetic fields.

The laser source is a VCSEL resonant with the Cs D2 line at 852 nm. The VCSEL injection current is directly modulated at 4.596 GHz (half of the Cs clock frequency) to produce two phase-coherent sidebands with a frequency splitting of 9.192 GHz. The 4.596 GHz signal is generated using a commercial microwave synthesizer driven by a 10 MHz quartz crystal local oscillator.

The linearly polarized output beam from the laser is collimated using a cylindrical lens in a 1-mm diameter beam and converted into circular polarization using a quarter-wave plate

978-1-61284-110-6/11/$26.00 ©2011 IEEE

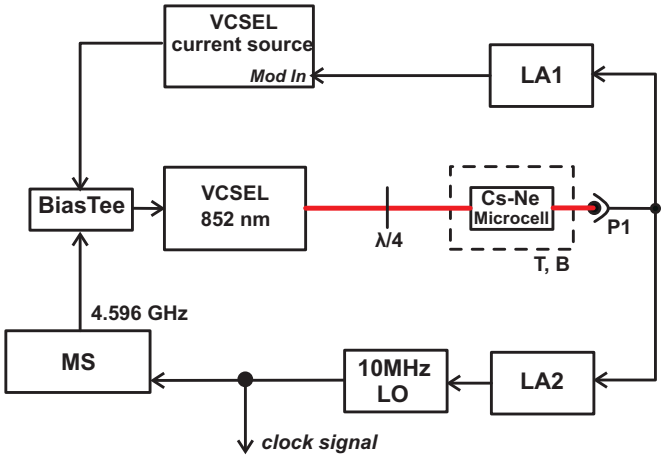


Fig. 1. *Experimental SetUp. Cs atoms with Ne buffer gas in the mm-scale cell interact with the microwave-modulated light from a VCSEL. The optical power transmitted through the cell is detected by the photodiode P1. The lockin amplifier LA1 is used to lock the VCSEL to the Cs optical resonance. The lockin amplifier LA2 is used to lock a quartz local oscillator (LO) driving a microwave synthesizer (MS) that generates the 4.596 GHz signal.*

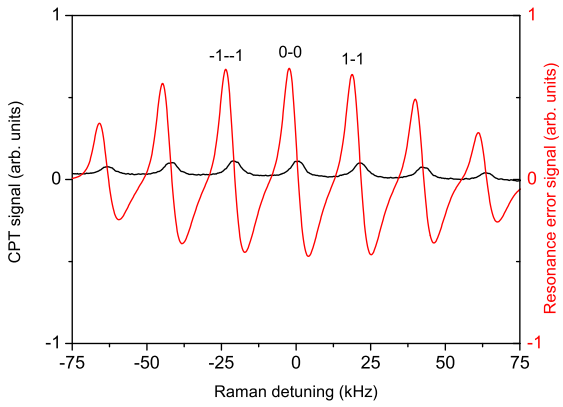


Fig. 2. *Zeeman spectrum obtained in a Cs-Ne microcell (dark line) and the corresponding error signal (red line) at the output of the lockin amplifier LA2. A longitudinal magnetic field of 2.9 μT flux density is applied.*

inserted in front of the cell. A neutral-density filter is used to adjust the laser intensity. The optical power transmitted through the Cs-Ne microcell is detected by a low noise Si photodiode (P1).

With the microwave modulation applied, the VCSEL frequency is locked on the position of maximum optical absorption by modulating the laser DC current at 65 kHz and demodulating the absorption signal on the photodiode with the lockin amplifier LA1. An error signal is then fed back to the VCSEL current driver to correct the laser injection current. The CPT resonance is detected by scanning the 4.596 GHz microwave signal frequency. The signal-to-noise (SNR) ratio of the resonance is increased by utilizing a similar frequency modulation-demodulation spectroscopy technique at 1-3 kHz with the lockin amplifier LA2. A voltage error signal can then be sent to the electric tuning port of the local quartz oscillator to lock the microwave synthesized signal on the atomic transition frequency.

III. CHARACTERIZATION OF CPT RESONANCES IN CS-NE MICROCELLS

In this section, we report the characterization of CPT resonances obtained in Cs-Ne microcells with various buffer gas pressures as a function of several experimental parameters such as the laser intensity I, the cell temperature T and the microwave power P_μ injected in the VCSEL.

Figure 2 shows a typical experimental Zeeman spectrum obtained in a Cs-Ne microcell as well as the corresponding error signal obtained at the direct output of the lockin amplifier LA2.

The 0-0 CPT resonance linewidth and height are extracted by fitting the CPT resonance error signal with the FM lineshape of a Lorentzian function [13]. The resonance contrast C is considered here as the ratio between the height of the resonance error signal and the background DC voltage at the direct output of the photodiode. The clock frequency stability can be measured by comparing the 10 MHz quartz local oscillator with a signal coming from a reference hydrogen maser. Frequency shifts measurements are performed by calculating the offset between the measured 0-0 resonance frequency and the unperturbed Cs atom frequency (9.192 631 770 GHz).

Cs-Ne microcells with different buffer gas pressures are tested. Cells 1, 2, 3 and 4 are filled with a Ne buffer pressure of 8.1, 6.7, 6.7 and 4.4 kPa respectively. All of them have the same dimensions. The static magnetic field B is 16 μT for all the following experiments.

A. Laser Intensity

Figure 3 shows the linewidth of the CPT resonance versus the total laser intensity for two different Cs-Ne microcells (cell 1 and cell 2).

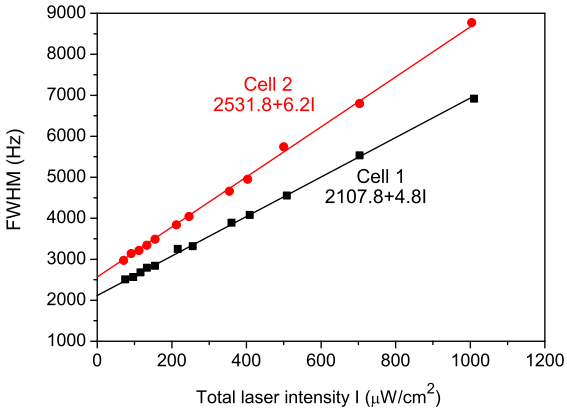


Fig. 3. *Linewidth of the dark resonance signal as a function of the total laser intensity for two Cs-Ne microcells with different Ne pressure (squares: cell 1, circles: cell 2). Solid lines are linear fit. Experimental parameters are: T=62.1°C, P_μ= 2.5 dBm.*

The CPT linewidth linearly increases with laser intensity

due to power broadening. The slope of the increase is lower for the cell with higher buffer gas pressure because of the increase of the optical linewidth $\Gamma^{\star}$. The zero-intensity linewidth is measured to be lower for the cell with higher buffer gas pressure. In this experimental set-up, the linewidth is limited by time-of-flight broadening because the laser beam is only 1-mm diameter [12]. The zero-intensity linewidth is expected to decrease with higher buffer gas pressures until the contribution of Cs-Ne collisions dominates. This only happens around 100 kPa in millimeter-scale cells.

We report in Fig. 4 the evolution of the ratio between the CPT resonance linewidth and contrast as a function of the total laser intensity for both cells 1 and 2.

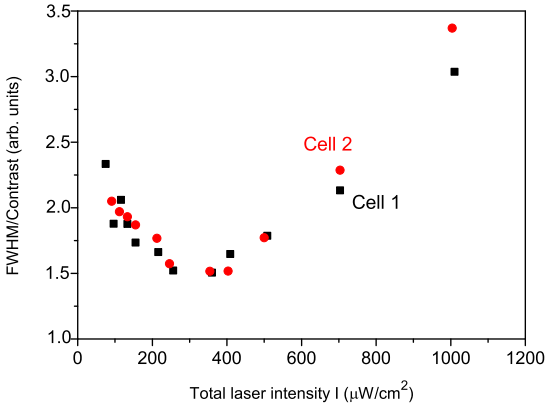


Fig. 4. *Ratio FWHM/contrast of the dark resonance as a function of the total laser intensity for two Cs-Ne microcells with different Ne pressure (squares: cell 1, circles: cell 2). Experimental parameters are: T=62.1°C, P_{μ}= 2.5 dBm.*

The ratio FWHM/contrast is minimized for both cells for intensities around 300-350 μW/cm^2. We observed that the noise on the detector increases strongly with increasing laser power W_i. This is due to the increase of the shot noise associated to the background signal, the increase of the laser AM noise that presents a $1/f$ slope at the modulation frequency but also due to the increase of the laser FM noise through a FM-AM conversion process in the vapor cell. In [14], it is claimed that the FM laser noise is the main limit to the short term frequency stability of a CSAC.

B. Cell Temperature

Fig. 5 shows the height of the dark resonance error signal versus the cell temperature for cells 1, 2 and 4.

For the three cells, we obtain a bell-shaped curve where the CPT resonance height is maximized in the 55-60°C range. A similar behavior has already been observed in cm-scale cells [15], [17], [16]. We see that the use of a higher buffer gas pressure gets the bell-shaped curve broader, reduces the CPT line height and shifts the temperature where the dark resonance height is maximum to higher temperatures. The signal is degraded by a factor 4-5 at 80°C compared to 60°C.

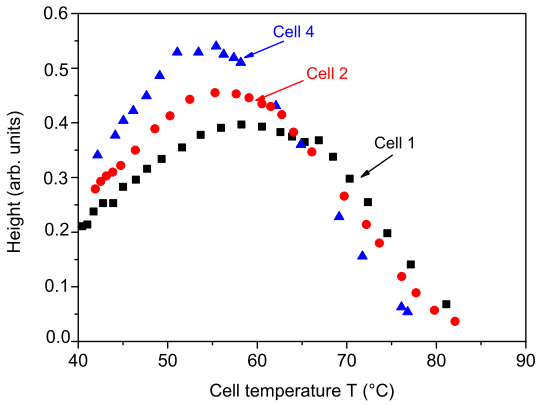


Fig. 5. *Height of the dark resonance error signal as a function of the vapor microcell temperature for three Cs-Ne microcells with different Ne pressure (squares: cell 1, circles: cell 2, triangles: cell 4). Experimental parameters are: I= 500 μW/cm^2, P_{μ}=2.5 dBm.*

It means that the clock short term frequency stability should be optimized at 60°C and degraded at 80°C.
The CPT linewidth is measured to decrease with higher buffer gas pressure and to slightly decrease with the cell temperature until about 80°C where the linewidth seems to rise poorly again.
We measured that the ratio FWHM/contrast is reduced when the buffer gas pressure is increased and is optimized in a temperature range between 58 and 66 °C for the three cells. The optimum temperature point increases when the buffer gas pressure is increased. Due to the reduction of the CPT signal height, this ratio is degraded around 78°C by a factor decreasing with higher buffer gas pressure. Note that no significant variation of the detected noise on the CPT resonance at the modulation frequency has been measured by changing the cell temperature.

IV. FREQUENCY SHIFTS

A. Light Shift

Light shift is known as a major limitation in the frequency stability of vapor cell frequency standards. We realized measurements of light shift in our Cs-Ne microcells. Figure 6 plots the frequency shift of the 0-0 resonance from the unperturbed Cs frequency as a function of the total laser intensity for three different microwave powers P_{μ} with the cell 2.

We show the possibility to cancel the first-order light shift by proper choice of the microwave power. For P_{μ} = 3.5 dBm, the light shift slope is measured to be only − 0.08 Hz/(μW/cm^2). In both other cases (2.5 dBm and 4 dBm), the slope is increased by a factor 6.1 but with opposite signs. Also, note that a slight frequency difference of 4.2 Hz between fitted curves extrapolated at zero-intensity is obtained validating our experimental measurements.

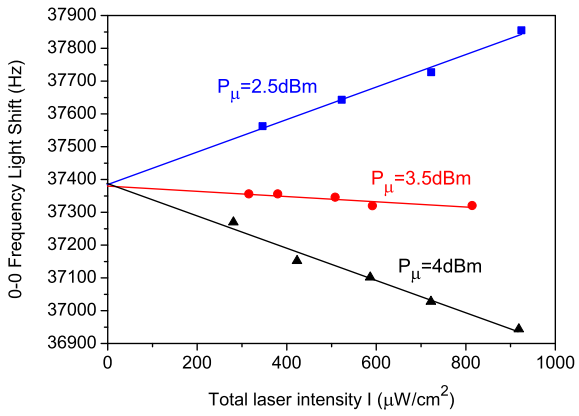


Fig. 6. *Light shift of the CPT resonance in the cell 2 for three different microwave power P_μ: 2.5 dBm (squares), 3.5 dBm (circles) and 4 dBm (triangles). Experimental parameters are: T=62°C. The solid lines are linear data fits.*

B. Collisional Frequency Shift

The 0-0 resonance frequency is measured for different temperatures for two Cs-Ne buffer gas microcells (cells 1 and 3) on Fig. 7. The measured frequencies are corrected for the Zeeman quadratic shift and for the light shift. The points are the experimental data. Data are fitted by a second order polynomial function.

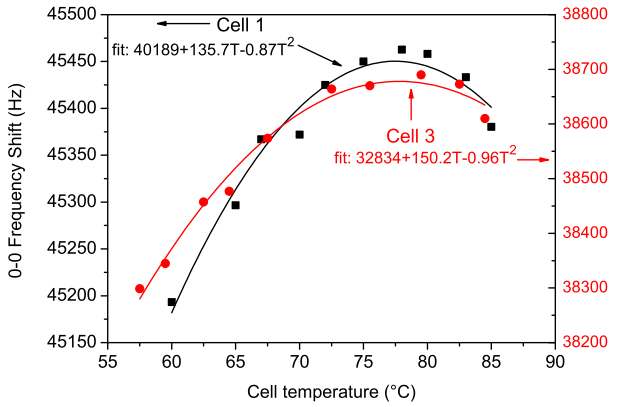


Fig. 7. *Frequency shift of the CPT resonance in cell 1 (squares) and cell 3 (circles) as a function of the cell temperature. The solid lines are data fits.*

The temperature dependence of the clock frequency is found to be cancelled for both cells around an inversion temperature of 78°C. These results agree well with those presented in [11] and confirm that the inversion temperature does not depend on the Ne buffer gas pressure. The fit of experimental points, with the gas pressure as fitting parameter and using coefficients given in [6], gives a Ne pressure of 8.1 kPa and 6.7 kPa for the cell 1 and cell 3 respectively at a temperature of 0°C. For the cell 1, the curvature is measured to be 1.7×10^{-10} K^{-2}. Assuming that the temperature control system allows to set the cell temperature T_c at $T_{inv} \pm 100$ mK with rms temperatures fluctuations dT =10mK around T_c, the clock frequency stability limit due to temperature variations is expected to be about 1.7×10^{-13}. These results are of great importance in view of the development of high long-term frequency stability miniature Cs atomic clocks with a unique Ne buffer gas cell. In this case, no suitable and well-controlled buffer gas mixture is required.

V. FREQUENCY STABILITY MEASUREMENTS

Fig. 8 shows preliminary measurements of the clock short term fractional frequency stability. For this purpose, the frequency stability of the 10 MHz output signal of the quartz local oscillator is measured using a digital 5120A Allan deviation test set [18] referenced by a hydrogen maser and plotted in Fig. 8.

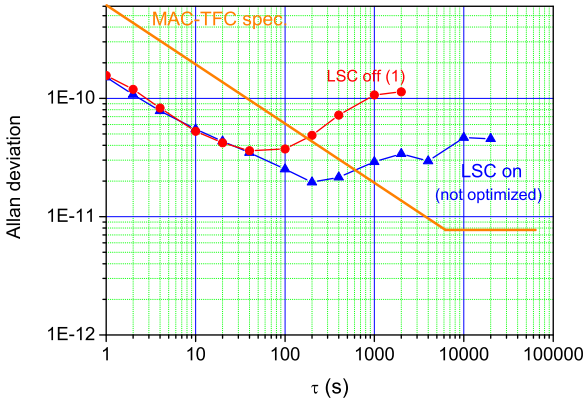


Fig. 8. *Allan standard deviation of the D_2 line VCSEL-based CPT clock using a Cs-Ne microcell. Experimental parameters are: T=60°C, P_μ=2.5 dBm. A measurement with or without light shift correction is shown.*

In a first measurement, the short-term frequency stability is given to be 1.5×10^{-10} $\tau^{-1/2}$ for averaging times between 1 and 30 s [12]. For short integration times, this result is comparable to those previously published [2] for miniature clocks using a VCSEL resonant with the Cs D_2 line and a mm-scale cell. The short-term frequency stability is expected to be significantly improved by the use of a laser resonant with the Cs D1 line [19] and was measured to be degraded by a factor 4 at 80°C. For integration times longer than 30 s, we observe a frequency random walk mainly attributed to huge temperature variations in the laboratory as well as laser intensity fluctuations.

Later, we started to investigate the possibility to implement a light shift correction (LSC) to improve the long-term frequency stability performances. Several advanced techniques, described in [20], [21], [22], were tested and allowed us to detect a zero-light shift point. In a first solution inspired from [20], the microwave power amplitude is modulated. This AM modulation is detected on the LA2 lockin amplifier output and can be minimized by adjusting the 4.596 GHz signal power.

The RF power for which the AM modulation vanishes is the zero-light shift point. This AM modulation can then be demodulated with a third lockin amplifier to generate a dispersive error signal crossing zero when the light shift is minimized and actively stabilize this point with an additional PI controller. We observed that this optimum point coincides well with the microwave power for which the CPT signal is maximized, i.e the point where the optical power contained in both first-order optical sidebands is maximum. In a second solution inspired from [22], the same lockin amplifier (LA2) is simultaneously used to lock the LO frequency to the clock transition and reduce the light shift. Indeed, it was observed that in presence of light shift, there are different zero-crossing frequencies for the in-phase and out-of-phase quadrature channels of the lock-in amplifier. When the in-phase channel is used to lock the local oscillator, the quadrature channel is proportional to the light shift near a zero-shift frequency. Note that both methods gave about the same optimum zero-light shift point. Figure 8 shows the improved frequency stability with LSC technique inspired from [20]. Despite the lack of optimization of the PI servo loop, the clock frequency stability is found to be 3×10^{-11} at 1000s and better than 5×10^{-11} at 20.000s. Further optimization remains to be done. Note also that the clock frequency should be better stabilized around 78°C where the temperature dependence of the clock frequency is greatly reduced.

VI. Conclusion

Cs-Ne microcells can be used for high performance miniature atomic clocks. For optimized clock frequency stability performances on the short term term, the clock should operate preferably with a VCSEL resonant on the Cs D1 line. For long term operation, the cell should operate at a temperature around 78°C to reduce greatly the Cs clock frequency temperature dependence. Moreover, advanced noise reduction techniques should be implemented to reduce frequency variations due to light shift effects. We preliminary investigated in this article a solution consisting in stabilizing the microwave power at a point where the light shift is cancelled.

Acknowledgment

This work is supported by a collaborative project MAC-TFC of the European Commission (FP 7, ICT program, grant n° 224132). The authors thank P. Abbe, C. Rocher and G. Michel for technical assistance.

References

[1] J. Vanier, "Atomic clocks based on coherent population trapping: a review", *Appl. Phys. B*, **81,**pp 421-442 (2005).

[2] S. Knappe, V. Shah, P. D. D. Schwindt, L. Hollberg, J. Kitching, L-A. Liew and J. Moreland, "A microfabricated atomic clock," *Appl. Phys. Lett.*, **85**, (6), pp. 1460–1462 (2004).

[3] R. Lutwak et al., "CSAC- The Chip Scale Atomic Clock", Proc. of the 7^{th} Symp. Freq. Standards and Metrology, edited by Lute Maleki (World scientific), pp. 454–462, (2008).

[4] R. H. Dicke, *Phys. Rev. Lett.*, **18**, pp. 472-475 (1953).

[5] F. Strumia, N. Beverini, A. Moretti and G. Rovera, "Optimization of the buffer gas mixture for opticaly pumped Cs frequency standard," *Proc. of the 1976 Freq. Contr. Symp.*, pp. 68–472 (1976).

[6] N. Beverini, F. Strumia and G. Rovera, "Buffer gas pressure shift in the m_F=0 m_F=0 ground state hyperfine line in Cs," *Opt. Comm.*, **37**, (6), pp. 394–396 (1981).

[7] J. Vanier, R. Kunski, N. Cyr, J. Y. Savard and M. Tetu, "On hyperfine frequuency shifts caused by buffer gases: Application to the optically pumped passive rubidium frequency standard," *Journ. Appl. Phys.*, **53**, (8), pp. 5387–5391 (1982).

[8] A. Douahi, L. Nieradko, J. C. Beugnot, J. Dziuban, H. Maillote, S. Guerandel, M. Moraja, C. Gorecki and V. Giordano, "Vapor Microcell for Chip Scale Atomic Frequency Standard," *Elec. Lett.*, **43**, (5), pp. 279–280 (2007).

[9] L. Nieradko et al., "New approach of fabrication and dispensing of micromachined cesium vapor cell," *Jour. Micro-Nano. MEMS and MOEMS*, **7**, (3), 033013 (2008).

[10] M. Hasegawa, R. K. Chutani, C. Gorecki, R. Boudot, P. Dziuban, V. Giordano, S. Clatot and L. Mauri "Microfabrication of cesium vapor cells with buffer gas for MEMS atomic clocks," *Sens. and Act. A: Phys.*, **167** , pp 594-601 (2011).

[11] D. Miletic, P. Dziuban, R. Boudot, M. Hasegawa, R. K. Chutani, G. Mileti, V. Giordano and C. Gorecki "Quadratic dependence on temperature of the Cs 0-0 hyperfine resonance in a single Ne buffer gas microfabricated vapor cell,"*Elec. Lett.*, **46,** (15), pp 1069-1071 (2010).

[12] R. Boudot, P. Dziuban, M. Hasegawa, R. K. Chutani, S. Galliou, V. Giordano and C. Gorecki, "Coherent Population Trapping resonances in Cs-Ne vapor microcells for miniature clocks applications", *Journ. Appl. Phys.*, **109,** 014912 (2011).

[13] G. C. Bjorklund, M. D. Levenson, W. Lenth and C. Ortiz, *Appl. Phys. B*, **32,** pp 145-152 (1983).

[14] R. Lutwak, D. Emmons, T. English, W. Riley, A. Duwel, M. Varghese, D. Serkland and G. Peake, *Proc. of the 34th Annual Precise Time and Time Interval Systems (PTTI 2003)*,San Diego, CA (2003).

[15] S. Knappe, R. Wynands, J. Kitching, H. G. Robinson and L. Hollberg *J. Opt. Soc. Am. B*, **18,** (11), pp 1545-1553 (2001).

[16] N. Castagna, R. Boudot, S. Guerandel, E. De Clercq, N. Dimarcq and A. Clairon, *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, **56,** (2), pp 246-253 (2009).

[17] S. Knappe, J. Kitching, L. Hollberg and R. Wynands, "Temperature dependence of coherent population trapping resonances", *Appl. Phys. B*, **74,** pp 217-222 (2002).

[18] http://www.symmetricom.com/

[19] S. Stähler, R. Wynands, S. Knappe, J. Kitching, L. Hollberg, A. Taichenachev and V. Yudin, "Coherent population trapping resonances in thermal ^{85}Rb vapor: D_1 versus D_2 line excitation", *Opt. Lett.*, **27,** (16), pp 1472-1474 (2002).

[20] V. Gerginov, S. Knappe, P. D. D. Schwindt, L. Hollberg and J. Kitching, "Atomic-based stabilization for laser-pumped atomic clocks", *Opt. Lett.*, **31,** (12) pp 1851-1853 (2006).

[21] V. Shah, V. Gerginov, P. D. D. Schwindt, S. Knappe, L. Hollberg and J. Kitching, "Continuous light shift correction in modulated CPT clocks", *Appl. Phys. Lett.*, **89,** 151124 (2006).

[22] B. H. McGuyer, Y.-Y. Jau and W. Happer, "Simple method of light-shift suppression in optical pumping systems", *Appl. Phys. Lett.*, **94,** 251110 (2009).

All-Optical Integrated Rubidium Atomic Clock

L. Maleki, A. A. Savchenkov, V. S. Ilchenko, W. Liang,
D. Eliyahu, A. B. Matsko, D. Seidel
OEwaves Inc., 2555 E. Colorado Blvd., Ste 400,
Pasadena, California, 91107 U.S.A.
Email: Andrey.Matsko@oewaves.com

N. P. Wells, J. C. Camparo, and B. Jaduszliwer
Electronics & Photonics Laboratory,
The Aerospace Corporation M2/253,
PO Box 92957, Los Angeles, CA 90009, U.S.A.
Email: Nathan.P.Wells@aero.org

***Abstract*— We report on progress in the development of an All-Optical Integrated Micro-Primary Atomic Clock (AO-IMPAC). The operattion of the clock is based on an inhomogeneously broadened rubidium optical transition used for stabilization of a Kerr optical frequency comb generated in a high-Q whispering gallery mode (WGM) microresonator. We have demonstrated the initial operation of this clock built in a physics package with dimensions $5 \times 6 \times 1.2$ cm. The clock's long term stability approaches 3×10^{-13} at about 2×10^4 s integration time. The measured value is limited by the noise floor of the measurement setup.**

I. Introduction

We introduced the concept and a preliminary design of an All-Optical Integrated Micro-Primary Atomic Clock (AO-IMPAC) at IFCS'2010 [1]. In this paper we report on progress towards implementing the design and evaluating the first demonstration of the clock operation. We have realized a bread-board model of this clock that is based on an optical frequency comb generated in a high-Q whispering gallery mode (WGM) microresonator, stabilized to an inhomogeneously broadened rubidium optical transition. The clock utilizes a low power single-mode DFB diode laser, self-injection locked to a mode of the microresonator. The optical frequency comb is generated in the resonator as a result of Kerr effect and four wave mixing process. The comb mode spacing effectively determines the RF output signal. Since the center frequency of a comb line and the comb's mode spacing both are determined by the refractive index and geometry of the microresonator, stabilizing the center frequency of a comb line stabilizes the clock frequency. In other words, locking a single frequency point of the comb is enough to stabilize the frequency spacing of the comb.

We have demonstrated the initial operation of this clock built in a $5\times6\times1.2$ cm physics package. The clock's long term stability approaches 3×10^{-13} at about 2×10^4 s integration time. The measured value is limited by the noise floor of the measurement setup locked to a commercial Rb clock. The maximum spectral width of the observed frequency comb was 4 nm, and the 24 GHz repetition frequency was generated with the vertical (not the basic) mode family of the WGM microresonator.

The comb was stabilized by locking it to the inhomogeneously broadened D1 line of Rb^{87} by means of a semiconductor laser. The lock quality and the frequency drift of the rubidium transition ultimately determine the performance of the system. The stability goal for this clock is to keep the time loss at the level of 32 ns over a period of one month. To achieve this timekeeping performance, the clock's random-walk frequency noise coefficient must be below 10^{-17} $s^{-1/2}$, and it must have a frequency aging rate below 10^{-15}/day. We found that a Doppler-broadening-induced shift mechanism is the most important parameter for stability, since it can lead to a temperature sensitivity of $4\times10^{-11}/^o$C, if the laser is locked to the peak of the absorption line. To reduce the temperature sensitivity by two orders of magnitude, we locked the laser to the so called isoclinic point of the D_1 line.

II. Operation principle and scheme of the clocks

Schematic of the clock setup is shown in Fig. (1). The goal of the set up is to stabilize the temperature of the resonator. Since the frequency drift of the hyper-parametric oscillation is primarily caused by temperature drift in the system, removing the temperature drift results in long term stabilization of the hyper-parametric oscillator.

Stabilization of the resonator temperature is realized via locking the frequency of a selected WGM to a Rb^{87} optical transition. We use the fact that the ratio of frequency of a high order optical mode and the free spectral range (FSR) of a monolithic WGM resonator does not depend on the resonator temperature or any resonator strain, to first approximation. The long term drift of the optical frequency is correlated with the drift of the FSR. Therefore, by stabilizing the optical frequency the FSR is also fixed.

The linewidth of the free running semiconductor distributed feedback (DFB) laser used in the experiment is on the order of a few megahertz. The bandwidth of the WGM is less than a hundred kHz. To pump the optical mode efficiently, we use the self-injection locking mechanism to lock the laser to the WGM, resulting is more than a thousand fold reduction of the laser linewidth. As a result, more than 30% of the power emitted by the laser is coupled to the mode.

It was shown that the frequency of hyper-parametric oscillation is determined by the FSR only, and does not depend on the parameters of the pumping laser light, if the optical absorption and corresponding heating of the resonator annulus is neglected. Thus drift of the laser frequency and the laser intensity noise do not influence the performance of the oscillator based on an ideal resonator. In a real system the light absorbed in the mode shifts the optical frequency and

978-1-61284-110-6/11/$26.00 ©2011 IEEE

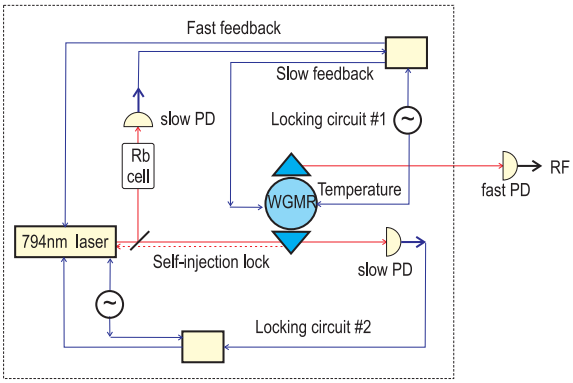


Fig. 1. Schematic of the demonstrated clock. A 794 nm semiconductor DFB laser is self injection-locked to a WGM. A slow PDH lock is used to keep the laser frequency at resonance with the WGM . The locked laser operates as an external cavity semiconductor laser, and can be tuned by tuning the frequency of the resonator. The laser/resonator combination is PDH locked to the ^{87}Rb D1 line. The locking scheme uses modulation of the frequency of the laser via temperature modulation of the WGM resonator, and the error signal is fed back to correct the temperature of the resonator. Since changing the laser current also changes the temperature of the resonator and the oscillation frequency, an additional lock loop is used for the laser.

the FSR frequency. Self-injection locking technique used to reduce the laser linewidth does not fix the value of detuning of the laser and the WGM. Thermal drift and drift of the current feeding the DFB laser result in drift of the detuning, and a corresponding drift of the WGM temperature due to change of the absorption of light in the resonator. Thus, we use a slow PDH-like locking circuit (locking circuit #2 in Fig. 1) to keep the laser frequency tuned to the specific position of the selected WGM.

Once the frequency of the laser follows the specific point of the WGM resonance, we lock the resonator to the atomic transition via the laser (locking circuit #1 in Fig. 1). This is accomplished by modulating the frequency of the optical mode thermally. The temperature modulation results in pure frequency modulation of the light emitted by the laser. The error signal is obtained by sending the modulated light through the atomic cell. The signal is subsequently processed and fed back to the resonator temperature. The modulation of the temperature is not the optimal solution since it also results in the modulation of the clock signal. Insertion of a power efficient phase modulator in front of the atomic cell, instead of modulation of the resonator temperature, would improve performance of the clocks.

The long term stability of the clock is determined by the thermal stability of the atomic transition. Any drift in the rubidium vapor temperature influences the stability of the D1 line in the Doppler-broadened optical absorption spectra, and compromises its frequency stability. We found that there exists a specific thermally insensitive point of the atomic resonance located in the vicinity of the local absorption minimum midway between the $5^2S_{1/2}(F_g = 2) \rightarrow 5^2P_{1/2}(F_e = 1, 2)$ absorption lines of ^{87}Rb vapor [2]. We verified experimentally that the relative frequency stability of the isoclinic point is better than -2×10^{-12} ^{o}K^{-1} while the stability of the point of the maximal absorption is on the order of 4×10^{-11} ^{o}K^{-1}. Therefore, the measured isoclinic point's temperature coefficient was an order of magnitude lower than the resonant transition. The thermal shift of the isoclinic point is probably even smaller since we could not find a difference between the isoclinic points temperature shift and zero at the 95% confidence level. Therefore, by keeping the averaged temperature of the cell stable with 10 mK accuracy the relative frequency drift of the system will be less than 2×10^{-14}.

III. Validation of Self-Injection Locking of a Laser with WGM Resonator

We performed an experiment in which we pumped a crystalline CaF_2 WGM resonator with 794 nm light emitted by a semiconductor laser that was self-injection locked [3] to a selected WGM. The light was sent in and retrieved out of the resonator using a glass coupling prism. The exiting light was collimated and injected into a single mode optical fiber to be further analyzed. The resonator was fabricated out of a z-cut CaF_2 wafer by mechanical polishing. The loaded quality (Q-) factor of the resonator exceeded 3×10^9.

We locked the semiconductor DFB laser to a WGM mode having a frequency in the vicinity of the isoclinic point of the D1 line of ^{87}Rb (Fig. 2). The self-injection locked range is a constant absorption level in the middle of the Rb absorption doublet. The absorption does not change since the frequency of the laser is constant and equal to the WGM frequency, while the current driving the laser changes (locking area). At the same time, the power of the light emitted by the laser decreases within the locking area since some amount of the light scatters back to the laser.

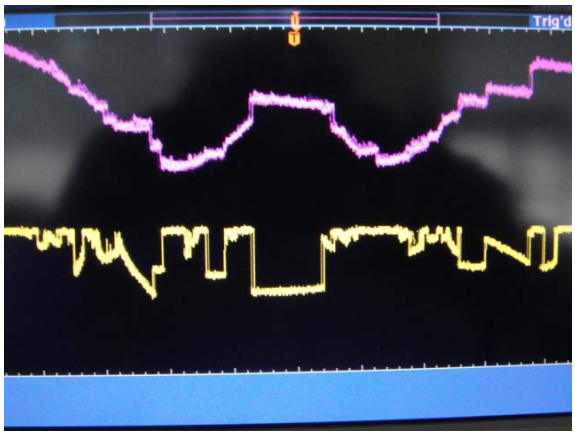


Fig. 2. Spectrum of the D1 transition of the rubidium (top) and the LI curve (optical output power (L) as a function of current that drives the laser (I)) of the DFB laser self-injection locked to the WGM resonator (bottom). The curves are shown on the oscilloscope screen.

IV. Demonstration of Stabilization of the Laser

Initially we demonstrated that the laser itself can be stabilized using the resonator. We then reduced the power of the laser such that the resonator operated in the linear regime, activated all the locking loops, and measured the frequency of

the laser during several hours. To lock the laser to the atomic transition, the laser frequency was modulated at the frequency of 0.23 Hz and a magnitude of about 100 MHz. Then the PDH-type lock scheme was implemented by feeding the signal from a lock-in amplifier back to the resonator temperature. To measure the stability of the laser we used another Rb cell as a frequency discriminator.

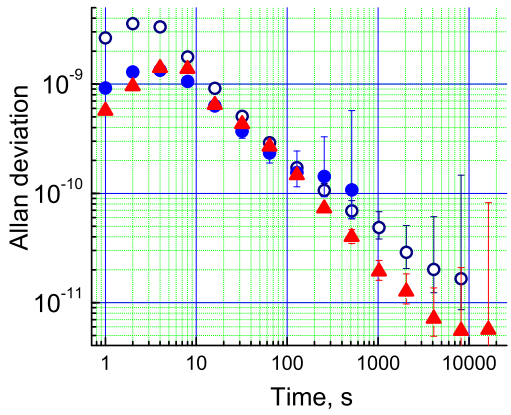


Fig. 3. Allan deviation of the frequency of the DFB laser locked to a WGM resonator and the rubidium cell. Several realizations are shown. All the realizations show averaging proportional to the integration time. Even though further averaging was expected we were unable to continue the measurements since the sensitivity floor of the measurement technique was reached.

V. Hyper-parametric oscillations at 794 nm

A nonlinear frequency conversion within the WGM resonator was used to make the system fully operate as a clock. Four-wave mixing (FWM) [4] and hyperparametric oscillation [5], [6] have been studied with WGM resonators with the generation of recently observed optical frequency (Kerr) combs [7]. Kerr combs have rapidly become the subject of extensive research as simpler alternatives to conventional combs produced with femtosecond mode-lock lasers. [7]-[14]. They have excellent uniformity as well as high repetition rate, which makes them ideal for many practical applications. We used such a comb for the implementation of the clock.

To generate the clock signal the power of the light pumping the resonator was increased so the optical frequency Kerr comb could be generated. The onset of hyperparametric oscillation was observed at 0.1 mW of optical power. Increasing the power to 0.3 mW resulted in the formation of a well pronounced optical frequency comb shown in Fig. (4). The optical spectrum of the comb demonstrates a clear asymmetry [15]. Depending on the position of the collimator collecting the light exiting the resonator, either higher or lower frequency components of the comb are nearly hidden in the noise. By changing the wavelength of the laser and locking it to different transverse modes of the resonator, we obtained the spectra shown in Fig. (4a and 4b). In some cases, the first optical sideband on the comb had higher power than the pump light exiting the resonator.

The observed optical comb spectra had several common features: i) they were asymmetric with respect to the pump frequency, i) the repetition frequency did not correspond to a multiple of the FSR of the resonator, contrary to earlier observations [8], [12]; ii) the envelope of the comb spectrum had unusually sharp edges. These features are novel, yet the comb lines are clearly self-phase locked since they produce a spectrally pure RF signal on a fast photodiode (Figs. 5 and 6), as in previously observed Kerr optical frequency combs [8].

A. Kerr comb on vertical modes

There are two nonlinear processes that result in the generation of a phase locked frequency comb in a WGM resonator possessing $\chi^{(3)}$ nonlinearity: hyper-parametric oscillation (modulational instability [16], [17], [18]) and FWM. The oscillation process involves two pump photons and two sideband photons and starts when the power of a continuous wave laser exciting the mode exceeds some threshold value [19], so that sidebands symmetrical in frequency are generated. Increasing the pump power results in the generation of multiple sidebands, phase locked due to the FWM process. The initial generation of sidebands widely separated by multiple FSR's could have a lower threshold compared with that of closely separated sidebands [8], [14], depending on the dispersion of the WGM spectrum.

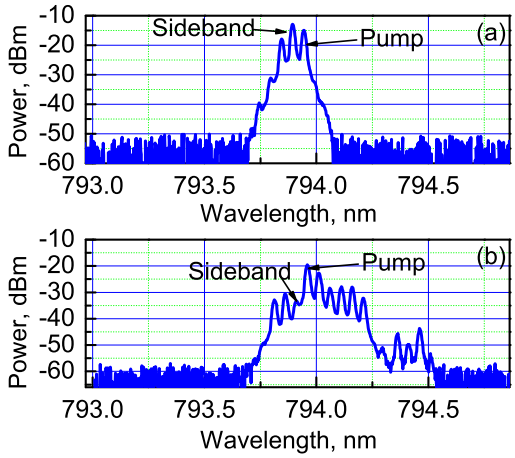


Fig. 4. Kerr optical frequency combs generated in the WGM resonator: (a) the power of the first generated sideband of the comb exceeds the power of the pump signal leaving the resonator; (b) the power of the same sideband is suppressed. To generate spectra (a) and (b) we pumped different WGMs of the same resonator. The spectra also are influenced by the position of the collimator collecting the light. No stimulated Raman scattering is observed. The resonator has 8 GHz free spectral range, while the optical frequency comb has 23.78 GHz repetition rate. The comb harmonics are self-phase locked, so they generate a coherent RF signals on a fast photodiode.

We found that i) the geometrical shape of the WGM resonator determines the frequency envelope as well as the repetition rate of the Kerr comb; ii) the observed components of the comb belong to different WGM families, i.e. have different spatial distribution; iii) the asymmetry of the spectrum

results from the significant spatial spread of the light emitted from the resonator and selectively coupled into the fiber. This novel type of Kerr comb could have, a higher repetition rate as compared with the FSR of the basic WGM sequence, as observed, and also a significantly lower repetition rate. The reason for such a flexibility is in the ability to modify the dispersion of the resonator with respect to the optical modes involved in the comb generation.

While it is generally believed that a Kerr comb is generated in the basic WGM family and has a relatively symmetric envelope, this is not the only option. A spheroidal resonator which has two nearly equidistant mode sequences characterized with FSR_l and $FSR_p = FSR_l(a-b)/b$, where b is the other semiaxis of the spheroid [20], may produce different results. Here, FSR_p can significantly differ from FSR_l depending on the ratio a/b. In our case $a/b \approx 4$, so that $FSR_p \approx 3FSR_l$, which is close to experimentally measured value $FSR_p/FSR_l = 23.78\ GHz/8\ GHz = 2.96$, where 23.78 GHz is the repetition frequency of the comb we observed. Therefore we conclude that the frequency components of the Kerr comb reported here are not generated in the same family of fundamental modes with different orbital momentum. Rather, the components belong to the family of modes with different vertical quantum number p and the same orbital momentum l.

B. The RF spectrum

We verified the spectral purity of the optical frequency comb and measured the repetition frequency using the RF signal produced by demodulation of the comb on a fast photodiode. The resolution of our optical spectrum analyzer did not exceed 8 GHz at 794 nm, so the optical spectrum was not suitable for measuremen. We increased the optical pump power, collected 2.5 mW of light leaving the resonator into a single mode fiber connected to a fast InGaAs photodiode. The photodiode generated several microWatt of 24 GHz RF signal that was analyzed using an RF spectrum analyzer. The recorded RF signal is shown in (Figs. 5 and 6).

The narrow linewidth RF spectrum generated by multiple comb lines suggests that the harmonics of the comb are self-phase locked in the sense of modulation-based combs [21].

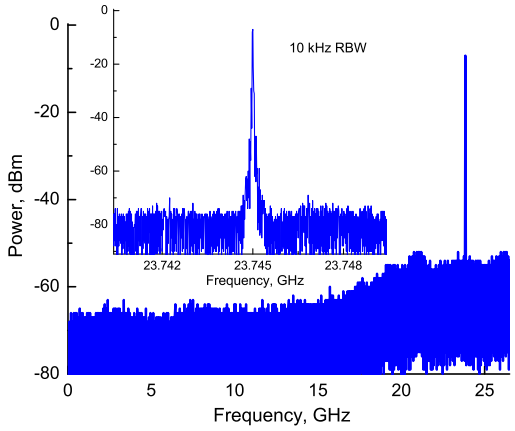


Fig. 5. The RF signal generated by demodulation of the optical frequency comb on a fast photodiode. Only one RF frequency component is generated.

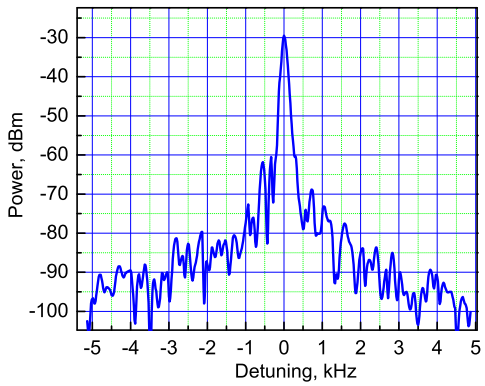


Fig. 6. The RF signal generated by demodulation of the optical frequency comb on a fast photodiode with lower resolution bandwidth than is shown in Fig. 5. Thermal drifts of the RF frequency limit the the resolution bandwidth of the measurement to 10 kHz. The instantaneous linewidth of the RF signal is 10 Hz, since the skirts of the RF signal can fitted with a Lorentzian curve of 10 Hz full width at half maximum.

VI. Locking the comb to the atomic transition

Activating the feedback loops shown in Fig. 1 locked the comb to the atomic transition, and the Allan deviation of the frequency of the locked signal could be measured with a frequency counter locked to a reference rubidium standard. The result of the measurement is shown in Fig. 7. The Allan deviation of the AO-IMPAC clock was averaging inversely proportional to the averaging time until it reached the sensitivity level of the measurement system determined by the Allan deviation of the reference Rb clock.

The observed behavior of the Allan deviation needs an explanation since it averages very fast as τ^{-1}, where τ is the integration time. When the Kerr comb runs freely the frequency drift behaves as $\tau^{1/2}$ (green line in Fig. 7). Such a drift results from the frequency random walk, so the phase noise spectrum is proportional to f^{-4} for $f \to 0$ (f is the spectral frequency). When the feedback is switched on, the phase noise should be divided by the product of the feedback gain and the PDH error function resulting from the lock to the atomic transition. If we assume that the error function goes as f^{-2} (the error-function for a Lorentzian) and the gain is constant at low frequencies (we found that the spectrum of the feedback gain is constant till 0.3 Hz), the phase noise in the closed loop will be proportional to f^{-2} for $f < 0.3$ Hz. This is the white frequency noise resulting in a $\tau^{-1/2}$ averaging of the Alan deviation signal. The expected behavior of the Allan deviation is illustrated by the black line in (Fig. 7). This is not what was observed with both the laser and the comb locked

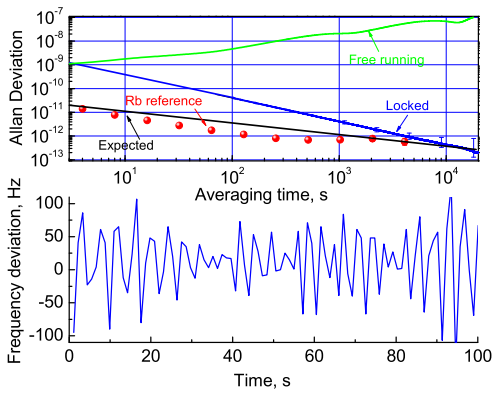


Fig. 7. Top: Allan deviation of the RF signal generated by the free running (green curve) and locked (blue curve) optical frequency comb; the Rb clock used in the measurement system (red dots); and expected averaging resulting from the PDH lock loop used in the experiment (black line). Bottom: Frequency signal from the counter.

to the atomic cell. The averaging apparently behaves as τ^{-1}.

The reason is that the resonator producing the comb is thermally modulated. The thermal modulation produces the frequency modulation of light needed to lock the laser and the resonator frequencies to the atomic transition. But, modulation of the resonator temperature changes the FSR of the resonator, as well, so the clock signal is also modulated with the same frequency (see the picture at the bottom of Fig. 7). The modulation frequency is approximately $f_1 = 0.238$ Hz. The averaged measured modulation magnitude of the 24 GHz signal is ± 70 Hz. The frequency counter sampling time is 1.1 s. Taking these values into account w the Allan deviation of a 24 GHz signal can be calculated, the frequency of which is modulated as $70\cos(2\pi f_1 t)$. The Allan deviation estimated in this way coincides with the Alan deviation measured in our experiment, which behaves as τ^{-1}. Hence, the measured result is determined by the imperfectness of the measurement system we used. The actual performance of the clock is likely better than what is measured.

VII. Conclusion

We have shown that a Kerr frequency comb excited in a WGM resonator can be used in the realization of a new type of optical atomic clock. The clock features an architecture using an optical transition in an atom or molecule, making it suitable for miniaturization with small power for operation. This is the goal of the AO-IMPAC clock, which will be further developed in the next phases of the DARPA IMPACT program.

Acknowledgment

The reported work was supported and approved for public release by DARPA MTO, distribution unlimited. The views, opinions, and findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

References

[1] L. Maleki, V. S. Ilchenko, M. Mohageg, A. B. Matsko, A. A. Savchenkov, D. Seidel, N. P. Wells, J. C. Camparo, and B. Jaduszliwer, "All-optical integrated atomic clock," Proc. IEEE IFCS'11, 119-124 (2010).

[2] N. P. Wells and J. C. Camparo, "^{87}Rb D1 isoclinic point," Phys. Rev. A 82, 062505 (2010).

[3] W. Liang, V. S. Ilchenko, A. A. Savchenkov, A. B. Matsko, D. Seidel, and L. Maleki, "Ultra-narrow linewidth external cavity semiconductor lasers using crystalline whispering gallery mode resonators," to be published in Optics Letters (2010).

[4] S. M. Spillane, T. J. Kippenberg, and K. J. Vahala, "Ultralow-threshold Raman laser using a spherical dielectric microcavity", Nature (London) **415**, 621-623 (2002).

[5] T. J. Kippenberg, S. M. Spillane, and K. J. Vahala, "Kerr-nonlinearity optical parametric oscillation in an ultrahigh-Q toroid microcavity," Phys. Rev. Lett. **93**, 083904 (2004).

[6] A. A. Savchenkov, A. B. Matsko, D. Strekalov, M. Mohageg, V. S. Ilchenko, and L. Maleki, "Low threshold optical oscillations in a whispering gallery mode CaF_2 resonator," Phys. Rev. Lett. **93**, 243905 (2004).

[7] P. Del'Haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, and T. J. Kippenberg, "Optical frequency comb generation from a monolithic microresonator," Nature (London) **450**, 1214-1217 (2007).

[8] A. A. Savchenkov, A. B. Matsko, V. S. Ilchenko, I. Solomatine, D. Seidel, and L. Maleki, "Tunable Optical Frequency Comb with a Crystalline Whispering Gallery Mode Resonator," Phys. Rev. Lett. **101**, 093902 (2008).

[9] I. S. Grudinin, N. Yu, and L. Maleki, "Generation of optical frequency combs with a CaF_2 resonator," Opt. Lett. **34**, 878-880 (2009).

[10] D. Braje, L. Hollberg, and S. Diddams, "Brillouin-enhanced hyperparametric generation of an optical frequency comb in a monolithic highly nonlinear fiber cavity pumped by a cw laser," Phys. Rev. Lett. **102**, 193902 (2009).

[11] P. Del'Haye, T. Herr, E. Gavartin, R. Holzwarth, and T. J. Kippenberg, "Octave spanning frequency comb on a chip," arXiv:0912.4890v1 [physics.optics] 24 Dec 2009.

[12] O. Arcizet, A. Schliesser, P. DelHaye, R. Holzwarth, and T. J. Kippenberg, "Optical frequency comb generation in monolithic microresonators," in *Practical Applications of Microresonators in Optics and Photonics*, A. B. Matsko, ed. (CRC Press, 2009), Chap. 11.

[13] A. B. Matsko, A. A. Savchenkov, W. Liang, V. S. Ilchenko, D. Seidel, and L. Maleki, "Whispering gallery mode oscillators and optical comb generators," Proc. of 7^{th} Symp. Frequency Standards and Metrology, ed. L. Maeki, pp. 539-558 (World Scientific, New Jersey, 2009).

[14] Y. K. Chembo, D. V. Strekalov, and N. Yu, "Spectrum and dynamics of optical frequency combs generated with monolithic whispering gallery mode resonators," Phys. Rev. Lett. **104**, 103902 (2010).

[15] A. A. Savchenkov, A. B. Matsko, W. Liang, V. S. Ilchenko, D. Seidel, and L. Maleki, "Kerr combs with selectable central frequency," Nature Photonics **5**, 293296 (2011).

[16] M. Haelterman, S. Trillo, and S. Wabnitz, "Additive-modulation-instability ring laser in the normal dispersion regime of a fiber," Opt. Lett. **17**, 745-747 (1992)

[17] S. Coen and M. Haelterman, "Modulational instability induced by cavity boundary conditions in a normally dispersive optical fiber," Phys. Rev. Lett. **79**, 4139-4142 (1997).

[18] S. Coen and M. Haelterman, "Continuous-wave ultrahigh-repetition-rate pulse-train generation through modulational instability in a passive fiber cavity," Opt. Lett. **26**, 39-41 (2001).

[19] A. B. Matsko, A. A. Savchenkov, D. Strekalov, V. S. Ilchenko, and L. Maleki, "Optical hyper-parametric oscillations in a whispering gallery mode resonator: threshold and phase diffusion," Phys. Rev. A **71**, 033804 (2005).

[20] M. L. Gorodetsky, A. E. Fomin, "Geometrical theory of whispering gallery modes," IEEE J. Sel. Top. Quantum Electron. **12**, 33-39 (2006).

[21] M. Kourogi, K. Nakagawa, and M. Ohtsu, "Wide span optical frequency comb generator for accurate optical frequency difference measurement," IEEE J. Quantum Electronics **29**, 2693-2701 (1993).

Low-Power Chip-Scale Rubidium Plasma Light Source for Miniature Atomic Clocks

Vinu Venkatraman, Herbert Shea
Microsystems for Space Technologies Laboratory
École Polytechnique Fédérale de Lausanne (EPFL)
Neuchâtel, Switzerland
vinu.venkat@epfl.ch

Yves Petremand, Nico de Rooij
Sensors, Actuators and Microsystems Laboratory
École Polytechnique Fédérale de Lausanne (EPFL)
Neuchâtel, Switzerland

Christoph Affolderbach, Gaetano Mileti
Laboratoire Temps-Fréquence
University of Neuchâtel
Neuchâtel, Switzerland

***Abstract*— We present the development, testing and characterization of a low-power chip-scale Rubidium (Rb) plasma light source designed to serve for optical pumping in miniature atomic clocks. The technique used is electrodeless capacitively coupled plasma (CCP) discharge, driven in a micro-fabricated Rb vapor cell. The device is electrically driven at frequencies between 1 and 36 MHz to emit 140 µW of stable optical power while coupling < 6 mW of electrical power to the discharge cell. To our knowledge this is the first reported Rb plasma emitted from a chip-scale device.**

I. INTRODUCTION

Compact Rb atomic clocks (overall volumes of 100-1000 cm^3), are widely used in many applications including navigation systems, network synchronizations and digital communications [1]. A common commercial implementation of such an atomic clock is the compact (~100 cm^3) Rubidium microwave frequency standard based on the Double-Resonance (DR) technique, using a glass-blown Rb plasma discharge lamp to achieve optical pumping. With the advent of microtechnology there has been an upsurge in portable devices (GPS receivers, for example) leading to a growing need for smaller, lower-cost, chip-scale atomic clocks with reduced power consumption, that show better frequency stability than the currently used quartz crystal oscillators. This has led to research towards miniaturizing atomic clocks to reach volumes of few cm^3 or below [2, 3], mostly based on the Coherent Population Trapping (CPT) approach that has been preferred over the traditional DR technique [4].

The majority of Rb DR clocks use Rb plasma discharge lamps as the light source for optical pumping due to the lamp's intrinsically correct frequency (Rb lines), technical simplicity, very slow frequency drift (aging), and long lifetime. However, because these glass-blown lamps usually consume several watts of electrical power and are not easily integrable with an ideal chip-stack configuration, chip-scale Rb DR clocks had been deemed a challenge and are relatively unexplored. The CPT technique was preferred for chip-scale clocks by using a laser diode (VCSEL) as the light source for atomic resonance [2, 3] because of their easy integration, good power efficiency, planarity and compact dimensions (~mm^2). However, the VCSEL approach requires critical temperature control, sophisticated control electronics for laser frequency stabilization, and complicated setup procedures [2].

The goal of this research is to develop a microfabricated planar low-power Rb plasma light source (as shown in operation in Fig. 1) to exploit the inherent advantages of a Rb discharge lamp, avoiding the issues faced by the VCSEL and providing an alternative and innovative design for chip-scale Rb atomic clocks.

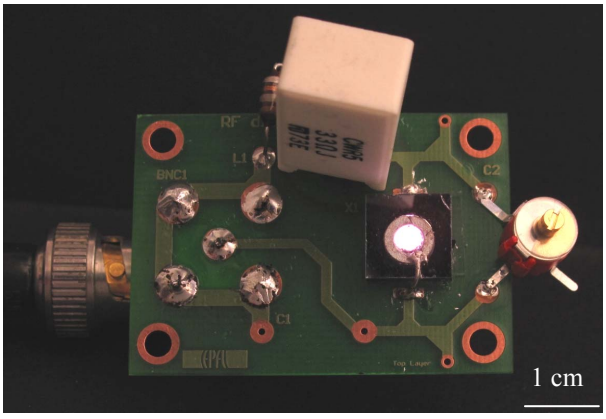


Figure 1. Photographic view of the Rb plasma light source emitting the Rb glow discharge

This project was funded by Swiss National Science Foundation *(SNSF)* Sinergia grant CRSI20_122693.

Microfabrication of hermetically sealed Rb cells with volumes on the order of 1 mm^3 has been a challenge, with several groups reporting different bonding techniques compatible with Rb filling, for instance anodic bonding [3, 5] or soldering [6, 7]. Such cells have been used as the reference cell in miniaturized atomic frequency standards (atomic clocks), but not as light sources – in part due to the additional complexity of igniting a stable plasma. While first studies on micro-cell plasma light sources reported on low-power rare-gas discharges, no Rb light was reported [8, 9].

A. *Double Resonance Rb Atomic Clock*

The fundamental concept behind timekeeping in atomic clocks is to exploit an atomic resonance as reference for stabilizing the frequency of a quartz oscillator. In a Rb atomic clock [6], this resonance is the "clock" transition between the two ground-state hyperfine levels of ^{87}Rb, at a microwave frequency near 6.83 GHz ($5^2S_{1/2}$ state, $|F=2, m_F=0\rangle$ to $|F=1, m_F=0\rangle$), detected on atoms maintained in a small Rb reference cell. In a double-resonance Rb clock, pump light from a lamp or laser resonant with the Rb D1 or D2 line (794.8 nm or 780.2 nm, respectively) irradiates the cell and creates a ground-state polarization in the atoms. The clock transition is then detected by applying a microwave field to the atoms via a microwave-cavity placed around the cell, and the resonance manifests itself as a narrow dip in pump-light power transmitted through the cell. The frequency of a voltage-controlled quartz oscillator is locked to this dip using phase-sensitive detection.

Our envisaged double-resonance Rb atomic clock consists of a micro-fabricated Rb plasma light source, a micro-fabricated Rb^{85} filter, a micro-fabricated Rb reference cell which is enclosed by a miniaturized microwave cavity, and a photodetector (Fig. 2). The Rb plasma light source intrinsically emits at the correct wavelengths, Rb D1 and D2 lines, as required in a Rb clock.

B. *Rb Plasma Glow Discharge*

Sustaining a plasma glow discharge in a Rb cell requires continuous ionization of Rb vapor atoms. A sufficiently strong electric field, when applied across a Rb vapor cell, can ionize the Rb vapor atoms to generate Rb ions and free electrons.

By maintaining a sufficiently strong electric field, the electrons and ions are accelerated towards the electrodes, and many of them collide with more neutral vapor atoms to generate positive ions and electrons through impact ionization before they reach the electrodes. This becomes an avalanche process, eventually initiating an electrical breakdown leading to several possible types of gas discharges [10], including the Rb plasma glow discharge emitted as radiation by Rb atoms during their radioactive decay from the excited state to the ground state.

The Rb vapor density is very low at room temperature (~10^{-6} millibars of vapor pressure at 30 ºC), hence requiring an extremely high field strength to ionize the Rb atoms (Paschen curve) [11]. In order to reduce the field and hence the electrical power to ignite the glow discharge, an inert low-pressure starter gas is added to the cell. The temperature of the cell is also increased to up to 150°C by using a resistive heater, in order to increase the Rb vapor density in the cell. As Rb is highly reactive with many gases, a few ten millibars of an inert gas are used as starter gas – in our case Ar, which has a relatively smaller diameter and high mean free path among noble gases with an ionization potential of 15.75 eV [10].

II. Device Microfabrication and Development

A. *Design and Microfabrication*

The design approach chosen for the Rb plasma light source is the electrodeless capacitively coupled plasma (CCP), sustained in a planar micro-fabricated Rb cell. Capacitive coupling produces stronger field lines in a planar cell cavity and therefore makes it easier to ignite a plasma than the more widely used inductive coupling which is more efficient for spherical cells.

The micro-fabricated Rb cell (outer dimensions of 1 x 1 x 0.3 cm^3), consists of a stack of three layers: Pyrex (500 µm thick), Silicon (2 mm), Pyrex (500 µm) which enclose a Φ = 5 mm, h = 2 mm cylindrical cavity in which few microliters of Rubidium and 70 millibars of Argon are hermetically sealed by a two-step anodic bonding process (Fig. 3). First, the cavity is created by DRIE etching a through-hole in the silicon wafer. Then, the etched silicon wafer is anodically bonded to the bottom Pyrex layer.

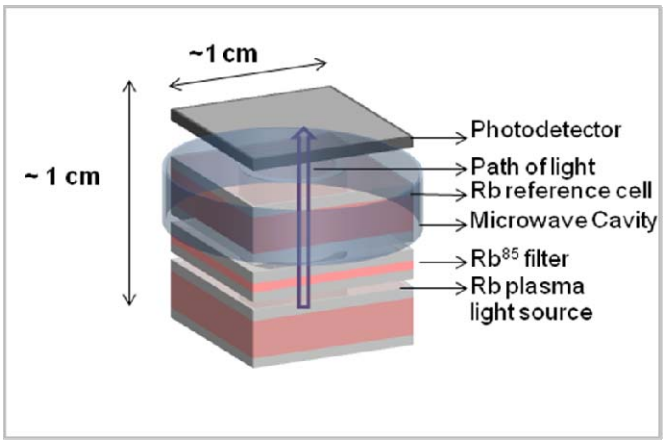


Figure 2. Schematic diagram of the envisioned Rb DR chip-scale atomic clock with the integrated Rb light source

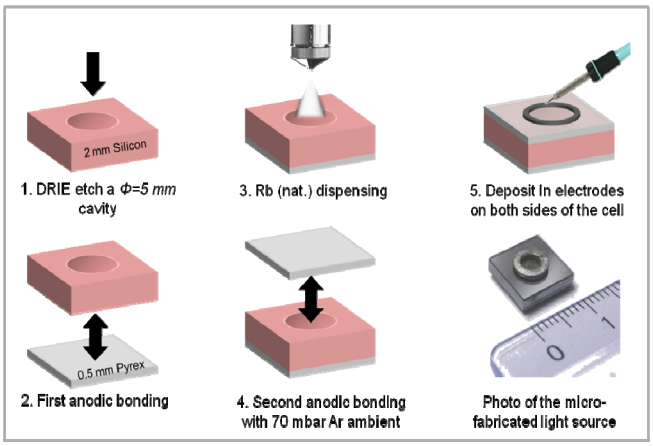


Figure 3. Schematic diagram showing the Rb light source microfabrication process flow

After this first bonding step, the natural Rb is dispensed into the cavity under vacuum by an atomic jet, 70 millibars of Argon are introduced in the chamber, and the top Pyrex layer is anodically bonded to the bottom layers to create the hermetically sealed Rb mini-cell [12].

A 200 µm thick 100% Indium solder ring or film is deposited on top of both sides of the Pyrex layers (i.e., outside the cell) to act as capacitive parallel electrodes to capacitively couple RF power to the cell. 100% Indium wets glass well and is hence a good choice for deposition on the Pyrex walls. The 100% Indium electrodes are deposited by placing and aligning Indium pre-form rings on top of Pyrex layers above the cavity and reflowing using a soldering iron. These external electrodes avoid electrode corrosion and might allow for lower power consumption when compared to the more traditional internal electrode designs [13]. Note that in a future wafer-level production of the plasma light source, the electrodes can be directly integrated onto the Pyrex layers used to close the cell by evaporation or sputtering and micropatterning techniques, and could be made of a suitable metal other than Indium in order to better withstand high temperatures. Our micro discharge cell design is thus compatible with wafer level manufacturing of many devices at a time, potentially allowing for greatly reduced costs per unit.

B. *RF Capacitively Coupled Plasma (CCP) Circuit*

The technique used to power the light source is RF electrodeless CCP. The microfabricated Rb vapor cell has an intrinsic capacitance of ~0.4 pF at low frequencies (<10 MHz) and up to ~1.6 pF at higher frequencies (<36 MHz), and a reactance of ~j85 kΩ at low frequencies (few MHz).

The output impedance of the RF generator used as signal source is (50 + j·0 Ω) hence requiring careful impedance matching of the load to equal the source to (50 + j·0 Ω) using L-C components. This enables high power and voltage amplification across the cell to achieve the required breakdown voltage with the lowest input electrical power.

Fig. 4 shows schematically the RF plasma ignition circuit connected to the impedance-matched load constituting the light source. The two main components in the impedance matching network are (1) a capacitor in parallel to the cell to reduce the overall capacitive reactance and (2) an inductor in series with the above capacitors to cancel out the reactive load to form a purely resistive load. The values and the Equivalent Series Resistance (ESR) of the components are carefully chosen to maximize the power transfer to the cell.

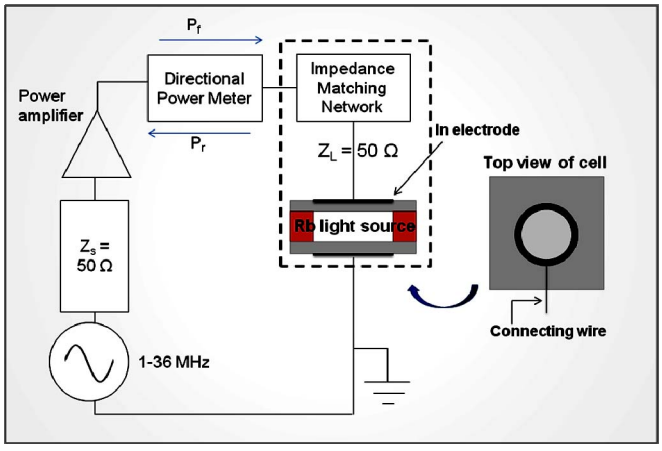


Figure 4. Schematic circuit of the RF electronic circuitry and impedance matching for plasma ignition

III. Device Characterization

The light source was impedance matched and tested at several frequencies over the 1 MHz to 36 MHz frequency range. The results reported in this article were obtained at 4.6 MHz.

A. *Output optical characteristics*

The output optical spectrum of the light source was measured using a spectrometer and is shown in Fig. 5. At cell temperatures above ~90 ºC strong Rb D1 lines (794.8 nm) and D2 lines (780.2 nm) are observed. These lines rapidly rise in intensity with increasing cell temperature, as expected with the rapidly rising Rb vapor pressure inside the cell.

The total optical power of the light source was 140 µW when 6 mW of electrical power was coupled to the cell. The total optical power emitted on the Rb D2 line was 15 µW, at a cell temperature of 150 ºC.

B. *Rb light power stability*

The stability of the output optical power of the Rb lines was recorded over time for the plasma light source operating at a temperature of 100°C. The short-term (seconds) and long term (hours) variation in optical power was measured.

The average short-term fluctuations in optical intensity of the Rb lines was found to be less than 0.1% and the long-term variations were found to be less than 5% (shown in Fig. 6). The primary cause of the long-term variations in optical power is the drift in resistance (2-3%) of the electrical components of the impedance matching network with time, mainly due to thermal effects. These long-term variations can be limited to less than 0.01% using a PID feedback loop that controls the input electrical power and/or the RF drive frequency to maintain a steady output light power level.

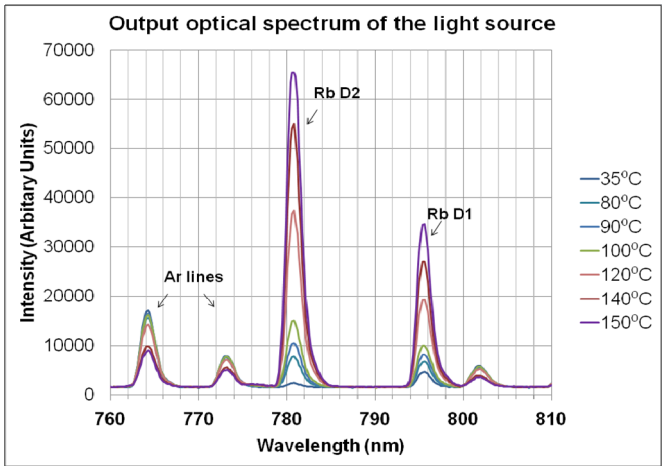


Figure 5. Optical spectrum of the light source at different temperatures between (35-150 ºC). The spectral width of the lines is limited by the resolution of the spectrometer used here.

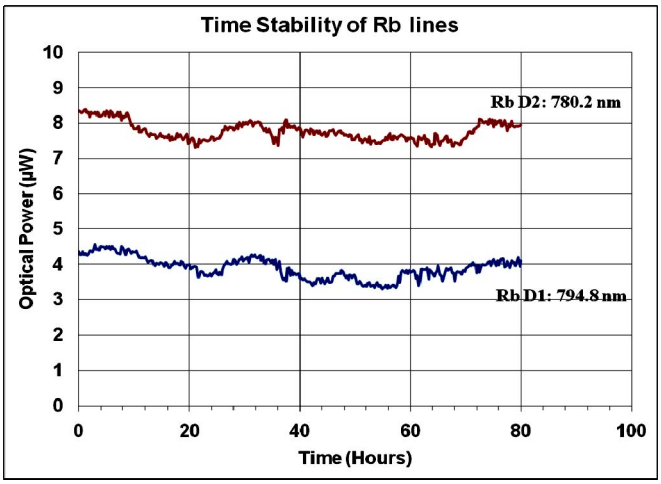


Figure 6. Stability of the Rb lines with time, measured at 100 ^{0}C

IV. Summary and Future Work

A stable and low-power chip-scale Rb plasma light source with potential for wafer-scale fabrication has been demonstrated, intended for use in chip-scale atomic clocks. The light source emits several μW of optical power at the 780.2 and 794.8 nm Rb lines from a 0.3 cm^3 chip stack. A successful Rb hermetic sealing technique through standard microfabrication processes has been demonstrated that withstands the plasma operation. The external electrode design ensures the longevity of the device by avoiding electrode erosion by the plasma.

When combined with a micro-fabricated ^{85}Rb filter cell [9], the presented plasma light source can be used for hyperfine optical pumping of Rb in a novel miniature atomic clock.

Testing of this light source as an optical pump in a test DR atomic clock setup to prove its functionality is planned. In the current circuit design, the efficiency of power coupled to the cell is low, thus current research is geared towards increasing this efficiency. Different starter gases and pressures are to be explored to find the best performance conditions.

Acknowledgment

We thank our project group members for their support. We acknowledge the SNSF R'Equip program for the funding of the High speed data acquisition test bench used in this study.

References

[1] J. Camparo, "The rubidium atomic clock and basic research", Phys. Today, vol. 60, pp. 33-39, 2007

[2] S. Knappe, "MEMS atomic clocks", Comprehensive Microsystems, vol. 3, pp. 571-612, 2007.

[3] A. Douahi, L. Nieradko, J. C. Beugnot, J. Dziuban, H. Maillote, S. Guerandel, M. Moraja, C. Gorecki and V. Giordano : "Vapour microcell for chip scale atomic frequency standard", Electronics Letters, vol. 43, pp. 279-280, 2007.

[4] A. M. Braun, et al., "RF-interrogated end-state chip-scale atomic clock", Proc. 39th Annual Precise Time and Time Interval Meeting (PTTI), USA, pp.233 – 247, 2007.

[5] S. Knappe, V. Shah, P. Schwindt, L. Holberg, J. Kitching, L. Liew and J. Moreland, "A microfabricated atomic clock", Appl. Phys. Lett., vol. 85, pp. 1460-1462, 2004.

[6] F. Vecchio, V. Venkatraman, H. Shea, T. Maedar and P. Ryser, "Dispensing and hermetic sealing Rb in a miniature reference cell", Sensors & Actuators: A. Physical, in press.

[7] Y. Pétremand, C. Schori, R. Straessle, G. Mileti, N. de Rooij and P. Thomann, "Low temperature indium-based sealing of microfabricated alkalicells for chip scale atomic clocks", Proc. of the European Frequency and Time Forum (EFTF), Noordwijk, the Netherlands, p.119, 2010.

[8] H. Guo and A. Lal, "Self-powered photon source", IEEE Proc. of the 12th Int. Conference on Solid State Sensors, Actuators and Microsystems, Transducers 2003, U.S.A, pp. 1474-1477, 2003.

[9] S. Wang, L. Lin and H. Guo, "Analysis and design of a micromachined Rb-85 filter in passive rubidium atomic clock", Proc. IEEE Int. Conference on Nano/Micro Engineered and Molecular Systems, China, pp. 911-914, 2009.

[10] E. Nasser, "Fundamentals of gaseous ionization and plasma electronics", Wiley-Interscience, 1971.

[11] F. Paschen, "Über die zum funkenübergang in luft, wasserstoff und kohlensäure bei verschiedenen drücken erforderliche potentialdifferenz", Weid. Annalen der Physik, vol.37, pp. 69-75, 1889.

[12] J. DiFrancesco, F. Gruet, C. Schori, C. Affolderbach, R. Matthey, G. Mileti, Y. Salvadé, Y. Petremand and N. De Rooij, "Evaluation of the frequency stability of a VCSEL locked to a micro-fabricated rubidium vapour cell", Proc. SPIE, vol. 7720, pp. 77201T, 2010.

[13] D. O. Wharmby, "Electrodeless lamps for lighting: a review", IEE-Proceedings, vol. 140, pp. 465-473, 1993.

MOT Loading Enhancement with Stimulated Light Forces

T.C. Liebisch, E. Blanshan, E.A. Donley, and J. Kitching
NIST Time and Frequency Division
325 Broadway
Boulder, Colorado 80305
email: edonley@boulder.nist.gov

***Abstract*—We demonstrate atom number enhancement in a magneto-optical trap (MOT) by use of bichromatic cooling to slow an atomic beam that is loaded into a MOT. Bichromatic cooling employs stimulated emission to apply strong cooling forces that are not limited by spontaneous emission. We demonstrate a factor of 3.5 increase in atom number captured from the atomic beam for a 1.5 cm cooling length. For a 1.5 cm effective cooling length, our technique yields MOT atom number enhancement that is about three times higher than the enhancement acheived via spontaneous emission.**

I. Introduction

While atom cooling and trapping have revolutionized atomic clocks and precision measurements, cold atoms are not yet used in commonplace technologies. The ability to create widely disseminated, high-performance compact atomic clocks and sensors based on cold atoms could revolutionize portable frequency standards. A fundamental difficulty in developing compact cold-atom devices is that the number of captured atoms in a conventional magneto-optical trap (MOT) scales as d^4, where d is the laser beam diameter [1]. Experiments performed with microfabricated pyramidal traps have demonstrated an $N \propto d^6$ scaling law, showing that the scaling can be even worse for small systems [2].

The strong dependence of the atom number on beam size is fundamentally set by the maximum spontaneous light force achievable in a MOT, $F_{spont} = \hbar k\gamma/2$ [3], where $\hbar k$ is the photon momentum, γ is the resonance linewidth, and $\gamma/2$ is the maximum spontaneous photon scatter rate. To overcome this limitation, we have employed stimulated emission forces to slow a ^{87}Rb atom beam and increase the atom number density at low velocities. In principle, the stimulated forces are limited only by the available laser power.

The stimulated forces are created by a bichromatic cooling technique. Experiments demonstrating use of the bichromatic force to slow an atomic beam were first done by Söding et al., who slowed a cesium beam over a distance of 10 cm with total laser power exceeding 100 mW [4]. Bichromatic cooling has also been studied for Rb atoms [5] and metastable He atoms [6]. A thorough review of laser cooling and trapping techniques that includes a review of bichromatic cooling can be found in Ref. [3]. To our knowledge, our experiment is the first to use the technique for MOT atom number enhancement.

The concept of bichromatic cooling is presented in Fig. 1, where a sample atom is moving with a velocity v_{atom} to the right. A light beam with two frequency components, $\omega = \omega_0 \pm \Delta - kv_{atom}$, strikes the atom from the right and a light beam with the frequency components, $\omega = \omega_0 \pm \Delta + kv_{atom}$, strikes the atom from the left, where $k = 2\pi/\lambda$ is the photon wavevector for light resonant with a cycling transition in the atom at resonance frequency ω_0. Due to the Doppler shift, the light from both the left and the right has frequency components of $\omega = \omega_0 \pm \Delta$ in the atom's frame of reference. For each bichromatic light beam, the two frequency components interfere to create a beat pattern with period $T = \pi/\Delta$. Tuning the intensity such that each frequency component has the Rabi frequency $\Omega = \Delta/4$ configures the light field into two counterpropagating traveling waves of pi-pulses. When the relative phase difference between the pulse trains is set such that an atom on average first absorbs a photon from the beams propagating opposite to its velocity, and is then de-excited due to stimulated emission into the other pulse train, the atom will be slowed with a $2\hbar k$ momentum loss per stimulated-emission cycle.

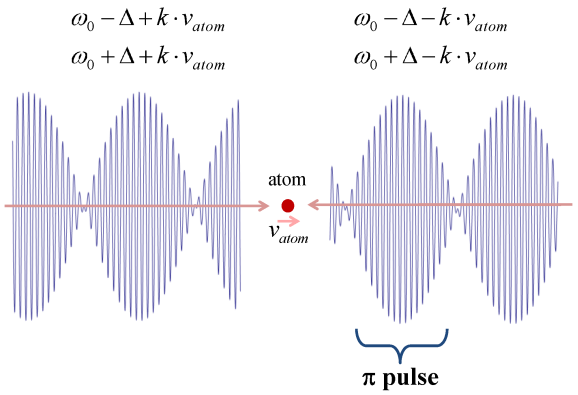


Fig. 1. The π-pulse picture of bichromatic cooling.

The relative phase of the beats from the two beams sets the preferred direction of photon momentum transfer. The optimum relative phase ϕ between the counterpropagating beat signals occurs at about $\phi \simeq \pm\pi/2$ ($+\pi/2$ for cooling, $-\pi/2$ for heating.) Spontaneous emission breaks the symmetry in the system. Since the atoms have more time after the second pulse to undergo spontaneous emission ($3T/4$), they

are preferentially excited by the first pulse and de-excited by the second pulse that arrives a period $T/4$ later. The velocity over which the bichromatic force profile acts extends over a large class, $\delta v = \Delta/k$, centered on the velocity v_{atom}.

II. Experiment

We slow atoms from a thermal beam and then capture them in a MOT. We have deliberately used low cooling powers for the MOT beams to simulate the small capture velocity that would result from the use of small MOT beams. The MOT is formed from three pairs of retro-reflected cooling beams with a 7 mm diameter and approximately 0.9 mW of laser power per beam ($I/I_{sat} = 1.4$). Under these conditions, we calculate a MOT capture velocity of $v_c \simeq 15$ m/s, which is equal to the value for a MOT of 2 mm diameter with fully saturated beams.

To test how bichromatic cooling scales with miniaturization, we performed experiments over short cooling lengths and with low laser powers. We use laser powers of < 2 mW per bichromatic frequency component, which is $\sim 0.05\times$ as large as the powers used in Ref [4]. We achieve the high intensities required for stimulated emission by focussing the light through two apertures of 0.5 mm diameter that collimate the atomic beam ($I/I_{sat} \simeq 600$). Through simulations of the force profile, we calculate a maximum stimulated force $12\times$ larger than the spontaneous force limit, with less than 8 mW of total laser power. We set the central position of the force profile to sit at 70 m/s so that cooling force extends from atoms with velocities of about 15 m/s ($\simeq v_c$) out to about 120 m/s.

We vary the effective slowing distance by changing the total interaction period of the cooling light with the atomic beam. This is accomplished by rapidly chopping the bichromatic cooling light with a radiofrequency (rf) switch that controls the rf power applied to acousto-optic modulators. We switch the cooling light on and off on time scales as short as ten microseconds. To vary the cooling length, we change the duty cycle with which we pulse the light. The full distance that the atomic beam travels in our system before entering the MOT volume is 7.5 cm, and we typically test the cooling methods down to cooling lengths of about 0.4 cm (5 % duty cycle).

To test the effectiveness of bichromatic cooling for loading a small MOT, we compare the number of captured atoms loaded with an atomic beam slowed by spontaneous forces versus one slowed by stimulated forces. For spontaneous slowing, we apply a strongly saturated monochromatic light beam red-detuned from resonance by approximately $2\pi \times 100$ MHz. We optimize the intensity and detuning to maximize the number of captured atoms. We quantify the the effects of the beam slowing in terms of MOT number enhancement, which we define as the ratio of captured atoms in the MOT with and without slowing light applied to the atomic beam.

III. Simple Model

We have performed simple modeling of the bichromatic cooling process with which we compare our experimental results. When the phase and Rabi frequency are optimized for a selected value of Δ, the force profile is nearly rectangular in shape with a width of Δ/k (see, for example, Refs. [4], [5]). The bichromatic force is averaged across the Gaussian beam profile since the Rabi frequency varies with intensity. The force profile is further averaged over the full slowing length of our system, because the relative phase of the π pulse trains varies by about 20 degrees. This averaging broadens the range of velocities over which the force acts by about 50 % and lowers the maximum force by about a factor of 1/2. For simplicity, we have approximated the ideal force profile with a rectangular function of constant force that we fit to the force profile determined via numerical modeling of the optical Bloch equations. For the spatially averaged profile, we have decreased the maximum force and broadened the addressed velocity class to fit the spatially averaged force.

To estimate the fraction of captured atoms in a MOT, we have calculated an approximate "effective capture velocity," v_{max}. To find v_{max}, we set the work done on the atoms by the bichromatic force over the cooling distance L equal to the initial kinetic energy of the atoms with a velocity v_{max} and we solve for v_{max}. This gives $v_{max} = (2R_{scatt}Lv_{rc})^{1/2}$, where R_{scatt} is the photon scatter rate and v_{rc} is the recoil velocity. Note that this is the same equation that is often used for the MOT capture velocity with laser beams of diameter L [1]. For stimulated bichromatic forces, R_{scatt} is determined from numerical modeling to be about $10\times$ greater than what is achieved via spontaneous forces for the powers used in our experiment. To estimate the fraction of the atomic beam captured in the MOT, we integrated a normalized Maxwell-Boltzmann distribution out to v_{max} or the upper edge of the calculated force profile, whichever is smaller.

Fig. 2 is a plot of the simulated fraction of atoms captured in a MOT versus duty cycle for spontaneous and stimulated cooling processes. The fraction captured with the bichromatic cooling method saturates at low duty cycles (effective cooling lengths) because the atoms are cooled very quickly by the strong bichromatic force, and they stop interacting with the cooling light when their velocity is lowered below the edge of the force profile. The strongly saturated spontaneous cooling profile is weaker than the ideal bichromatic force by about an order of magnitude, but it acts over a much larger velocity range, which is why the fraction of captured atoms continues to rise for larger duty cycles, whereas the bichromatic fraction captured saturates.

IV. Results

A detailed paper on our experimental results is currently in preparation [7]. By use of bichromatic forces, we have demonstrated an increase in atom number captured from an un-cooled atomic beam by a factor of about 3.5 for a 1.5 cm effective cooling length, and by up to a factor of 5.5 for longer cooling lengths. Consistent with Fig. 2, the technique yields MOT atom number enhancement that is about $3\times$ higher than the enhancement achieved with spontaneous forces for a 1.5 cm effective cooling length.

Fig. 3 shows example data for MOT number enhancement versus duty cycle achieved with and without an applied

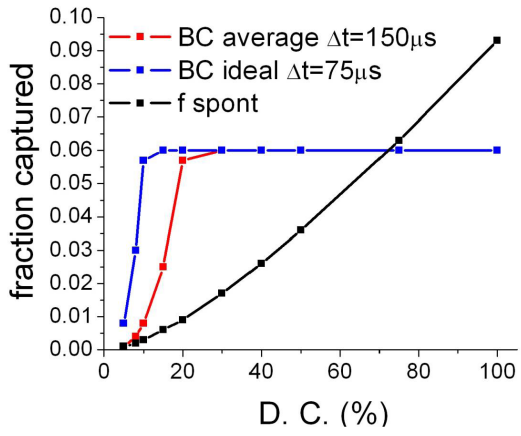


Fig. 2. Simulations of the fraction of captured atoms versus duty cycle for spatially averaged and ideal bichromatic force profiles as well as for a spontaneous force profile. To convert from duty cycle to effective cooling length (in centimeters), multiply the x-axis by 0.075. $\Delta t = m\delta v/F$ is the time required to slow an atom from the upper to the lower limit of the force profile of width $\delta v = \Delta/k$. The fraction captured in the MOT without any slowing of the atomic beam is 0.0002.

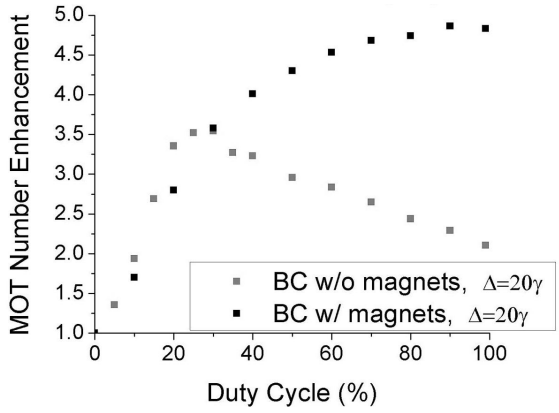


Fig. 3. Measurements of the MOT number enhancement versus duty cycle with (black squares) and without (gray squares) a magnetic guiding potential between the apertures.

magnetic guiding field. The unguided data are qualitatively similar to Fig. 2, but the enhancement actually decreases for the higher duty cycles after a maximum of 3.5× is reached for a 20 % duty cycle (1.5 cm effective cooling length). In other experiments, the MOT number enhancement actually *increased* when the bichromatic force was made *weaker* – particularly for the higher duty cycles (data not shown).

We attribute these dependencies on force magnitude and cooling length to cold-atom loss. Consistent with this, the MOT number enhancement that we observe is about 60 times smaller than what we would expect. We suspect that the main loss channel is transverse spreading of the atomic beam. The divergence angle of the uncooled atomic beam is set by the aperture geometry in the atomic beam collimator to approximately 10 mrad. The divergence angle for a given atom is equal to its transverse velocity divided by its longitudinal velocity. When an atom is slowed, its divergence angle increases and it is less likely to emerge from the second pinhole in the collimator.

One of the data sets in Fig. 3 was measured with a linear quadrupole field applied between the collimator apertures by two permanent magnets. The magnetic field gradient was $\sim$ 20 G/cm, which is strong enough to guide atoms with transverse velocities of up to 0.1 m/s through the second collimator aperture. Since the maximum transverse velocity for the atoms addressed by the cooling light is $\sim$ 1.2 m/s, we would not expect to guide all of the atoms with the quadrupole field, but we would expect to increase the enhancement for the proper cooling-light polarization. With the guiding field, we increased the MOT number enhancement for the higher duty cycles from about 2× to about 5× (Fig. 3). As expected, the increase in MOT atom number enhancement was only observed for the proper cooling-light polarization.

V. Conclusion

By use of bichromatic cooling, with our current experimental parameters we should be able to slow 6 % of the atoms in an atomic beam to MOT capture velocities in distances of under 1 cm.

We are currently working on experimental design changes that will alleviate the problem of cold atom loss, after which we will try the method in a compact system with trapping volumes of order 1 mm^3. Ultimately, we expect to optimize our system to be able to load an optical molasses with 1 mm diameter beams with 100× more atoms than if loaded from a background vapor.

Acknowledgment

We thank M. Partlow and H. Metcalf for the code that performs the numerical modeling of the bichromatic force. We acknowledge funding from NIST, DARPA, and NRC. This paper is a contribution of the NIST, an agency of the U.S. government, and is not subject to copyright.

References

[1] K. E. Gibble, S. Kasapi, and S. Chu, "Improved magneto-optic trapping in a vapor cell," Opt. Lett. **17**, 526-528 (1992)

[2] S. Pollock, J. P. Cotter, A. Laliotis, F. Ramirez-Martinez, and E. A. Hinds, "Characteristics of integrated magneto-optical traps for atom chips," arXiv:1101.0686 (2011)

[3] H. J. Metcalf and P. van der Straten, *Laser Cooling and Trapping*, Springer-Verlag, New York (1999)

[4] J. Söding, R. Grimm, Y. B. Ovchinnikov, Ph. Bouyer, and C. Salomon, "Short-distance atomic beam deceleration with a stimulated light force," Phys. Rev. Lett. **78**, 1420 (1997)

[5] M. R. Williams, F. Chi, M. T. Cashen, and H. Metcalf, "Measurement of the bichromatic optical force on Rb atoms," Phys. Rev. A **60**, R1763 - R1766 (1999)

[6] M. Partlow, X. Miao, J. Bochmann, M. Cashen, and H. Metcalf, "Bichromatic slowing and collimation to make an intense helium beam," Phys. Rev. Lett. **93**, 213004 (2004)

[7] T. C. Liebisch, E. Blanshan, E. A. Donley, and J. Kitching, "Boosting the MOT atom number by use of stimulated light forces," In Preparation, (2011)

CPT-pump-probe measurement of the Cs clock transition DC Stark shift

Jean-Luc Robyr, Paul Knowles, and Antoine Weis
Department of Physics
University of Fribourg, CH-1700 Fribourg
Email: jean-luc.robyr@unifr.ch

***Abstract*— We report progress in measuring the third order electric polarizability of the Cs ground states using a Ramsey pump-probe experiment on coherent population trapped (CPT) atoms in a thermal atomic beam. We give a short description of the apparatus as well as the Fourier transform method used to monitor the phase and frequency of the Ramsey signal. Analysis of a typical data set is shown, proving the consistency of the method. We show that the motional magnetic field phase shift can be used to test the reliability of the electric field modeling. Finally, we give a preliminary value for the Cs ground state polarizability and compare it to previous published values of the DC Stark shift.**

I. Introduction

The second is defined by the unperturbed ground-state hyperfine frequency splitting in ^{133}Cs (ISO 31–1). Realizing the second requires control and understanding of the systematic effects [1], [2] that perturb that transition. One effect is the AC Stark shift arising from the blackbody radiation spectrum of the finite temperature clock environment. In practice, the blackbody radiation shift can be calculated from the Cs ground state third-order polarizability $\alpha_0^{(3)}(F)$ which we measure in a CPT-based Ramsey pump-probe experiment similar to [3] and [4] on a Cs thermal atomic beam. The polarizability induces a shift in the resulting Ramsey fringe as a function of the DC electric field applied in the free evolution region of the atom's flight.

The energy shift produced by a DC electric field $\mathbb{E}$ is described using the electric polarizability α as $\Delta E(n, L_J, F, m_F) = -\frac{1}{2}\alpha\mathbb{E}^2$. The constant α is traditionally expanded in power series, both in multipole order (subscript) and perturbation order (superscript). Neglecting the (small) quadrupole hyperfine interaction, the lowest order terms (i.e., third) affecting the $\Delta m_F = 0$ ground-state Cs hyperfine transitions, produce a frequency shift $\delta\nu = -\frac{1}{2h}\alpha\mathbb{E}^2$ [5], [6] with

$$\alpha = \frac{16}{7}\alpha_0^{(3)}(F{=}4) + \alpha_2^{(3)}(F{=}4)\frac{3m_F^2-16}{28}f(\theta)\,, \qquad (1)$$

where $\alpha_0^{(3)}$ and $\alpha_2^{(3)}$ are the third order scalar and tensor polarizabilities, m_F is the magnetic quantum number defined by the quantization axis (magnetic field), and $f(\theta) = 3\cos^2(\theta) - 1$ is a function of the relative angle between $\hat{\mathbb{E}}$ and the quantization axis. Note that $f(\theta)$ averages to zero for isotropic blackbody radiation, and that the Ramsey method makes the shift sensitive to the effective average squared field, $\langle\mathbb{E}^2\rangle$. For a 20 kV/cm static electric field, the expected scalar shift in Cs is 830 Hz, and the tensor shift is about 3 Hz.

A. Motional magnetic field

Atoms flying through the electric field $\vec{\mathbb{E}}$ experience an additional magnetic field $\vec{B}_{mot}$ given by

$$\vec{B}_{mot} = -\frac{\vec{v}}{c^2}\times\vec{\mathbb{E}}\,,$$

where $\vec{v}$ is the velocity of the atom. This motional field produces a Zeeman shift of the $\Delta m_F = 0$ hyperfine transitions but due to its linear velocity dependence this shift results in a phase change

$$\Delta\phi = (g_{F=4} - g_{F=3})m_F\mu_B\frac{\langle\mathbb{E}\rangle}{\hbar}\frac{L}{c^2} \qquad (2)$$

of the Ramsey resonance, where g_i is the Landé factor of level i, m_F is the magnetic quantum number of both ground states (our experiment presently probes phase shifts of levels with identical m_F), $\langle\mathbb{E}\rangle$ is the mean electric field seen by the atoms between the pump and the probe zones, separated by L. This phase shift produces a small displacement of the Ramsey resonance center frequency and thus mimics an atomic frequency shift. However, from this measured phase, the mean electric field value can be inferred and then compared to an electric field model prediction to give confidence in the latter.

B. Extracting frequency and phase

The Ramsey resonance lineshape $S(\omega_{RF})$ can be expressed as the integral of the phase accumulated by each velocity class during the interaction with the electric and magnetic fields weighted by the atomic velocity distribution, $\rho(v)$. The accumulated phase is given by a term oscillating at the difference between the microwave ω_{RF} and atomic ω_{at} frequencies and includes a constant offset ϕ.

$$S(\omega_{RF}) = \int_0^\infty \rho(v)\cos\left[(\omega_{RF}-\omega_{at})\frac{L}{v} + \phi\right]dv\,. \qquad (3)$$

Taking the Fourier transform of $S(\omega_{RF})$ yields

$$\mathcal{F}[S(\omega_{RF})] \propto \rho(T)\exp[i(\omega_{at}T + \phi)]\,, \qquad (4)$$

where $\rho(T)$ is the atomic time-of-flight distribution. The argument of the complex exponential contains information on both atomic frequency ω_{at} and phase ϕ of the Ramsey signal $S(\omega_{RF})$ permitting their clear separation and extraction. The

978-1-61284-110-6/11/$26.00 ©2011 IEEE

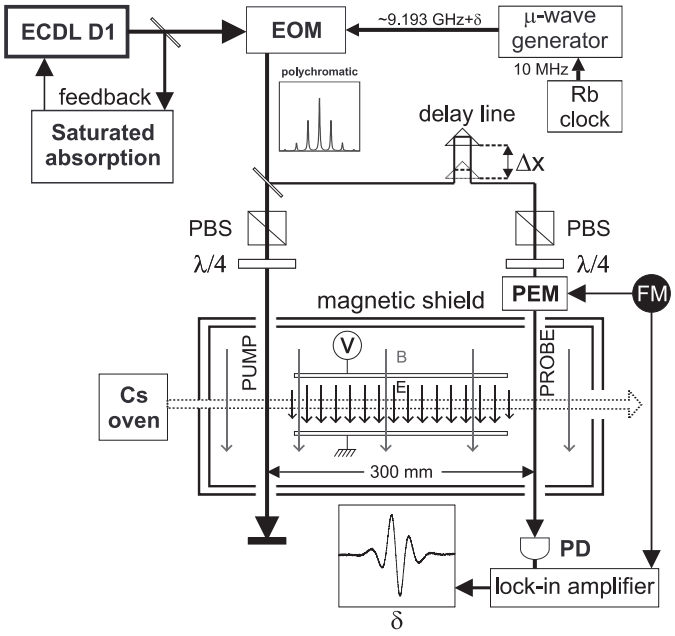


Fig. 1. Experimental setup for the measurement of the Stark shift. In order to produce a phase coherent bichromatic light field, a Cs D_1 extended cavity diode laser beam is phase modulated at the Cs hyperfine frequency using a electro-optic modulator (EOM). The modulated laser beam is split into a strong pump and a weak probe and sent to a Cs atomic beam experiment to perform Ramsey interrogation. The probe-beam transmitted intensity is recorded while scanning the microwave frequency. The DC Stark shift is obtained applying an electric field to the atoms during their flight between the pump and the probe zones.

Fourier transform is therefore a powerful tool to distinguish between Stark frequency shift and motional phase shift: similar techniques have been used in [7], [8].

II. Experimental setup

The main elements of the present experimental scheme have already been described in detail in [9] and here we will focus on the new elements. Figure 1 shows the actual experimental configuration. A Cs D_1 extended cavity diode laser is stabilized to a chosen Cs hyperfine transition (normally the F=3 $\longrightarrow$ F'=3) using saturated absorption in an atomic vapor cell. The light is sent through a commercial lithium niobate electro-optic modulator (EOM), which accepts a microwave signal (whose frequency can be scanned around the Cs ground state hyperfine frequency ~ 9.2 GHz), and produces a frequency modulated spectrum, containing multiple sidebands all separated by the modulation frequency. The resulting spectrum generates coherent population trapping (in this case, the carrier and the $-\omega_{RF}$ bands are used). The modulated beam is split into a strong pump beam and a weak probe beam. The pump and probe beams pass through polarization optics to create the desired circular polarization before being sent to the thermal Cs atomic beam. The interaction between the light and the atoms occurs in a magnetically shielded environment with a 3.6 μT DC magnetic field. The atomic beam first interacts with the pump laser beam creating a coherence between the Cs hyperfine ground states which then evolves

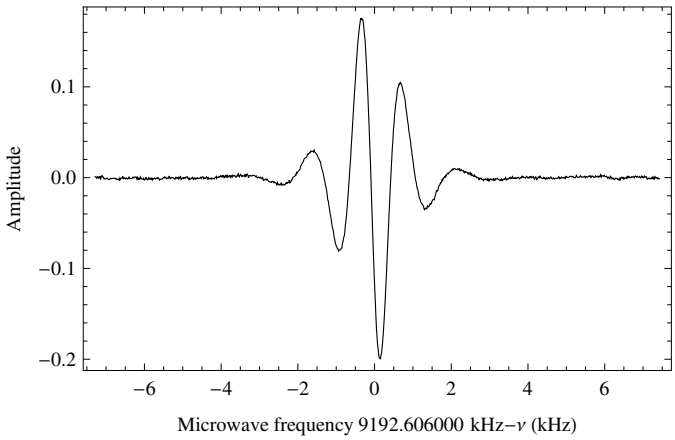


Fig. 2. Averaged Ramsey fringe pattern (15 scans) of the Cs hyperfine transition $m_F = +1 \rightarrow m_F = +1$, with a central fringe period of ~ 1 kHz. The dispersive lineshape is due to the difference of microwave phases between the pump and probe zones and can be controlled by the optical delay line.

over a 300 mm path length during which it passes between two 260 mm long electrodes with given potential difference. The electric field between the plates shifts the atomic coherence oscillation frequency due to the Stark perturbation. Finally the phase of the coherence is probed by recording the transmitted laser intensity of the probe beam in the second interaction zone while scanning the microwave frequency, thus producing typical Ramsey spectra (Fig. 2).

A. Electric field

The electric field is produced by applying a known voltage to one of the two conductive rectangular float glass plates, while the other is grounded. The plate separation is defined by 6.065 mm thick optical flats. Two grounded metal plates, with apertures collimating the Cs beam, are placed at the entrance and exit of the capacitor to shield the light interaction zones from electric fringe fields. The applied high voltage is measured by a calibrated digital voltmeter (with resolution of 10^{-4}), and remains stable to better than that resolution for the typical measurement time of 10 minutes. A technical problem prevented the use of negative voltages exceeding -2000 V. Finite element electric field modeling is used in the data analysis to extract the polarizability [10].

III. Experimental results

A. Ramsey fringe data

Ramsey fringe spectra are recorded for each of the seven $\Delta m_F = 0$ hyperfine transitions, split by the 3.6 μT field. Due to the $\sim$ 1 ms precession time in the fields, the period of the central fringe is about 1 kHz. Each resonance spectrum is obtained ramping the microwave frequency over ± 7.5 kHz around the center of the fringe. The background is removed using a quadratic fit function. The phase of the Ramsey resonance can be accurately controlled by an optical delay line. Figure 2 shows a typical Ramsey fringe spectrum for the $m_F = +1 \rightarrow m_F = +1$ hyperfine ground state transition.

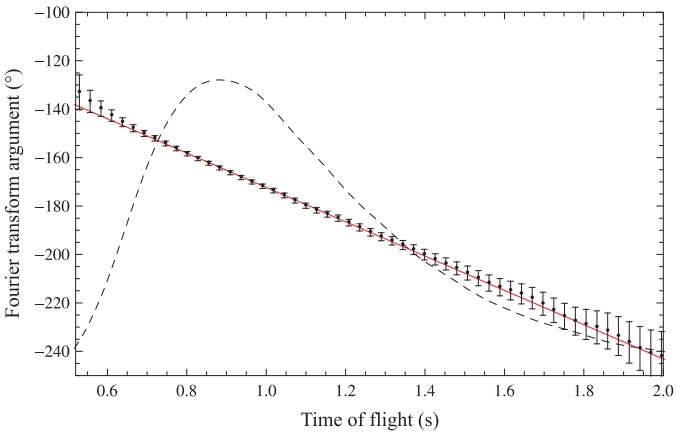


Fig. 3. Fourier transform data of recorded Ramsey signal showing the linear dependence of the argument $\omega_{at}T + \phi$ in Eq. (4) in function of the time of flight. Each data point is weighted proportionally to the time of flight distribution in the beam (dashed line). A weighted linear fit is used to extract both the atomic frequency ω_{at} and the Ramsey fringe phase ϕ.

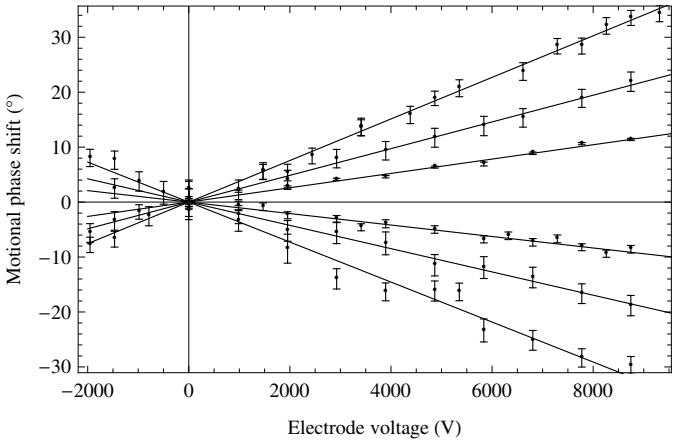


Fig. 4. Measurement of the Ramsey resonance phase as function of the applied electrode voltage for hyperfine transitions with Δm_F=0 and m_F=± 1, ±2, ±3. The linear fit of these data gives access to the mean electric field experienced by the atoms on their flight between the pump and probe interactions.

B. Fourier transform of Ramsey data

A discrete Fourier transform is applied to the Ramsey spectra. To avoid noise coming from asymmetric ends of the spectrum, a Tukey windowing function is applied to the data before transformation. The transformed data become functions of the atomic time-of-flight with a minimal time interval between two data points determined by the microwave scan span. However, since the applied frequency window goes to zero and there is no signal beyond the span limits, the effective span can be arbitrarily extended during the Fourier transform. As consequence, we are not limited in the time resolution of the Fourier transform. Figure 3 shows typical data from the Fourier transform's argument $\omega_{at}T + \phi$ in Eq. (4). Each data point is attributed a weight proportional to the corresponding signal intensity from the time-of-flight distribution $\rho(T)$ (dashed line in Fig. 3). The atomic frequency ω_{at} as well as the Ramsey fringe phase ϕ are extracted by a weighted fit of a straight line.

C. Motional magnetic field

The motional magnetic field produces a phase change of the Ramsey resonance as a function of the applied electric field, as described above. The Fourier transform analysis of the Ramsey spectra for different applied voltages gives access to this phase shift. Figure 4 shows the motional phase shift of the six Δm_F=0 hyperfine transitions with m_F=± 1, ±2, ±3 as fit with a linear function. As expected from Eq. (2) the slope depends linearly on m_F. The linear fits therefore give measurements of $\langle \mathbb{E} \rangle$ which can be compared to the mean electric field value obtained from the field modeling, useful to verify the accuracy of the latter. The quantity compared is the effective electrode separation d_{eff} defined as

$$\langle \mathbb{E} \rangle = \frac{V}{d_{\text{eff}}},$$

where $\langle \mathbb{E} \rangle$ is the mean electric field between the pump and probe interaction and V the applied voltage. Table I shows that the model and measurements agree. We currently work on improving the statistical and systematic uncertainties of the measurement.

	Field modeling	Measurement
d_{eff}	5.90(2) mm	5.83(8) mm

TABLE I

VALUES FOR THE EFFECTIVE DISTANCE BETWEEN THE ELECTRODE PLATES OBTAINED WITH THE ELECTRIC FIELD NUMERICAL MODELING AND FROM THE MOTIONAL MAGNETIC FIELD PHASE SHIFT MEASUREMENT.

D. Stark shift

The Cs ground state Stark shift is obtained by measuring the change of atomic frequency ω_{at} as a function of the applied electrode voltage using the method described above. Figure 5 shows the Stark frequency shift for the m_F=+1 → m_F=+1 Cs ground state hyperfine transition. A second order polynomial is fit to the data for extracting the quadratic coefficient, itself proportional to the electric polarizability of the Cs ground states. The linear coefficient is consistent with zero, as expected. The voltage is converted into the effective mean of the squared electric field $\langle \mathbb{E}^2 \rangle$ seen by the atoms between the pump and probe interactions using numerical modeling of the electric field. A value of the total electric polarizability for each of the seven Cs ground states hyperfine transition is obtained and the polarizability difference of any pair of grounds states can be extracted using Eq. (1). Preliminary data with fit are shown in Fig. 6. The relative orientation of the electric and magnetic field for these measurement is $\theta = 85(1)^{\circ}$. A preliminary value for $\alpha_0^{(3)}$ has been extracted from the fit and is compared, in Table II, with the two latest DC Stark shift measurements (which between themselves showed a discrepancy of more than 5 σ).

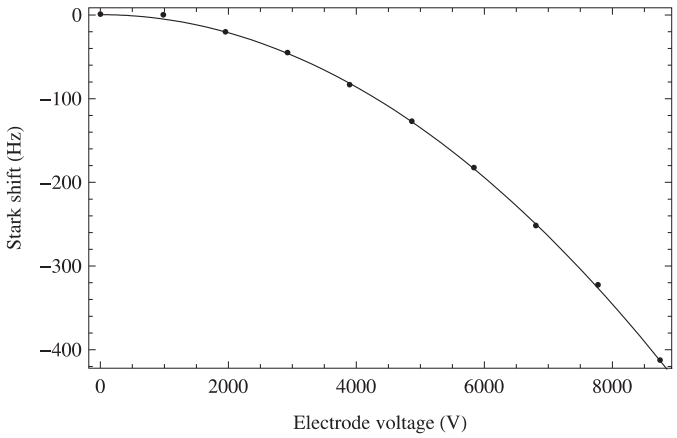


Fig. 5. Cs ground state Stark frequency shift of the $m_F{=}+1 \rightarrow m_F{=}+1$ hyperfine transition. Statistical uncertainties on the points are smaller than the point size. The continuous line is a second degree polynomial fit.

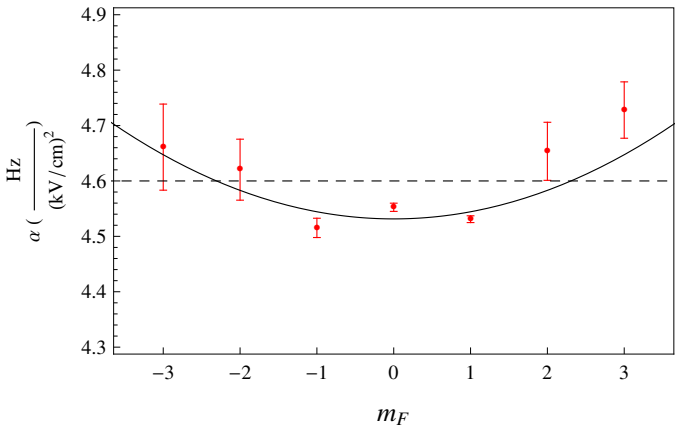


Fig. 6. Cs ground state electric polarizability extracted for each of the seven hyperfine transitions $(4, m_F \leftrightarrow 3, m_F)$. Error bars come from the measurement statistics with sizes depending on the signal to noise ratio of the Ramsey fringe, which is higher for low m_F transitions. The continuous line shows the fit with Eq. (1). The dashed line shows only the scalar contribution in α from Eq. (1).

IV. Conclusion

The Fourier analysis of the Ramsey spectra phase and center frequency has been successfully applied to measured data, extracting the Stark frequency shift, and the motional phase needed to test the electric field model which itself permits the third order Cs ground state polarizability $\alpha_0^{(3)}$ to be inferred. Our preliminary value for $\alpha_0^{(3)}$ was compared to previous measurements of the DC Stark shift. Further work will focus on improving the statistical uncertainties of the measurements and characterize and reduce the systematic effects.

Acknowledgment

The authors thank the mechanical and electronics workshops of the Fribourg University Physics Department for their skillful support, and the Swiss National Science Foundation (SNF #200021–117841) for funding this project. Thanks go to Z. Andjelic, of ABB Corporate Research, Baden, for the field modeling.

Measurement of $\alpha_0^{(3)}$ $\left(\frac{\text{Hz}}{(\text{kV/cm})^2}\right)$		
Rosenbusch *et al.*	1.997(4)	[11], [12]
Godone *et al.*	1.78(4)	[13]
this work (preliminary)	2.01(2)	

TABLE II

Comparison of our preliminary measurement of the third order electric polarizability and two previous measurements of the DC Stark shift in the Cs ground states, which, between them, show a $>5\,\sigma$ discrepancy.

References

[1] R. Wynands and S. Weyers, "Atomic fountain clocks," *Metrologia*, vol. 42, pp. S64–S79, 2005. [Online]. Available: http://www.iop.org/EJ/abstract/0026-1394/42/3/S08

[2] J. H. Shirley, W. D. Lee, and R. E. Drullinger, "Accuracy evaluation of the primary frequency standard NIST–7," *Metrologia*, vol. 38, no. 5, pp. 427–458, 2001. [Online]. Available: http://stacks.iop.org/0026-1394/38/427

[3] J. E. Thomas, P. R. Hemmer, S. Ezekiel, C. C. Leiby, R. H. Picard, and C. R. Willis, "Observation of Ramsey Fringes Using a Stimulated, Resonance Raman Transition in a Sodium Atomic Beam," *Phys. Rev. Lett.*, vol. 48, no. 13, pp. 867–870, Mar. 1982. [Online]. Available: http://prola.aps.org/abstract/PRL/v48/i13/p867_1

[4] J. Fuchs, G. J. Duffy, W. J. Rowlands, and A. M. Akulshin, "Electromagnetically induced transparency in ^{6}Li," *J. Phys. B*, vol. 39, no. 17, pp. 3479–3489, 2006. [Online]. Available: http://stacks.iop.org/0953-4075/39/3479

[5] S. Ulzega, "Theoretical and experimental study of the Stark effect in the ground state of alkali atoms in helium crystals," Ph.D. dissertation, Department of Physics, University of Fribourg, Switzerland, 2006, unpublished.

[6] S. Ulzega, A. Hofer, P. Moroshkin, and A. Weis, "Reconciliation of experimental and theoretical electric tensor polarizabilities of the cesium ground state," *Europhys. Lett.*, vol. 76, no. 6, pp. 1074–1080, Dec. 2006, see errata in [14]. [Online]. Available: http://dx.doi.org/10.1209/epl/i2006-10383-2

[7] G. Di Domenico, L. Devenoges, A. Stefanov, A. Joyet, and P. Thomann, "Fourier analysis of Ramsey fringes observed in a continuous atomic fountain for in situ magnetometry," *ArXiv e-prints*, Apr. 2011.

[8] J. H. Shirley, "Velocity distributions calculated from the fourier transforms of ramsey lineshapes," *IEEE Transactions on Instrumentation and Measurement*, vol. 46, no. 2, pp. 117–121, Apr. 1997.

[9] J.-L. Robyr, P. Knowles, and A. Weis, "Stark shift of the cs clock transition frequency: a new experimental approach," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 57, no. 3, pp. 613–617, Mar. 2010.

[10] Z. Andjelic, POLOPT http://www.poloptsoftware.com/htm/modules_portfolio.htm.

[11] P. Rosenbusch, S. Zhang, and A. Clairon, "Blackbody radiation shift in primary frequency standards," in *Frequency Control Symposium, 2007 Joint with the 21st European Frequency and Time Forum. IEEE International*, vol. 29, no. 1, June 2007, pp. 1060–1063.

[12] E. Simon, P. Laurent, and A. Clairon, "Measurement of the Stark shift of the Cs hyperfine splitting in an atomic fountain," *Phys. Rev. A*, vol. 57, no. 1, pp. 436–439, Jan. 1998. [Online]. Available: http://prola.aps.org/abstract/PRA/v57/i1/p436_1

[13] A. Godone, D. Calonico, F. Levi, S. Micalizio, and C. Calosso, "Stark-shift measurement of the $^2S_{1/2}$, F=3→F=4 hyperfine transition of ^{133}Cs," *Phys. Rev. A*, vol. 71, no. 6, p. 063401, 2005. [Online]. Available: http://link.aps.org/abstract/PRA/v71/e063401

[14] S. Ulzega, A. Hofer, P. Moroshkin, and A. Weis, "Erratum: Reconciliation of experimental and theoretical electric tensor polarizabilities of the cesium ground state," *Europhys. Lett.*, vol. 78, no. 6, p. 69901, June 2007. [Online]. Available: http://dx.doi.org/10.1209/0295-5075/78/69901

An Efficient AT-cut Quartz Crystal Resonator Design Tool for Activity Dip in Working Temperature Range

Shih-Yung Pao
TXC Corporation
Taoyuan, Taiwan/R.O.C.
sypao@txc.com.tw

Qiao-Qiao Pan, Min-Chiang Chao
TXC(Ningbo) Corporation
Ningbo, Zhejiang/China

***Abstract*—The work of an efficient design tool for engineers to enhance the frequency stability performance of AT-cut quartz crystal resonator in whole working temperature range is presented. Based on Mindlin's 2D AT-cut plate model, the design tool is a finite element analysis(FEA) program to simulate quartz plate vibration with partial electrodes(mass loading), finite dimensions(free edge), piezoelectric and temperature effects. The simulation and experiment results, including mode chart, electrical response and frequency-temperature characteristics, of a 3.2mm by 2.5mm 40MHz resonator are shown in this study. Finally, an activity-dip-contour-map which locates the high frequency-temperature stability zone is constructed and it guides engineers to do a robust design.**

I. INTRODUCTION

As a key frequency device in almost all electric circuits, AT-cut quartz resonator should be very accurate and high stability, and therefore could not be designed by experiments or trial and error only. Mindlin[1],[2] did a series of work in crystal plate vibration as the base of AT-cut quartz resonator analysis. Lee[3] constructed a FEA program based on Mindlin's 2D model to analyze pure mechanical crystal plate vibration without piezoelectric effects. Yong and Wang[4],[5] extended the work to 3D FEA and some details, like frequency-temperature characteristics and quality factor. Today, some commercial FEA software could also simulate quartz plate vibration. However, it is time consuming to analyze quartz resonator by general commercial FEA programs because of the anisotropic structure and aspect ratio (length/width to thickness) of the quartz plate. Hence, an efficient and accurate enough design tool for modern quartz crystal resonator is always desired.

In this study, a 2D AT-cut quartz plate finite element model based on Mindlin's 2D theory is built. By the weak coupling of quartz, the pure mechanical vibration and piezoelectric effects are considered separately in the FEA to enhance the calculation efficiency. The simulation mode charts by 2D and 1D model show 2D model could contain more complete modes of quartz crystal plate. Due to including piezoelectric effects in the FEA program, the simulation electrical response (at room temperature) and frequency activity dip (in working temperature range) match the experiment results very well.

Today's high volume quartz crystal resonators' business is quite challenging. Cost reduction and maintain high quality is the key for survive. Engineers should take care wide temperature range activity dips of resonators, and this makes argues inside about design and process control. For shortening the development time and saving the cost of frequency temperature stability testing, an activity-dip-contour-map is introduced. The ordinate and abscissa of the map are the length and width of the electrode (or quartz plate), respectively. And the contour level shows the frequency activity dip (in ppm) due to working temperature change. The simulation activity-dip-contour-map is similar to the experiment results and could help engineers to choose a comfortable zone for producing.

II. TWO-DIMENSIONAL AT-CUT QUARTZ PLATE MODEL

Considering a rectangular thin crystal plate shown in Figure 1. the width, length and thickness of the plate are 2c, 2a and 2b, respectively, and it is partially plated by metal as electrodes with 2W by 2L dimension.

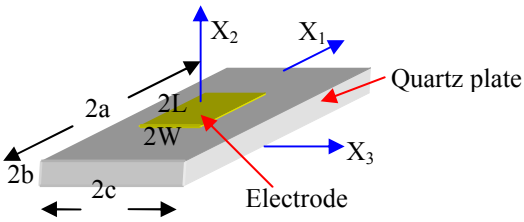


Figure 1. A rectangular crystal plate with electrodes

The equations of motion obtained from Mindlin's 2D plate theory[6]

$$\begin{cases} c_{ijkl}^{(0)}\left(u_{k,li}^{(0)}+\delta_{2l}u_{k,i}^{(1)}\right)+e_{kij}^{(0)}\varphi_{,ki}^{(0)}+T_j^{(0)}=\rho\ddot{u}_j^{(0)} \\ e_{kij}^{(0)}\left(u_{j,ki}^{(0)}+\delta_{2i}u_{j,k}^{(1)}\right)-\varepsilon_{ij}\varphi_{,ij}^{(0)}+D^{(0)}=0 \\ c_{abcd}^{(1)}u_{d,ca}^{(1)}+e_{cab}^{(1)}\varphi_{,ca}^{(1)}-3b^{-2}\left[c_{2bkl}^{(0)}\left(u_{l,k}^{(0)}+\delta_{2k}u_l^{(1)}\right)+e_{a2b}^{(0)}\varphi_{,a}^{(0)}+e_{22b}^{(0)}\varphi^{(1)}\right]+T_b^{(1)}=\rho\ddot{u}_b^{(1)} \\ e_{cab}^{(1)}u_{b,ca}^{(1)}-\varepsilon_{bc}\varphi_{,bc}^{(1)}-3b^{-2}\left[e_{2ijl}^{(0)}\left(u_{j,i}^{(0)}+\delta_{2i}u_j^{(1)}\right)-\varepsilon_{a2}\varphi_{,a}^{(0)}-\varepsilon_{22}\varphi^{(1)}\right]+D^{(1)}=0 \end{cases} \quad (1)$$

where:

$u^{(0)}$, $u^{(1)}$: the zero and first order of mechanical displacement, and $i,j,k,l=1,2,3$; $a,b=1,3$

$\varphi^{(0)}$, $\varphi^{(1)}$: the zero and first order of electrical potential

$T^{(0)}$, $T^{(1)}$: the zero and first order of stress

$D^{(0)}$, $D^{(1)}$: the zero and first order of electrical displacement

c, e, ε: the components of elastics stiffness, piezoelectric stress constant and dielectric permittivity of AT-cut quartz plate, respectively

Defining zero order mechanical strain (ν), mechanical stress (P, Q), electric field (β) and electric stress (P_E, Q_E); and first order mechanical strain (χ), mechanical stress (M), electric field (η) and electric stress (M_E) as shown in Figure 2. and Figure 3.

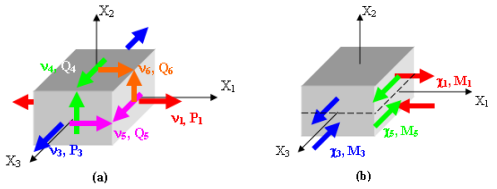


Figure 2. The sketch of strains and stresses: (a)zero order; (b) first order

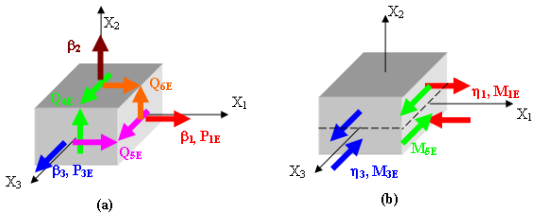


Figure 3. The sketch of electric field and stresses: (a)zero order; (b) first order

equation (2) could be express by these stress items[7]:

$$\begin{cases} P_{1,1}+Q_{5,3} & +P_{1E,1}+Q_{5E,3} & =\rho(1+R)\ddot{u}_1^{(0)} \\ Q_{4,3}+Q_{6,1} & +Q_{4E,3}+Q_{6E,1} & =\rho(1+R)\ddot{u}_2^{(0)} \\ P_{3,3}+Q_{5,1} & +P_{3E,3}+Q_{5E,1} & =\rho(1+R)\ddot{u}_3^{(0)} \\ M_{1,1}+M_{5,3}-Q_6 & +M_{1E,1}+M_{5E,3}-Q_{6E} & =\rho(1+3R)\ddot{u}_1^{(1)} \\ M_{3,3}+M_{5,1}-Q_4 & +M_{3E,3}+M_{5E,1}-Q_{4E} & =\rho(1+3R)\ddot{u}_3^{(1)} \\ D_{1,1}^{(0)}+D_{3,3}^{(0)} & -D_{d1,1}^{(0)}-D_{d3,3}^{(0)} & =-D^{(0)} \\ D_{1,1}^{(1)}+D_{3,3}^{(1)}-D_2^{(0)} & -D_{d1,1}^{(1)}-D_{d3,3}^{(1)}+D_{d2}^{(0)} & =-D^{(1)} \end{cases} \quad (2)$$

introducing a interpolation function for finite element model, and by energy densities in terms of nodal displacement and variational principle, obtained the equations of motion for one element in matrix form:

$$\begin{cases} [M^e][\ddot{U}^e]+[d_m^e][\dot{U}^e]+[K_m^e][U^e]+[K_p^e][\Psi^e]=[\hat{F}^e] \\ [K_p^e]^T[U^e]+[K_d^e][\Psi^e]=[\hat{D}^e] \end{cases} \quad (3)$$

where

$[\mathrm{M}^e]$: element mass matrix; $[\mathrm{M}_m^e]$: element elastic stiffness; $[\mathrm{d}_m^e]$: element damping matrix; $[\mathrm{K}_p^e]$: element piezoelectric coupling matrix; $[\mathrm{K}_d^e]$: element dielectric stress matrix; $[\hat{\mathrm{F}}^e]$ and $[\hat{\mathrm{D}}^e]$ are the element mechanical force and surface electric displacement.

III. FINITE ELEMENT ANALYSIS PROCESS

Assembling all elements of FEM model to obtain the system equations in matrix form are

$$\begin{cases} [M][\ddot{U}]+[d_m][\dot{U}]+[K_m][U]+[K_p][\Psi]=[\hat{F}] \\ [K_p]^T[U]+[K_d][\Psi]=[\hat{D}] \end{cases} \quad (4)$$

Based on weak piezoelectric coupling of quartz, the elastic displacement and electrical potential are solved separately [8] to enhance the calculation efficiency. The FEA process steps are

(i) Free vibration eigenfrequency analysis

$$[M][\ddot{U}]+[K_m][U]=0 \quad (5)$$

Solve (5) as an eigenproblem to get eiengmodes U_1, U_2, … U_n and the natural frequency f_1, f_2, … f_n.

(ii) Static electric potential due to applied voltage

$$[K_d][\Psi]=[\hat{D}'] \quad (6)$$

Give the surface-charge $[\hat{D}']$ as electric excitation signal to solve potential field $[\Psi]$ in the crystal plate.

(iii) Piezoelectric coupling stress due to potential field

$$[K_p][\Psi]=[\hat{F}'] \quad (7)$$

Substituting $[\Psi]$ form step (ii) into (7) to obtain $[\hat{F}']$, where $[\hat{F}']$ is the amplitude of piezoelectric coupling stress due to applied potential. In the next step, $[\hat{F}']$ is the external force to the pure mechanical system.

(iv) Force vibration frequency response analysis

$$[M][\ddot{U}]+[d_m][\dot{U}]+[K_m][U]=[\hat{F}']\cos\omega t \quad (8)$$

where $[\hat{F}']\cos\omega t$ is a harmonic external stress. Equation (8) is solved by mode superposition method to get the harmonic elastic displacement response

$$U=a_1U_1+a_2U_2+\ldots+a_nU_n$$

where U_1, U_2, … U_n are the nature modes in step (i) and a_1, a_2, … a_n are weights of each mode.

(v) Electric displacement and response

$$[K_p]^T[U]+[K_d][\Psi]=[\hat{D}] \quad (9)$$

Substituting U in step (iv) and $[\Psi]$ in step (iii) into (9) to obtain $[\hat{D}]$, the electric displacement response of the system.

(vi) Others

In the post process of the FEA, the admittance curve and equivalent circuit parameters, C_0, C_1, R_1 and L_1 .could be obtained.

IV. Simulation and Experiment Results

A 3.2mm by 2.5mm 40MHz and 7.0mm by 5.0mm 61.44MHz resonator design examples, including mode charts, frequency response verification, temperature stability verification and activity-dip-contour-map, are presented in this section.

A. Mode charts by simulation

The simulation mode charts of a 3.2mm by 2.5mm 40MHz resonator by 2D FEA model and 1D model are shown in Figure 4. and Figure 5. respectively. The former could calculate more complete modes of AT-cut quartz plate and make the simulation more accurate.

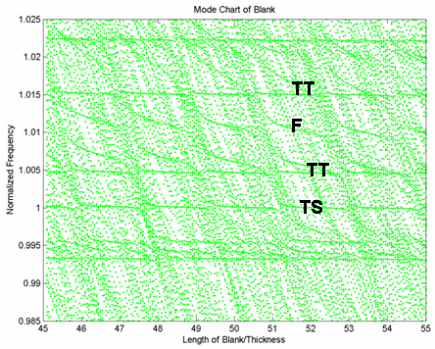


Figure 4. A 3.2mm by 2.5mm 40MHz simulation mode chart by 2D FEA model

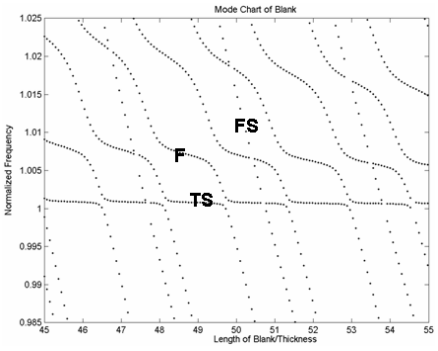


Figure 5. A 3.2mm by 2.5mm 40MHz simulation mode chart by 1D mode

B. Verification in freauency response

Considering the piezoelectric effects as a force excitation in (8), the program calculates the modes "existence obviously". The response Figure 6. shows there are three modes will be measured easily in real products as Figure 6.

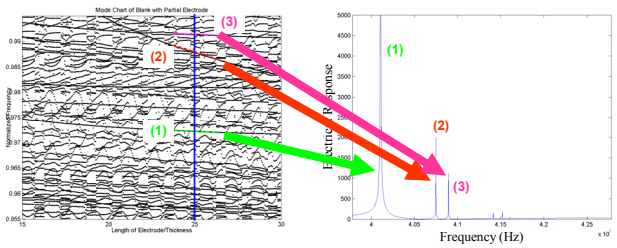


Figure 6. Simulation electrical response

We can plot the mode shape of U1(displacement in X1-axis direction) of modes (1), (2), and (3) in Figure 6. along length and width of the quartz plate, and identify mode (1) is thickness shear and mode (2) and (3) are thickness shear inharmonic overtones. The results were also check by using experiment data. In Figure 7. the left and right chats are the response by simulation and experiment, respectively. The frequency of three strong modes are marked on the simulation charts (the frequency in the parenthesis are experiment data). These two results matched very well.

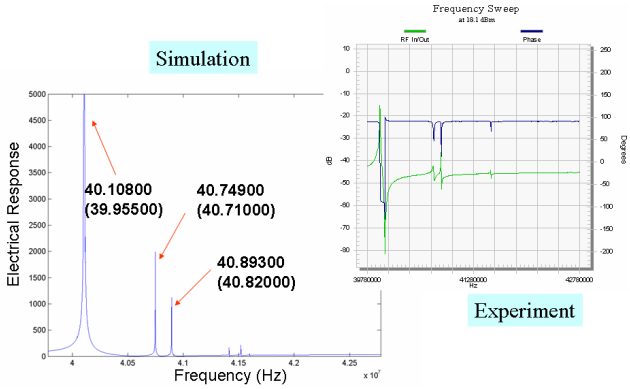


Figure 7. Verification in freauency response

C. Verification in frequency-temperature stability

In Figure 8. the left and right chats are the frequency-temperature stability curves by simulation and experiment, respectively. There are activity dips around 16 °C and 40 °C in both calculating and testing data; that means this FEA design tool could simulate the frequency-temperature characteristics very well in the whole working temperature range.

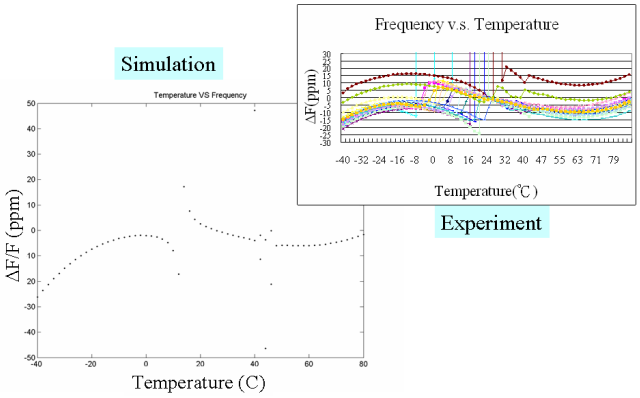


Figure 8. Verification in freauency temperature stability

D. *Activity dip contour map*

Traditionally, a crystal resonator designer should work hard to test many dimension combinations of quartz plate and electrode to verify the frequency-temperature stability performance of each size. These trial and error works make the cost of products development high; however, it spends too much time to simulate temperature characteristics by general commercial FEA program.

The design tool in this study could calculate the activity dip across the whole working temperature fast. Because of its high efficiency, it could finish the simulation for hundreds of different size in a few days with a common PC.

Figure 9 is a sketch of activity-dip-contour-map. The abscissa and ordinate are the length and width of crystal plate or electrode. The vertical axis is the activity dip of each size. The valleys on the map are "good" zones, but the peaks on the map are "bad" zones.

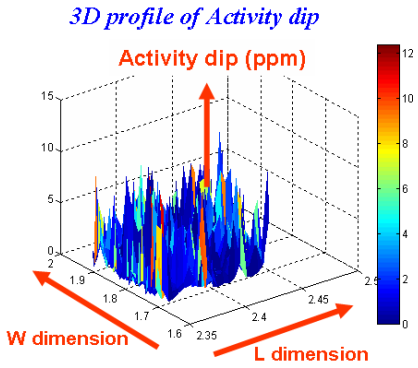


Figure 9. Activity-dip-contour-map

In Figure 10. the left and right chats are the activity-dip-contour-map by simulation and experiment, respectively. It is a 7.0mm by 5.0mm 61.44MHz design example. The main simulation conditions are:

(i) 16 steps in width, 16 steps in length

(ii) -30 ^{o}C to 85^{o}C in 2^{o}C step

(iii) Total simulation time: 5 days by PC 2.4GHz CPU with 4GB memory

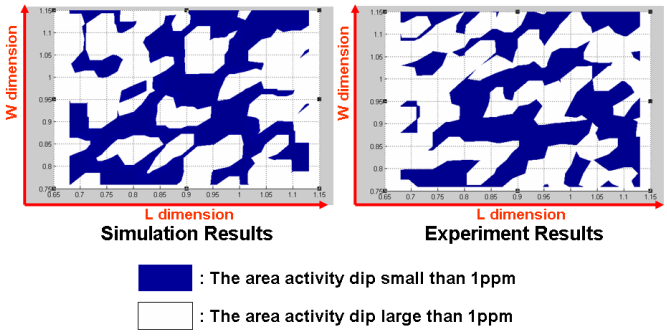


Figure 10. Simulation design map v.s. testing data

Due to some limitation of the FEA model and process variation, the activity-dip-contour-map by calculation and testing does not match very well, but it still could guide engineers to start their work in a better base for robust design.

V. CONCLUSION

A 2D Mindlin plate FEA model considering mechanics, piezoelectrical, and dielectrical effects is constructed. As quartz weak coupling characteristics, the simulation efficiency could be enhanced by calculating the mechanics and electrical response separately. The simulation results of frequency response in room temperature and the activity dip characteristics in working temperature are verified. The simulation tool can provide direct-feeling about design robustness and still need to improve to consider more factors inside model. More study activities between company and academic resources are needed.

REFERENCES

[1] R.D. Mindlin, "High frequency vibrations of plated, crystal plates," *Progress in* Applied Mechanics, The Prager Anniversary Volume, Macmillan, New York, pp. 73–84, 1963.

[2] R.D. Mindlin and W.J. Spencer, "Anharmonic, thickness-twist overtones of thickness-shear and flexural vibration of rectangular, AT-cut quartz plates," *The Journal of the Acoustical Society of America*, vol. 42, no. 6, pp 1268-1277, Dec. 1967.

[3] P.C.Y. Lee, C. Zee, and C.A. Brebbia, "Thickness-shear, thickness-twist, and flexural vibration of rectangular AT-cut quartz plates with patch electrodes," in *Proc. of Freq. Control Symp.*, pp 108-119, 1978.

[4] J. Wang, Y.K. Yong, and T. Imai, "High order plate theorey based finite element analysis of the frequency-temperature relations of quartz crystal resonators," in *Proc. of Freq. Control Symp.*, pp 956-963, 1998.

[5] Y.K. Yong, W. Wei, M. Tanaka, and T. Imai, "Three dimensional finite elements and their relationships to Mindlin's higher order plate theory in quartz crystal plate," in *Proc. of Freq. Control Symp.*, pp 791-794, 2001.

[6] R.D. Mindlin, "High frequency vibration of piezoelectric crystal plates," *Int. J. Solids Structures*, vol. 8, pp 895-906, 1972.

[7] S.Y. Pao, *A theoretical and experimental study of advanced AT-cut quartz and piezoelectric thin film resonator*, thesis in the Institude of Applied Mechanics, for the Degree of Doctor of Philosophy at National Taiwan University, Taipei, Taiwan, June, 2009.

[8] S.Y. Pao, M.K. Chao, C.S. Lam, and P.Z. Chang, "An efficient numerical method in calculating the electrical impedance different modes of AT-cut quartz crystal resonator," in *Proc. of Freq. Control Symp.*, pp 396-400, 2004.

Recent Investigations on BAW Resonators at Cryogenic Temperatures

Maxim Goryachev, Serge Galliou, Joël Imbaud, Roger Bourquin, Bernard Dulmet and Philippe Abbé
Frequency and Time Department
FEMTO-ST Institute, UMR CNRS 6174
26 Chemin de l'Epitaphe, 25000 Besançon, France
Email: maxim.goryachev@femto-st.fr

***Abstract*— This work presents the results of investigations on different aspects of cryogenic operation of BAW resonators. For the quartz crystal resonators the losses mechanisms, the temperature sensitivity, the amplitude-frequency effect are described for liquid helium temperatures. The corresponding problems are discussed. To overcome some of these problems, the possible solution of operation at higher modes is considered. Some of these higher modes exhibit outstanding quality factors. The highest (for BAW resonators) quality factor value and quality factor-frequency product are measured for the 11th and 13th overtones respectively. In addition, two LGT resonators have been characterized in a wide temperature range. Finally, some preliminary results on utilization of cryogenic quartz resonators as a part of a frequency stabilization closed loop system are given.**

I. Introduction

The quality factor (Q-factor) of resonators and the frequency stability of frequency sources are closely related. Thus, the time and frequency community is always eager for new resonators with higher Q values of any physical nature, i.e. electromagnetic, acoustic, mechanical, etc. Bulk acoustic wave (BAW) quartz resonators at room temperatures have rather high $Q \cdot f$ product of about $15 \cdot 10^{13}$ for the frequency band of 1-100 MHz. The properties of these devices at room temperatures are widely investigated and their limitations are known [1]–[6]. Indeed, the Q-factor is not the only parameter that counts for the frequency stability or even for resonator applicability in frequency sources. Others may be frequency-temperature characteristics, nonlinear phenomena, vibration sensitivity, required environment, etc. Contrary to that, though it is well-known that quartz resonators have much higher Q-factor at cryogenic temperatures, their properties and limitations at these temperatures have not yet been deeply investigated. This work is dedicated to this problem. First, the working conditions and measurement methods for BAW resonators are discussed. Second, three vibration modes, A, B and C are considered at liquid helium temperatures for quartz SC-cut [7] BVA (Boîtier à Vieillissement Amélioré) [8] resonators optimized for working at their 3rd or 5th overtones at 5 or 10 MHz at usual temperatures. An important point here is the temperature sensitivity of this devices. Also, this section concerns amplitude-frequency and self-heating effects and associated limitations. In addition, some results for two higher overtones are given. The next section gives the results of investigation on resonators of a new synthetic material known as Lanthanum Gallium Tantalate (LGT). Finally, the frequency stability results of one possible reference system with a cryogenic quartz resonator are considered.

II. Experimental Set-up

Since the first investigations on cryogenic crystal resonators, the cryocooling technology has made a considerable progress. Thus, for example, cryostats can be substituted with cryogenerators. This substitution simplifies the cryogenic system operation and does not superposes any limitations on experiment time duration. But cryogenerators themselves introduce additional environmental disturbances, such as temperature fluctuations and vibrations mainly due to the helium pulsation at the frequency of 1.7 Hz.

A. Cryogenic System

For all the presented measurements a two-stage pulse-tube cryocooler is used (F-70H from Sumitomo Heavy Industries). The cryocooler is equipped with two shields: a vacuum chamber and an anti-radiation shield to prevent the heat exchange by means of air conduction and radiation respectively. The vacuum level inside the chamber is about 10^{-7} mBar. Both primary and secondary pumps are cut off from the chamber during measurements. Nominally, the cryocooler can absorb up to 1 W at the second stage at the temperature of 4 K. The temperature of the second stage can go down to almost 3 K if no additional thermal losses are present. The thermal losses are mainly introduced with the connecting cables needed for measurement purposes and environmental control. The resonators are inserted inside a copper block attached to the second pulse-tube stage [6].

B. Environmental Control

As it is said above, two main environmental disturbances are temperature fluctuations and vibrations. The temperature fluctuations of the cryocooler second stage due to helium pulses are about 0.3 K. The power spectral density of temperature fluctuations of the second stage in the steady state regime consists of f^{-2} region for frequencies less than about 10^{-2} Hz and f^0 region for higher frequencies. In addition, a clear spurious frequency of 1.7 Hz and its harmonics are present due to the temperature pulsation. To improve the thermal conditions passive and active instability suppression

978-1-61284-110-6/11/$26.00 ©2011 IEEE

methods are used. First, a thermal filter is installed between the second stage and a device under test (DUT). The filter reduces the temperature pulsation down to 15 mK, but increases the absolute temperature difference between these parts. Second, the active closed loop temperature regulation is implemented with a 332 Temperature Controller from Lake Shore, a CERNOX sensor and 20 Ohms heating resistor. The sensor and the actuator are installed in a copper block under the filter. The control system eliminates the f^{-2} region in the temperature fluctuation spectrum and reduces pulses down to 5 mK.

Another type of environmental disturbances is vibration. The permanent pulsation of the helium inside the pulse-tube causes DUT vibrations with a nominal amplitude of $A = 10\mu$m and frequency $f = 1.7$ Hz. The influence of the vibration on resonator parameters and noise can be reduced simply by choosing a proper angle of resonator orientation relatively to the main (vertical) axis of the vibration. No other vibration compensation system is installed for the presented measurements.

C. Measurement Techniques

The resonator characteristics are measured with a 10Hz-500MHz network analyzer HP 4195A equipped with its impedance test adapter. This method is preferable to traditional π-network method for this kind of measurements because the resonator parameters, in particular its equivalent motional resistance, change significantly with temperature. This fact results in an impedance mismatch in the π-network. The measurement capability provides a frequency resolution of 1 mHz, a phase resolution of 0.01 degree and a magnitude resolution of 0.01 dB. The minimum power of the signal source is -50 dBm. A distinctive feature of the network analyzer method is the need for compensation of long connecting cables. Indeed, the resonators are connected to the analyzer with 1 m coaxial cables passing a temperature range from 4 K to the ambient temperature. To compensate cable effects, three calibration references - the open circuit, the short cut and a 50 Ohms resistor - are put close to DUTs at the end of identical cables. Measurements are made with -40 dBm source power for 3rd and 5th overtones and -50 dBm for higher overtones.

The network analyzer is used to trace a so-called GB-plot, i.e. real and imaginary parts of the device admittance $Y(j\omega) = G(\omega) + jB(\omega)$ versus frequency $f = 2\pi\omega$ near a resonance of interest. These data are acquired by means of the GPIB connection and Labview software. After that the parameters of the traditional series $R - L - C$ lumped resonator model are optimized to fit the measured data.

III. SC-cut Quartz Resonators at Cryogenic Temperatures

Quartz BVA resonators are the subject of investigation of the present section, because it has been already shown that electrodeless resonators exhibit higher quality factors than electrode-deposited resonators [5]. More precisely the plano-convex SC-cut resonators optimized at room temperatures are chosen here as the most modern devices (active part diameter: 11 mm, thickness: 0.54 mm, curvature radius: 300 mm).

A. Resonator Losses at Cryogenic Temperatures

The main property of the devices of interest is the total losses that are inversely proportional to the device quality factor Q. The total losses (actually measured results) are a sum of all losses presented in the device, i.e. the material losses, mounting losses, electrical losses, etc. Concerning material part, this type of losses are represented in perfect material with three physical mechanisms [9]–[13]: acoustic wave-phonon dissipation (usually dominant intrinsic loss mechanism in insulating resonators), thermoelastic dissipation (pronounced in plates or rods when the wavelength of flexural vibrations is larger than the thickness of the rod or the plate), phonon-electron dissipation (usually negligible in insulators). Since the first type of the listed material losses is dominant in the devices of interest, it is considered further. The losses due to interaction of the acoustic wave and thermal phonons can be described with two distinct approaches. The first one, introduced by Akheiser [12], considers the interaction of a macroscopic strain wave in an insulator crystal with thermal phonons. Dependency of the phonon frequency on the strain leads to transport of the phonons between macroscopic regions of a crystal restoring the thermal equilibrium. This process is accompanied by acoustic wave energy loss. This approach is valid when the phonon life time is much smaller than a period of the acoustic wave, i.e. $\tau\omega \ll 1$, where τ is a thermal phonon life time and ω is an acoustic wave angular frequency. According to Akheiser approach the material $Q \cdot f$ product at high temperature is found as follows:

$$f \cdot Q = \frac{\rho V_a^2 \left(1 + (\omega\tau)^2\right)}{C_v T \gamma^2 \tau}, \tag{1}$$

where $f = 2\pi\omega$, V_a is the wave velocity, h the Planck constant, γ the Grüneisen parameter, C_v the volumetric heat capacity, T the absolute temperature, ρ the density. The second approach, proposed by Landau and Rumer [13], considers the acoustic wave as a parallel beam of interacting acoustic phonons. The interaction process involves acoustic quanta and individual thermal phonons. It is shown that the acoustic phonon attenuation is a result of three phonon interaction and is proportional to the angular frequency ω in the most studied case (also assuming that $\omega\tau \gg 1$ which is relevant at low temperature) [10]. So, the product of the quality factor and the resonance frequency is given by

$$f \cdot Q = \frac{15\rho V_a^5 h^3}{\pi^5 \gamma^2 \tau K^4 T^4}\omega, \tag{2}$$

where K is the Boltzmann constant. According to this approach the material quality factor does not depend on the angular frequency ω.

It is clear that if the acoustic wave-thermal phonon attenuation is dominant for some range of temperatures, the dependence of the losses on temperature can show which regime is valid for the considered material in given conditions. To

do so, for the first experiments, SC-cut BVA quartz resonators optimized for working on their C-mode (about 5 MHz at room temperature) are characterized between 3.8 and 20 K. The results for 5th overtones of three different modes are shown in Fig. 1.

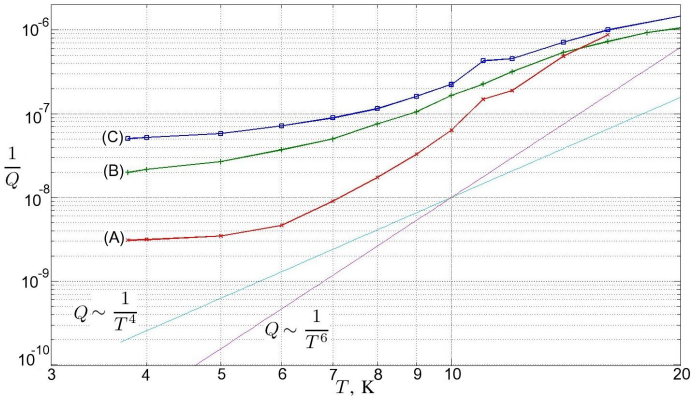


Fig. 1. Measured device losses for 5th overtones of different modes: A-mode (quasi-longitudinal), B-mode (fast quasi-shear), C-mode (slow quasi-shear).

The dependence of the losses on temperature (Fig. 1) clearly shows three regions: a plateau for temperatures less than 6 K (device losses are dominant, such as acoustic diffraction at the resonator surface or damping at the edge of the resonant disk), a linear region in the log-log scale (attributed to acoustic wave - phonon losses) and the second plateau for temperatures about 20 K (region of the maximum absorption [4]). The medium region clearly shows the dependency of the type $Q^{-1} = aT^n + constant$, where n is close to 4 for quasi-shear modes and to 6 for a quasi-longitudinal mode. Though for the latter mode, the dependence does not strictly confirms the slope predicted by the Landau-Rumer theory, the corresponding losses regime is actually dominant in quartz material at the liquid helium temperature. The increase of the power n has been already observed [10], [11].

The highest measured Q-values among 5th overtones of three modes are $325 \cdot 10^6$ for the A-mode, $50.5 \cdot 10^6$ for the B-mode, $1.99 \cdot 10^6$ for the C-mode. This sequency corresponds to the list of modes sorted out from the best trapped mode to the less trapped one according to the Stevens and Tiersten model [14]. Among all modes of 3rd and 5th overtones the best result is achieved for the 5th overtone of the A mode with $Q \cdot f$ product of about $5.1 \cdot 10^{15}$ which is the highest result in comparison with previous investigations [6]. But already starting at this stage, the problems of practical utilization arise. Since at cryogenic temperatures the resonators have very low values of equivalent motional resistance (e.g. about 3 Ohm for the 5A mode), minimization of the active load is required in order to preserve the high quality factor. This considerably complicates the choice of signal sources and amplifiers.

B. Temperature Sensitivity

Another important characteristic of a resonator is its temperature sensitivity. Indeed, to minimize the influence of the environment on the frequency source, the sensitivity of the resonance element has to be minimized. Fig. 2 shows the fractional frequency temperature sensitivity for the three modes considered in the previous subsection.

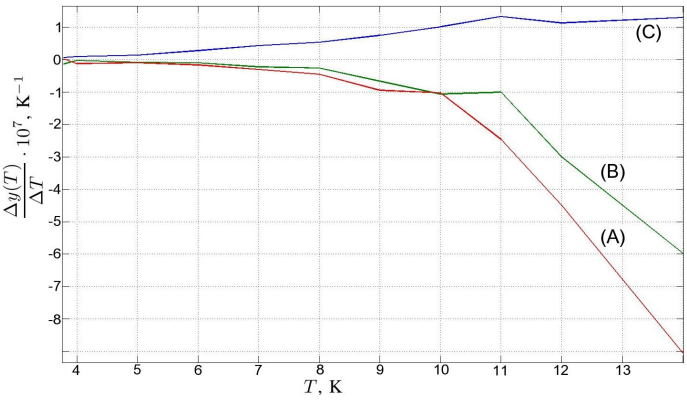


Fig. 2. Fractional frequency temperature sensitivity: A, B and C-modes.

As it is clearly seen from Fig. 2 the investigated modes does not have any turn-over point. The absence of a turn-over point poses serious problems in terms of crystal temperature stabilization. In general, the smaller sensitivity corresponds to lower temperatures for these devices. At 4 K the sensitivities are $2 \cdot 10^{-9}$ K^{-1} for the A-mode, $3 \cdot 10^{-9}$ K^{-1} for the B-mode and $1 \cdot 10^{-8}$ K^{-1} for the C-mode. These results suggest that to obtain a desired fractional frequency stability of 10^{-13}, in best case, the temperature has to be controlled within 0.5 mK.

C. Amplitude-Frequency Effect

BAW resonators are nonlinear devices in nature. Resonator properties considerably depend on the power level of the excitation signal. This dependence becomes even more pronounced with the increase of the quality factors, since the resonance frequency shift is proportional to the product of the dissipated power and Q. To characterize this effect in cryogenic temperatures, the impedance of the resonator under test is measured with different levels of an excitation signal for the 5th overtone of the A-mode (that with the highest Q). The results are shown in Fig. 3.

Fig. 3 suggests that already for the dissipation power of $P = 0.22\mu$W (Curve (2)) the impedance curve distortion exists. Furthermore, for $P = 1.6\mu$W the resonator self-heating effect becomes apparent. Also, it can be seen that for an increasing power P, the equivalent motional resistance also increases. These effects can even lead to temperature-amplitude oscillations. These oscillations are characterized by two states of low and high dissipation powers corresponding also to higher and lower values of motional resistance. In addition, the higher driving signal levels degrades resonator noise characteristics [15].

The strong amplitude-frequency and self-heating effects arise another problem concerning effective utilization of cryogenic quartz crystal resonators in actual applications. Indeed, in order to avoid this effect the driving signal has to be kept

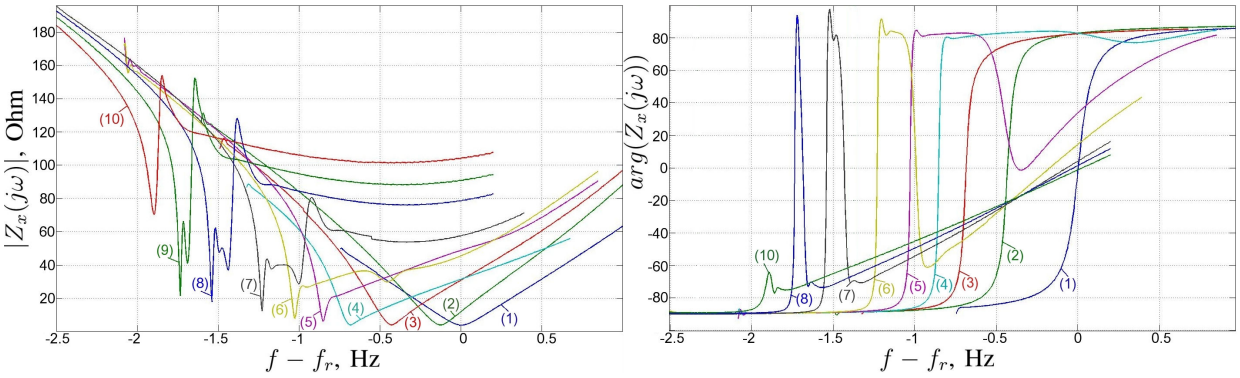


Fig. 3. Impedance of the resonator near the A-mode on 5th overtone for different dissipation powers: (1) $P = 0.09\mu$W, $I = 0.154$mA (2) $P = 0.22\mu$W, $I = 0.237$mA, (3) $P = 0.42\mu$W, $I = 0.324$mA (4) $P = 0.64\mu$W, $I = 0.403$mA, (5) $P = 1.6\mu$W, $I = 0.479$mA (6) $P = 2.26\mu$W, $I = 0.538$mA, (7) $P = 4.87\mu$W, $I = 0.624$mA (8) $P = 9.20\mu$W, $I = 0.719$mA, (9) $P = 12.35\mu$W, $I = 0.757$mA (10) $P = 17.20\mu$W, $I = 0.494$mA.

very low, but this fact degrades the system signal-to-noise ratio as well as it requires additional amplification stages. This problem is even worse lower values of the resonator equivalent motional resistance.

D. Higher Overtones

In order to deeply investigate the abilities of devices under tests, some higher resonator harmonics have been characterized at cryogenic temperatures. Thus, Fig. 4 demonstrates the presence of different anharmonics of the 11th overtone of the A-mode. The corresponding indexes for identifying the anharmonics are marked on the figure. The best anharmonic mode has a quality factor of $417 \cdot 10^6$ at the frequency of 68823167 Hz with an equivalent resistance of 19 Ohms. The estimated $Q \cdot f$ product for this resonance equals to $2.87 \cdot 10^{16}$ Hz. Thus, the DUT working at this resonance at cryogenic temperatures has the highest reported quality factor ever measured on BAW resonators.

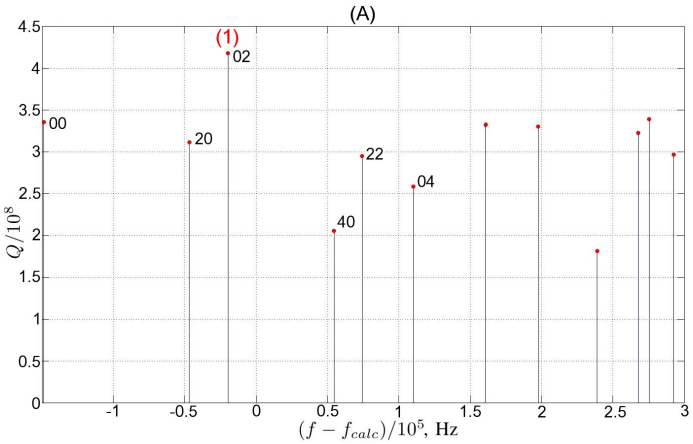


Fig. 4. Different anharmonics of the 11th overtone of the the A mode ($f_{calc} = 68.842620$ MHz).

In the same way the 13th overtone of the A-mode is characterized. Here only three anharmonics are found. The best of them has a Q value of $377 \cdot 10^6$ at 81392723 Hz with an equivalent resistance of 30 Ohms. Thus, the highest $Q \cdot f$ measured on BAW resonators can reach $3.07 \cdot 10^{16}$Hz

So, these modes could be perspective for a practical utilization. But the issue of the selection or rejection of the undesired modes should be solved first in autonomous systems such as oscillators. Also, higher Q values result in a stronger amplitude-frequency effect [16].

IV. LGT Resonators at Cryogenic Temperatures

The LGX material family is promising for designing BAW resonators with high quality factors. Among this materials the Lanthanum Gallium Tantalate (LGT) and Lanthanum Gallium Silicate (LGS) are the most commonly known. Nevertheless, their properties have not been yet investigated for temperatures lower than 80 K. In this section, investigation results on BVA-type plano-convex Y-cut LGT resonators (active part diameter: 10.2 mm, thickness: 0.65 mm, curvature radius: 100 mm) are presented. The attempts to characterize their LGS counterparts have been abandoned since they show a decrease of quality going from room to cryogenic temperatures.

A. LGT Resonator Losses

The characteristics of two LGT resonators of similar origin have been measured in the wide temperature range $(4-80$ K). The results for their quality factors are shown in Fig. 5. The maximum Q-factor is just about $7.2 \cdot 10^6$ that is only 5 times higher than its value at room temperature. Unlike in the case of quartz resonators, device losses dependence on temperature seem to be rather uniform. The only physical phenomenon clearly seen is the peak of losses at about 40 K. It can be attributed to relaxation processes due to the presence of impurities. In addition, a very important difference between two resonators is observed. These results suggest that the LGT resonator technology is far from being optimized in terms of different types of losses: device losses cover material losses. On the other hand, further technology development might lead to considerable increase of their quality factors in cryogenic conditions.

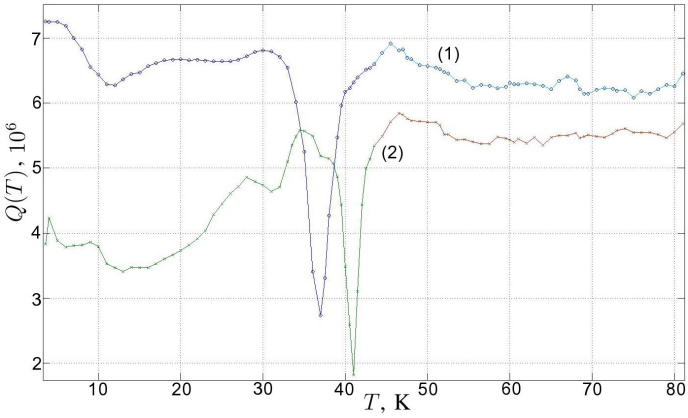


Fig. 5. LGT resonator quality factor Q as a function of temperature T (two resonators).

B. LGT Resonator Temperature Sensitivity

The dependence of the fractional stability on the temperature is also examined for the LGT resonators (see Fig. 6). Both resonators have a turn-over point between 6 and 7 K which is advantageous for creating a high quality temperature stabilization system.

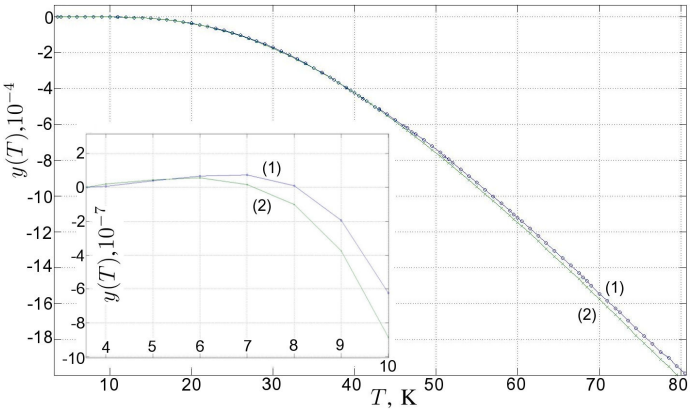


Fig. 6. Fractional frequency y versus temperature T for two LGT resonators.

V. Frequency Stabilization System Based on Cryogenic Quartz Resonator

This section shows the preliminary results on a system based on a quartz resonator working at liquid helium temperatures. For this purpose a frequency stabilization system based on an external high quality resonator is implemented [17]. The general scheme of this system is shown in Fig. 7. The system consists of a voltage-controlled oscillator used as a reference for a frequency synthesizer whose output signal frequency corresponds to the resonant frequency of the BAW device. The resonator is put into a so-called π-network implemented at cryogenic conditions and connected with the long cables with the other components. The signal filtrated by the resonator is compared with the original one (also passed through the cryogenerator to compensate the parasitic effects of cables) with a double balanced mixer. The resulting signal is filtrated and used to correct the frequency variation of the VCO. In fact, only the resonator π-network and the through connection for the original signal channel are placed inside the cryogenerator.

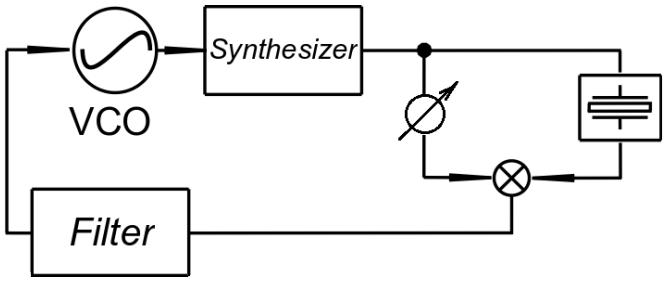


Fig. 7. External frequency discriminator system.

The system is designed for working on the resonator 5th overtone of the A-mode preserving a loaded Q of only $16{\cdot}10^6$. The frequency stability results in terms of Allan deviation are shown in Fig. 8. This figure compares the results measured for the stand-alone voltage-controlled oscillator and the closed loop system after different periods of continuous cryogenic system operation. The stability improvement is clearly seen for the long averaging times τ, though for short τ the stability is degraded due to internal loop noise sources. Comparing two cases of closed loop system, the 12-hours case (measurements are made within first 12 hours of cryogenerator permanent operation) exhibits a τ^{+1} slope of the Allan deviation that corresponds to the linear frequency drift. The 60-hours case (measurements after at least 60 hours of operation) uncover a $\tau^{+1/2}$ behavior that is a result of the VCO f^{-4} or cryogenic resonator f^{-2} phase noises. Both types of noise can be attributed to resonance frequency sensitivity to the temperature random-walk noise [18]. These results suggests the presence of very long time constants in the system (probably due to thermal effects).

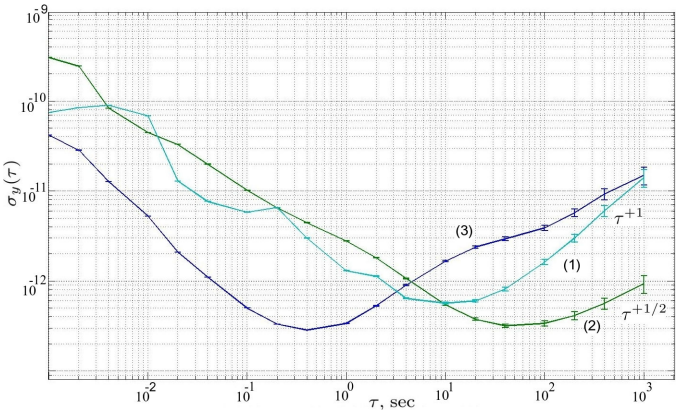


Fig. 8. Allan deviation of the closed-loop frequency stabilization system measured at first 12 hours of cryogenerator operation (1), measured after at least 60 hours of cryogenerator permanent operation (2) and compared with a stand-alone oscillator (3).

Though experimental conditions are not so ideal, tests with a passive reference as well as an oscillator based on the

presented devices are presently in progress. The aim of these tests is to achieve the frequency stability corresponding to the measured values of the quality factor.

VI. Conclusion

Though quartz resonators can achieve extraordinary values of quality factors at cryogenic conditions, deeper investigations show the presence of severe problems concerning practical utilization of these devices. The difficulties come from the small values of equivalent motional resistances, the strong amplitude-frequency effect, the self-heating, the absence of a temperature turn-over point. Some of these disadvantages could be overcome for higher overtones, such as 11th or 13th, exhibiting even higher Q values. The LGT resonators are still perspective candidates for cryogenic operation, since the device losses overshadow material losses. The preliminary investigations on the passive frequency stabilization system based on cryogenic BAW resonators show rather moderate results due to a non-optimized environment, internal noise sources and a reduced quality factor.

Acknowledgment

This work is supported by Conseil Régional de Franche-Comté (Convention No. 2008C_16215). The authors wish to thank the council for its grant and financial support as well as the staff of the Time and Frequency Department of FEMTO-ST Institute, Besançon, France for the help and fruitful discussions.

References

[1] N. Smagin, "Quartz crystal resonators with quality factors of about $120 \cdot 10^6$ at 2K," *Izmeritel Tekh. SSSR*, no. 9, 1960.

[2] G. Mossuz, "Study and achievement of a quartz crystal oscillator at very low temperature," *UFC PhD Thesis*, July 1975.

[3] G. Robichon, J. Groslambert, J.J. Gagnepain, "Frequency stability of quartz crystal at very low temperatures: preliminary results," in *38th Annual Freq. Contr. Symp.*, 1984, pp. 201-205.

[4] A. El Habti, "Study of resonators and oscillators at very low temperatures," *UFC PhD Thesis*, Jan. 1993.

[5] J. J. Suter, "Acoustic loss phenomena in alpha quartz resonators over the $1.4 - 77$K temperature range," *J. appl. Phys.*, vol. 63, no. 11, pp. 5594-5595, June 1988.

[6] S. Galliou, J. Imbaud, R. Bourquin, N. Bazin, P. Abbé, "Outstanding Quality Factors of Bulk Acoustic Wave Resonators at Cryogenic Temperature," in *Proc. EFTF 2008*, Toulouse, France, April 2008.

[7] E. P. Eernisse, "Quartz Resonator Frequency Shifts Arising from Electrode Stress," *Proc. 29th Ann. on Freq. Contr. Symp.*, Fort Monmouth, NJ, pp. 1-4, 1975.

[8] R. J. Besson, "A new electrodeless resonator design," *Proc. 31th Ann. on Freq. Contr. Symp.*, Fort Monmouth, NJ, pp. 147-152, 1977.

[9] A. El Habti, F. Bastien, "Low temperature limitation on the quality factor of quartzresonators," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 41, no. 2, pp. 250-255, 1994.

[10] H. J. Maris, "Interactions of sound with thermal phonons in dielectric crystals," *Physical Acoustics*, vol. 5, pp. 279-345, Academic Press, 1968.

[11] M. F. Lewis and E. Patterson, "Microwave Phonon-Attenuation Measurements in Quartz, *Phys. Rev.*, 19, 3, pp. 703-711, 1967.

[12] A. Akheiser, "On the absorption of sound in solids, *J . Phys. USSR*, no.1, p. 277, 1939.

[13] L. Landau and G. Rumer, "Absorption of sound in solids," *Phys. Z. Sowjetunion*, 11, 18, 1937.

[14] D. S. Stevens and H. F. Tiersten, "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variations," *J. of the Amer. Stat. Assoc.*, vol. 79, no. 6, pp. 1811-1826, 1986.

[15] J. J. Gagnepain, M. Oliver, F. L. Walls, "Excess Noise In Quartz Crystal Resonators," in *Proc. 37th Ann. Sym. on Freq. Contr.*, pp. 218-225, 1983.

[16] R. Bourquin, D. Nassour, D. Hauden, "Amplitude frequency effect of SC-cut quartz trapped energy resonators ", *Proc. 36th Ann. on Freq. Contr. Symp.*, Philadelphie, USA, pp. 200-207, 1981.

[17] C. McNeilage, E. A. Ivanov, P. R. Stockwell, J. H. Searls, "Review of feedback and feedforward noise reduction techniques," in *Proc. IEEE Inter. Freq. Contr. Symp.*, p. 196, 1998.

[18] A. Ballato, J. R. Vig, "Static and Dynamic Frequency-Temperature Behaviour of Singly and Doubly Rotated, Oven-Controlled Quartz Resonators," in *Proc. Ann. on Freq. Contr. Symp.*, pp. 181-188, 1978.

Nonlinearities for Parametric Pumping of Quartz UHF Oscillators

R.L. Kubena[1], Y-K Yong[2], D.J. Kirby[1] and R.J. Joyce[1]
[1]HRL Laboratories, Malibu, California, U.S.A.
[2]Rutgers University, Piscataway, New Jersey, U.S.A.

***Abstract*— The feasibility of a parametric amplification method for improving the *Q* of UHF quartz oscillators is considered. A fundamental thickness shear resonator operating at about 500 MHz was studied for estimating the magnitude of fractional change in stiffness, $\Delta k/k$ that could be obtained at resonance and off-resonance. The $\Delta k/k$ is employed in a parametric pumping of the thickness shear mode of vibration. A relationship between the $\Delta k/k$ at resonance and $\Delta k/k$ at off-resonance was derived. For a 532 MHz fundamental thickness shear resonator, a $\Delta k/k$ = $3.8\text{x}10^{-4}$ /V at resonance, and a $\Delta k/k$ = $1.6\text{x}10^{-7}$ /V at off-resonance was found. The off-resonance data compares well with measured data of $\Delta k/k$ = $1.7\text{x}10^{-7}$ /V. Our off-resonance study of $\Delta k/k$ established that it is independent of the sign of electric potential drive that is $\Delta k/k$ is a rectified excitation and therefore appears at twice the excitation frequency.**

The parametric amplification phenomenon is governed by the Mathieu equation. MATLAB Simulink models of the Mathieu equation were developed to establish the baseline criteria for parametric amplification to improve the resonator *Q*. The resonator model was excited at a frequency ω_A over a range that included the natural frequency ω_0 and the third overtone ω_3. The parametric drive frequency ω_P was set equal to $2\omega_A/n$ (n=1, 2, 3, ….). The model results showed that for parametric drive frequencies of $\omega_P = \omega_A$ (for fundamental mode operation), and $\omega_P = \omega_A/3$ (for third overtone operation), the $\Delta k/k$ needed for parametric amplification is in the range of 0.001 to 0.003 for a resonator *Q* of 9,000 to 15,000. It was observed that there was shift of the resonance frequency with the parametric amplification resulting from the change in dc stiffness $\Delta k/k$ of the thickness shear mode.

I. INTRODUCTION

It is well known that flicker phase modulation (PM) noise in the sustaining circuit of oscillators causes flicker frequency modulation (FM) noise to the oscillator output frequency given by $L_{osc}(f_f) = L_{ckt}(1\ Hz)f_f^2/(4f_f^3Q_L^2)$ [1]. Thus, increasing the loaded Q_L can help in reducing the phase noise. Previous reports have shown that parametric amplification can be used to increase the *Q* of silicon nano-resonators [2-4]. However, to date very few studies have been made for shear-mode quartz resonators. In this paper, we investigate whether parametric amplification is feasible for improving the Q of UHF quartz resonators.

Sponsored by DARPA Microsystems Technology Office (MTO).
Program: Dynamics-Enabled Frequency Sources (DEFYS).
Issued by DARPA/CMO under Contract No. HR0011-10-C-0109.

Fundamental and third overtone thickness shear-mode resonators are investigated to establish the criteria needed for parametric amplification. Specifically the fractional change in stiffness $\Delta k/k$ in the Mathieu equation that governs parametric amplification and the resonator Q are considered. MATLAB Simulink models for the Mathieu equation were developed to study the relationships between the parametric frequencies, $\Delta k/k$, and the resonator *Q*. A study of the $\Delta k/k$ off-resonance was needed to establish the magnitude of the change in stiffness for given applied voltage levels.

II. MECHANICAL NONLINEARITIES IN QUARTZ

Quartz is known to have small mechanical nonlinearities in its elastic stiffness. These nonlinearities in quartz resonators are exhibited in well documented phenomena such as their frequency-temperature [5, 6], acceleration sensitivity [7, 8] and drive level dependency [9, 10]. The nonlinear mechanical effects are governed by the piezoelectric equations with finite strains and nonlinear elastic constants as given below in Equations (1) – (3):

A. Strain-displacement, and electric field-potential relations

$$S_{ij} = \frac{1}{2}(u_{j,i} + u_{i,j} + u_{k,i}u_{k,j} + u_{k,j}u_{k,i})$$
$$E_i = -\phi_{,i} \tag{1}$$

Here S_{ij} is the mechanical finite strain for the relatively large mechanical displacements u_i. E_i and φ are respectively the electric field and electric potential.

B. Constitutive relations

$$T_{ij} = (C_{ijkl} + C_{ijklmn}s_{mn} + C_{ijklmnop}s_{mn}s_{op})s_{kl} + \eta_{ijkl}\dot{s}_{kl} - e_{kij}E_k$$
$$D_i = e_{ijk}s_{jk} + \varepsilon_{ik}E_k \tag{2}$$

where T_{ij}, C_{ijkl}, C_{ijklmn}, $C_{ijklmnop}$, η_{ijkl}, e_{kij}, D_i, ε_{ik} are respectively the mechanical stress, linear elastic constants, third order nonlinear elastic constants, fourth order nonlinear

elastic constants, viscosity constants, piezoelectric stress constants, electric displacements, and dielectric permittivity. The values of the material constants of quartz are well known and could be found in the literature, except for the fourth order nonlinear elastic constants. The fourth order elastic constants are needed for modeling the Duffing effect at very high drive voltages. Since most of the fourth order elastic constants are not well known, we neglect the fourth order elastic constants in our present models.

C. *Equations of motion and electrostatics*

$$\begin{aligned} (T_{ij} + T_{jk}u_{i,k})_{,j} + \rho b_i &= \rho \ddot{u}_i \\ D_{i,i} &= 0 \end{aligned} \tag{3}$$

The equations of motion and electrostatics are for any piezoelectric volume V where for boundary conditions on the surface S enclosing the piezoelectric volume we specify surface traction $P_i = n_j(T_{ij} + T_{jk}u_{i,k})$ or mechanical displacements u_i, and electric displacements D_i or electric potential φ.

III. Nonlinear Piezoelectric Model of an AT-cut Quartz Plate

The governing equations (1) - (3) along with the necessary material constants for the AT-cut quartz were implemented in a COMSOL finite element tool. Figure 1 shows a model of the 532 MHz fundamental thickness shear AT-cut quartz resonator. The mode shape of the energy trapped fundamental thickness shear mode is shown in Figure 2.

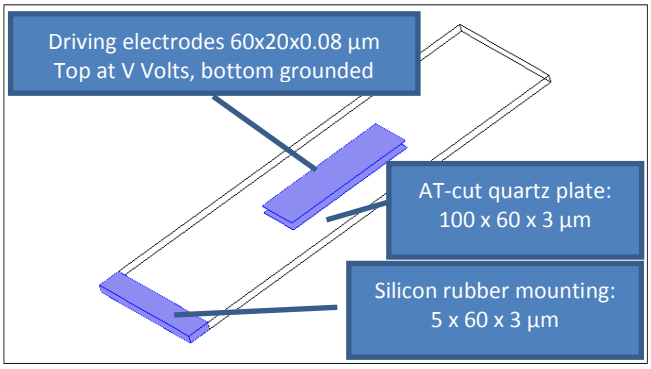


Figure 1. Model of a 532 MHZ AT-cut quartz plate resonator with dimensions for electrodes and mounts.

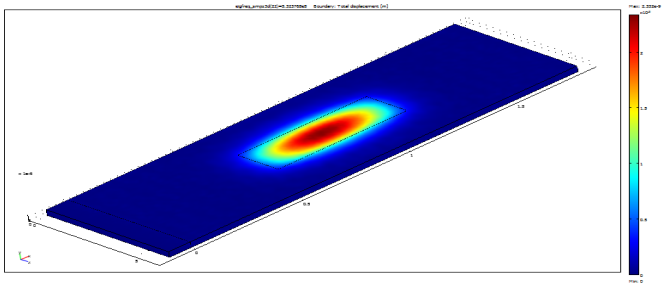


Figure 2. Mode shape of the energy trapped fundamental thickness shear mode in a 532 MHz AT-cut quartz plate resonator.

A. *Drive level dependency of the thickness shear mode resonator*

The effects of nonlinear elastic stiffness are observed in the phenomenon of drive level dependency, that is, the resonant frequency changes with the drive level of electric potential at the top electrode of quartz plate in Figure (2). Figure 3 shows the frequency response of the AT-cut plate resonator at two drive levels: 1 mV and 100 mV. A change in resonant frequency of 18.8 ppm at 100 mV was observed in the data. The resonant frequency is given by:

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{4}$$

where k is the modal stiffness, and m is the modal mass. Here a small frequency change with drive level electric potential is due to a small change in the modal stiffness Δk. Hence a fractional change in frequency can be related to a fractional change in modal stiffness, and using a binomial expansion:

$$\frac{\Delta f}{f} \cong \frac{1}{2}\frac{\Delta k}{k} \tag{5}$$

Therefore the fractional change in modal stiffness $\Delta k/$k is 37.6 ppm at 100 mV, or with normalization to 1 V,

$$\frac{\Delta k}{k} = 3.76x10^{-4}/\mathrm{V} \tag{6}$$

for the 532-MHz fundamental thickness shear mode of vibration of the AT-cut quart plate resonator represented in Figures (1) and (2).

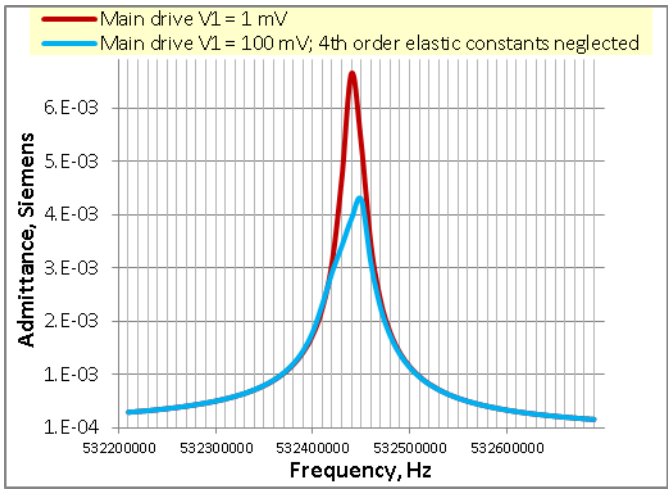


Figure 3. Frequency deviation versus AC electric potential for a 532 MHz AT-cut quart plate resonator.

B. Nonlinear effect of static electric potential on the shear strain in the AT-cut quartz plate resonator

The parametric amplification of a vibration mode in a mechanical resonator requires that the modal stiffness k of the vibration mode can be varied by a pumping excitation; that is k is nonlinear of the form:

$$k(1+\frac{\Delta k}{k}\sin(\omega_P t)) \tag{7}$$

where ω_P is the frequency of the parametric pumping excitation and t is time. Here $\Delta k/k$ is the fractional change in modal stiffness due to the parametric excitation.

For the thickness shear mode of Figure (2), and a 1 V AC applied at the plate resonance frequency, Eqn. (6) gives $\Delta k/k = 3.76\text{x}10^{-4}$. A more general description for parametric excitation may not involve pumping at a mechanical resonance; we therefore first investigated the $\Delta k/k$ at frequencies not near the resonance frequency, that is, at "off-resonance" frequencies, or at DC electric potentials. In this case, we also investigated whether $\Delta k/k$ is dependent on the sign of the drive level electric potential.

The thickness shear mode of vibration in the AT-cut quartz plate is governed predominantly by the shear strain S_{xy} (or S_{12}); therefore a study of the effects of DC electric potential at the top electrode (Figure 1) on the shear strain S_{xy} is useful. Figure 4 shows the COMSOL finite element results of positive and negative electric potential, respectively, on the shear strain S_{xy} at the midpoint of the quartz plate. Two second order regression curve fittings were performed on the COMSOL data: one for the positive electric potential, and the other for the negative electric potential. Two regression equations are shown on the graph, one for positive electric potential and the other for negative electric potential. We note that while the slope of both regression equations is the same, the coefficient of the nonlinear term changes sign with the electric potential. This change in sign indicates that the $\Delta k/k$ for the shear strain S_{xy} is *not* dependent on the sign of the drive level electric potential and that it is a rectified function. The $\Delta k/k$ modulation then appears at twice the excitation frequency.

The regression equations can be written in the form:

$$V = 1.76x10^6 S_{xy}(1 \pm 0.126 S_{xy}) \tag{8}$$

Since the shear strain S_{xy} in eqn. (8) is approximately $5.7\text{x}10^{-7}$ /V, we calculate that

$$\frac{\Delta k}{k} = 0.126 S_{xy} = 7.16x10^{-8}/V \tag{9}$$

C. Quality factor Q of a mode of vibration, dynamic Δk/k versus static Δk/k

If a thickness shear resonator is represented by a single degree of freedom in shear strain S_{xy}, then the ratio of the dynamic shear strain at resonance at drive V to its static shear strain at DC V is equal to its quality factor Q, that is:

$$\frac{S_{xy}\big|_{dynamic}}{S_{xy}\big|_{static}} = Q \tag{10}$$

where $S_{xy}\big|_{dynamic}$ is the peak shear strain at resonance for a given electric potential V_{AC} drive, and $S_{xy}\big|_{static}$ is the static shear strain at the same electric potential, V_{DC}.

Since Eqn. (9) shows that the $\Delta k/k$ is linearly dependent on S_{xy} we could use Eqn. (10) to relate the $\Delta k/k$ of Eqn. (6) that was obtained at peak resonance to a $\Delta k/k$ at "off-resonance". For a given electric potential, the $\Delta k/k$ at peak resonance is equal to Q times the $\Delta k/k$ "off-resonance":

$$\frac{\Delta k}{k}\bigg|_{dynamic} = Q\frac{\Delta k}{k}\bigg|_{static} \tag{11}$$

The Q for the fundamental thickness shear mode in Figure 3 was 24000 which upon substitution into Eqn. (11) we can calculate the equivalent static $\Delta k/k$ from the dynamic $\Delta k/k$ of Eqn. (6):

$$\frac{\Delta k}{k}\bigg|_{static} = 1.57x10^{-7}\ /\text{V} \tag{12}$$

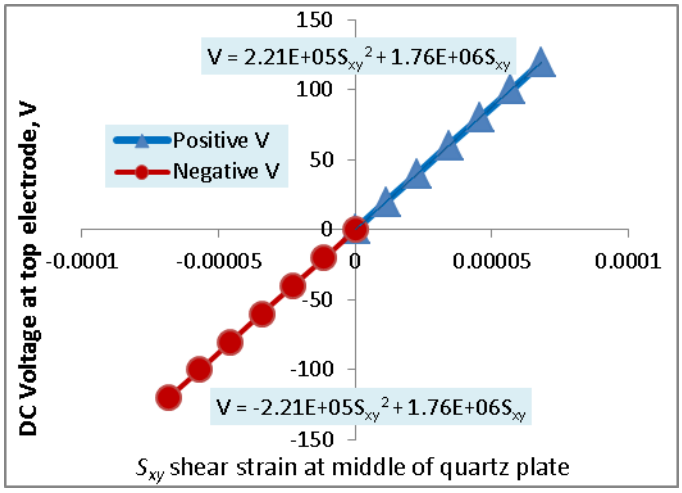


Figure 4. Shear strain s_{xy} at the midpoint of the AT-cut plate versus the static electri potential.

The static $\Delta k/k$ in Eqn. (12) is about two times larger than the static $\Delta k/k$ in Eqn. (9). This agreement is reasonable because the static $\Delta k/k$ of Eqn.(9) is for the shear strain S_{xy} at one point in the plate while the static $\Delta k/k$ of Eqn.(12) is for the thickness shear mode that is predominantly governed by the shear strain S_{xy} distributed across the thickness of the plate.

Equation (11) is useful for estimating the $\Delta k/k$ for parametric excitations at "off-resonance" frequencies. In designing experiments for parametric amplification of the thickness shear mode, the equation is useful for determining the feasibility of parametric pumping at "off-resonance" frequencies.

IV. Mathieu Equation, and Parametric Amplification of Mechanical Resonators

We discuss Mathieu equation in terms of a single degree of freedom resonator in Equation (13). We could modify the spring stiffness k to a nonlinear spring stiffness $k[1+(\Delta k/k)\sin(\omega_P t)]$ so that the governing equation for the resonator is

$$\frac{F}{m}\sin(\omega_A t) = \ddot{u} + \frac{\omega_0}{Q}\dot{u} + \omega_0^2(1+\frac{\Delta k}{k}\sin(\omega_P t))u \qquad (13)$$

where F is the force on the mass m, u is the displacement, and Q is the quality factor. Here ω_A is the activation frequency, while ω_P is the parametric frequency. Equation (13) is the Mathieu equation that yields parametric amplification for parametric frequencies [11]

$$\omega_P = 2\omega_0 / n \text{ for } n = 1,2,3,.... \qquad (14)$$

A. *MATLAB Simulink model of the Mathieu equation*

A MATLAB Simulink model for the Mathieu equation was developed to study the effects of parametric amplification on the Q of a resonator. Figure 5 shows a MATLAB Simulink model of the Mathieu equation. Figure 6 shows sample input acceleration to the model (top graph), the "rectified" parametric excitation (middle graph) using a ω_0 drive (n = 1), and the sample output displacement Z (bottom graph).

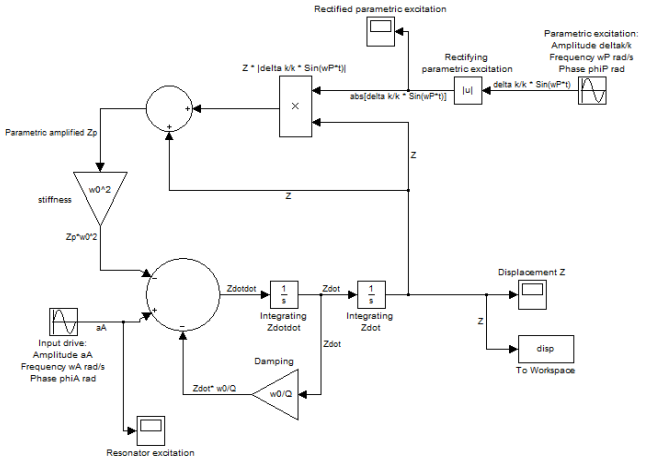


Figure 5. MATLAB Simulink model of the Mathieu equation.

B. *Effects of parametric amplification on the Q of a resonator*

We study the effects of parametric amplification on the Q of a resonator using various values of $\Delta k/k$ and two modes of pumping: (a) pumping the fundamental using a ω_0 drive (n = 1), and (b) pumping the third overtone using a ω_0 drive (n = 3).

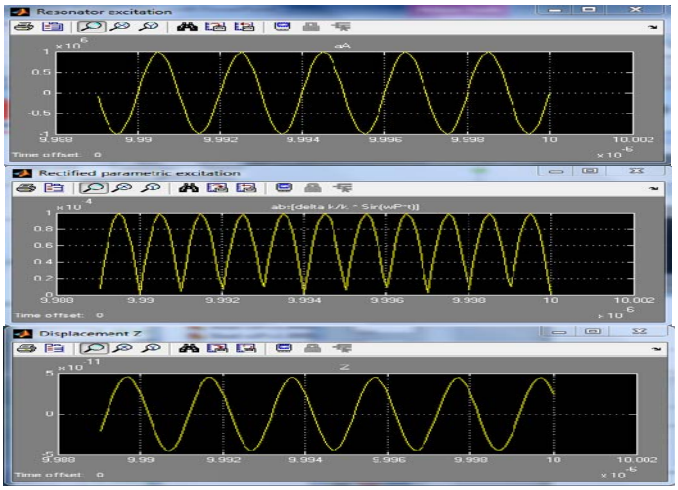


Figure 6. Sample input acceleration (top), parametric excitation (middle), and output displacement (bottom) for the MATLAB Simulink model of the Mathieu equation. ω_0 drive (n = 1).

Figure 7 shows the frequency response of the resonator with a natural frequency ω_0 equal to 500 MHz and an initial Q equal to 15000. In this case, the parametric pumping frequency ω_P is set equal to the activation frequency ω_A. This is equivalent to the pumping of the fundamental thickness shear mode using a ω_0 drive (n = 1). We observe that the starting Q = 15,000 of the resonator could be increased to Q = 70,000 for a $\Delta k/k$ = 0.001.

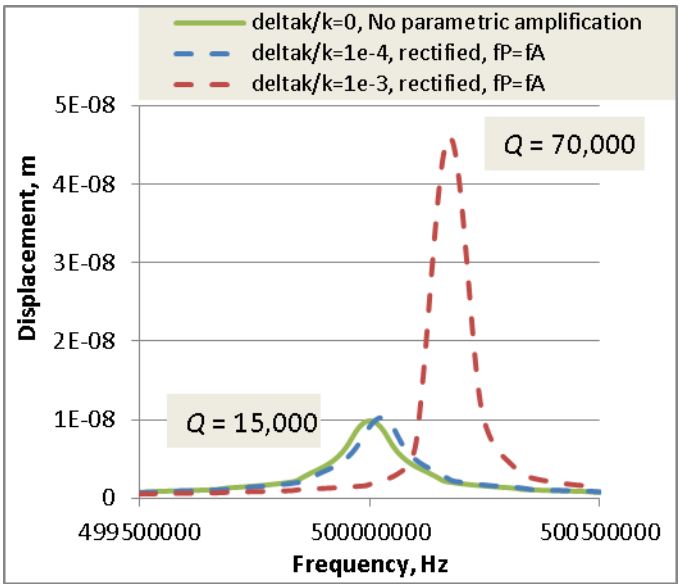


Figure 7. Simulink model result of Mathieu equation with parametric amplification, $\omega_P = \omega_A$ and n = 1.

For our thickness shear resonator when n = 1, there will be parametric self pumping since the excitation force F will yield a $\Delta k/k$ that is "rectified" so that the parametric frequency ω_P is effectively $2\omega_0$. However such parametric self-pumping is

likely limited by the "Duffing" effect if higher order nonlinear terms are included in the modeled dynamics.

Figure 8 shows the frequency response of the resonator with a natural frequency ω_0 equal to 1500 MHz and a Q equal to 9000. Here the parametric pumping frequency ω_P is set equal to one third of the activation frequency ω_A. This is equivalent to the pumping of the third overtone thickness shear mode (~1500 MHz) using a fundamental thickness shear mode (500 MHz) ω_0 drive (n = 3). This will require a resonator design with the third overtone frequency $\omega_3 \sim 3\omega_0$.

We observe that the resonator Q could be increased from 9,000 to Q = 13,000 for a $\Delta k/k$ = 0.003. It was observed that in both Figs. 7 and 8, there was considerable shifting of the resonance frequency; however this frequency shift could be due to the rectified nature of the fractional change in stiffness $\Delta k/k$ of the mode of vibration.

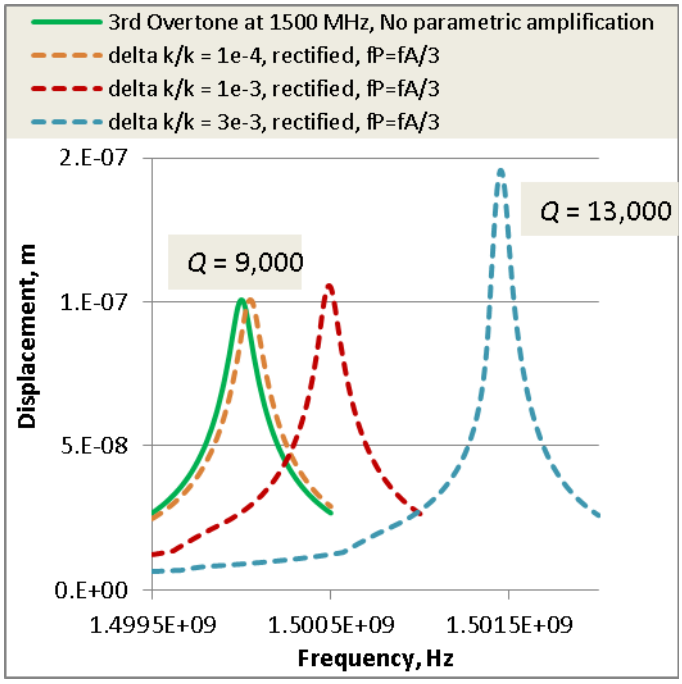


Figure 8. Simulink model result of Mathieu equation with parametric amplification, $\omega_P = \omega_A/3$, n = 3

V. Experimental Results for Nonlinear Shiftness Modulations of Quartz Thickness Shear Resonators.

AT-cut quartz shear mode resonators operating at 507-MHz were used to estimate the magnitude of the fractional change in stiffness, Δk/k, obtained under non-linear drive conditions and also to investigate parametric amplification effects.

A typical admittance measurement involves a single swept frequency source while monitoring the device current and voltage. The magnitude of the admittance Y of the device is then $|Y| = |I/V|$. This measurement can readily be performed with a single source Agilent admittance meter. However, measurement of the admittance of a resonator driven simultaneously with two different frequencies presents a challenge in that the typical admittance meter can no longer provide the dual excitation frequencies. Therefore, an experimental set-up was designed and built specifically to accommodate the dual frequency inputs and provide admittance information of the device-under-test.

A schematic of the experimental set-up for the non-linear drive studies is presented in Figure 9. In this case, the quartz resonator is excited by an input derived from a Holzworth HS300 2A 2-channel synthesizer. The Holzworth synthesizer contains a single OCXO crystal oscillator and provides two drive signals of differing frequency, amplitude and phase. These two drives are phase coherent in that they are driven by the same internal reference which eliminates any drift between them. These signals are swept in frequency and summed in a power combiner to serve as the excitation for the resonator. An Agilent DSO9404A digital storage oscilloscope with differential active probes monitors the voltage and current associated with the resonator and the device admittance Y is computed through post processing. The voltage across the resonator is measured with the first active probe and the current is derived by monitoring the voltage across a 10Ω sense resistor using the second active probe.

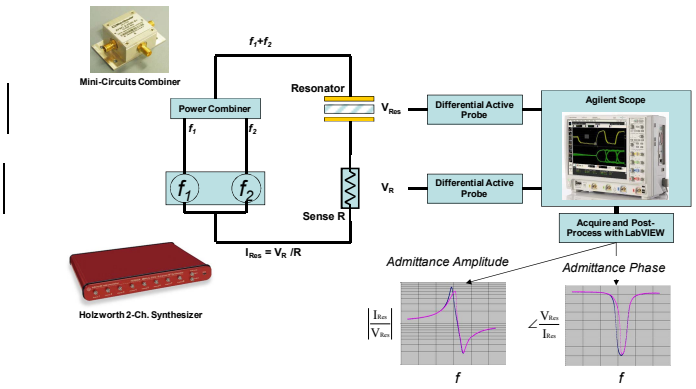


Figure 9. Schematic of the experimental set-up for parametric amplification studies.

A photograph of the experimental set-up (Figure 10) shows the power combiner, the quartz resonator and sense resistor mounted and interrogated by the two differential active probes, along with two directional couplers which are included to provide isolated monitoring of the Holzworth synthesizer outputs. The system permits both on- and off-resonance studies of the non-linear effects in the quartz resonator. We are also able to investigate various configurations of excitation and parametric pump frequencies including the n=1 and n=3 conditions. Additionally, replacing the Holzworth synthesizer with a dc bias voltage source allows static $\Delta k/k$ value estimates.

This new configuration was tested by exciting the resonator around the fundamental mode with a low drive voltage. Figure 11 shows the measured admittance response of an AT-cut UHF 507 MHz quartz resonator with Q ~ 21K. This validates the experimental procedure and illustrates the basic device natural frequency (ω_0) response without non-linear effects.

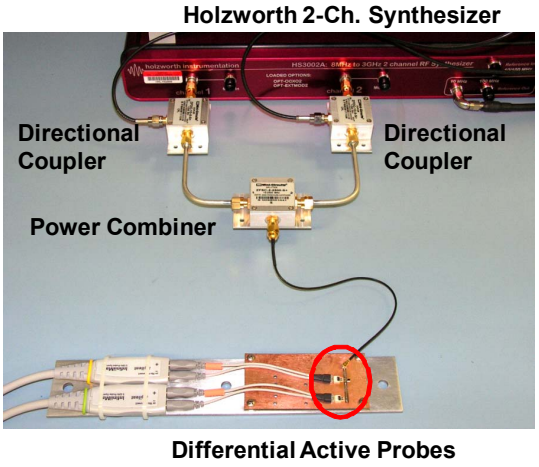


Figure 10. Photograph of the 507 MHz quartz resonator driven with a Holzworth HS300 2A 2-channel synthesizer and interrogated using differential active probes.

In order to estimate the value of $\Delta k/k$ obtained through off-resonance pumping, a varying polarity dc bias voltage was applied to the device. Recalling that $\Delta k/k = 2\Delta f/f$, we can estimate the $\Delta k/k$ achieved through measurement of the frequency shift Δf for ± dc voltages applied to the device. From the result in Figure 12, we observe a small slope of approximately a 0.0824 ppm /V corresponding to a change in the $\Delta k/k$=1.65x10^{-7} /V. This measured value agrees well with the theoretically calculated value of $\Delta k/k$=1.57x10^{-7} /V. From this result we conclude that off-resonance parametric pumping will require voltages which far exceed realistic values that can be easily implemented. However, on-resonance pumping, in which we utilize the high Q of the device, can provide significantly larger $\Delta k/k$ values and hence a more practical parametric pumping condition.

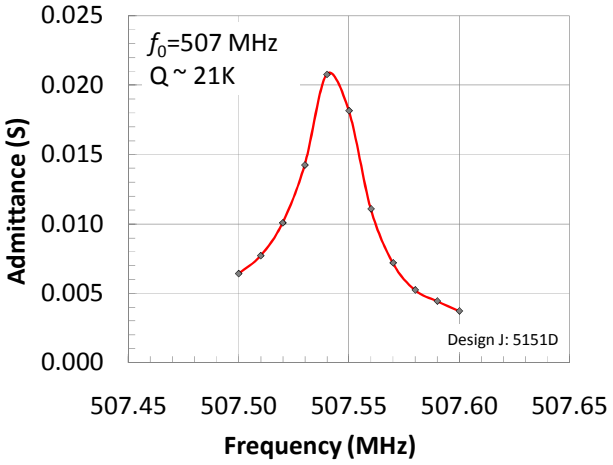


Figure 11. Admittance response of an AT-cut UHF 507 MHz quartz resonator at the fundamental thickness shear resonance.

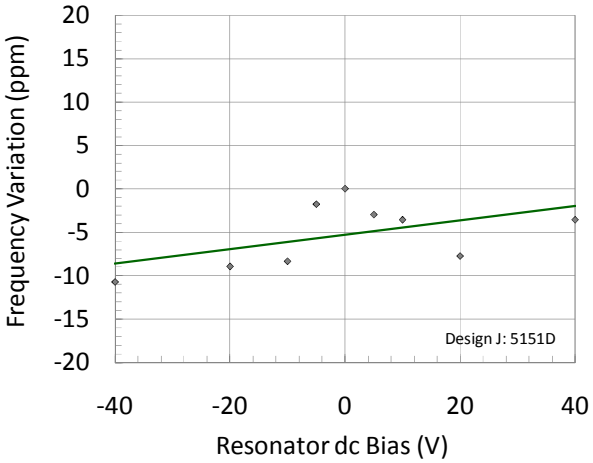


Figure 12. Resonator ω_0 frequency variation versus applied dc bias voltage for a AT-cut UHF 507 MHz device. A small positive slope is observed indicating prohibitively high voltage requirements for off-resonance parametric pumping.

Recalling that $\omega_P = 2\omega_0/n$, where ω_0= natural resonance frequency of the resonator, and ω_P = pump frequency, which in this case is the frequency at which the $\Delta k/k$ occurs. Therefore for n=1,2 &3 the corresponding $\Delta k/k$ occurs at $2\omega_0$, ω_0, & $2\omega_0/3$ respectively, and the parametric excitation frequency falls at ω_0, $\omega_0/2$, & $\omega_0/3$ for rectified $\Delta k/k$. Pumping at $\omega_0/2$ is an off-resonance condition which will be inefficient. Thus, we have two scenarios which satisfy the on-resonance pumping condition. Consider a pump at ω_0, for which one can use the rectified $\Delta k/k$ at $2\omega_0$ to either pump the fundamental mode at ω_0 (n=1) or the 3rd overtone at ω_3 (n=3). Note that in order to implement the latter case of n=3, the 3rd harmonic of the device should be well matched (within the linewidth) of the 3rd overtone mode.

VI. Summary

In summary, new simulation tools have being developed for a complete COMSOL model of the quartz non-linearities including computation of $\Delta k/k$ for resonance and off-resonance conditions using the second and third order elastic constants. New experimental techniques have been developed to directly measure the amplified response and $\Delta k/k$ values at UHF frequencies with multiple inputs applied to a resonator simultaneously. We have determined that the $\Delta k/k$ change is a rectified function of the excitation voltage and its measured magnitude agrees closely with our predictions.

Parametric amplification in quartz shear-mode resonators may be possible with reasonable voltages if the strain due to a resonant response is utilized. The off-resonant pumping condition appears to be impractical using reasonable voltages. Future work will investigate using the fundamental mode to parametrically pump the third overtone for improved phase noise.

We acknowledge the support of D. T. Chang, H. D. Nguyen, F. P. Stratton, T. J. Boden, B. T. Holden, M. F. Boag-O'Brien, H. C. Moyer from HRL Laboratories, LLC for device fabrication and electronics development.

REFERENCES

[1] J. Vig, "Quartz Crystal Resonators and Oscillators for Frequency Control And Timing Applications - A Tutorial", 2008

[2] M. Zalalutdinov, et. al., "Autoparametric Optical Drive for Micromechanical Oscillators," Applied Physics Letters, 79, 695, 2001.

[3] L. Grasser, H. Mathias, F. Parrain, X. Le Roux, and J. Gilles, "MEMS Q-Factor Enhancement Using Parametric Amplification: Theoretical Study and Design of a Parametric Device," DTIP of MEMS & MOEMS, Stresa, Italy, April 25-27, 2007.

[4] D. Rugar and P. Grutter, "Mechanical Parametric Amplification and Thermomechanical Noise Squeezing," Phys. Rev. Lett., 67, 699, 1991.

[5] R. Bechmann, A. D. Ballato, and T. J. Luzaszek, Higher-order temperature coefficients of the elastic stiffnesses and compliances of the alpha-quartz, *Proceedings of the IRE*, Vol.50, No. 8, 1962, pp.1812-1822.

[6] Y-K Yong, J. Wang and T. Imai, On the accuracy of Mindlin plate predictions for the frequency-temperature behavior of resonant modes in AT- and SC-cut quartz plates, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 46, No. 1, January 1999, pp 1-14.

[7] [3] E.D.Fletcher and A.J.Douglas, "A Comparison of the Effects of Bending Moments on Vibrations of AT and SC- Cuts of Quartz", *Proceeding of the Frequency Control Symposium,* 1972, pp.346-350.

[8] Y.K.Yong, and M.S.Patel, Application of a DC-Bias to Reduce Acceleration Sensitivity in Quartz Resonators, *International Journal of Applied Electromagnetics and Mechanics*, 21, 2005, pp.1-14.

[9] J.J.Gagnepain, J.C.Poncot and C.Pegeot, "Amplitude Frequency Behavior of Doubly Rotated Quartz Resonators," Proceedings of 31st Annual Symposium on Frequency Control, US Army Command, Fort Monmouth, New Jersey, pp.17-22, 1977.

[10] M.S. Patel, Y-K Yong and M. Tanaka, "Drive Level Dependency in Quartz Resonators", Mihir Patel, Yook-Kong Yong and Masako Tanaka, International Journal of Solids and Structures, Volume 46, Issue 9, 2009, pp. 1856-1871.

[11] A. Grebennikov and N. O. Sokal, "Switchmode RF Power Amplifiers," Elsevier Inc. 2007, pp. 46 – 51.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

Collective Fabrication of 20 MHz Resonators by Deep Reactive Ion Etching on 3'' Quartz Wafers

JJ. Boy[1], H. Tavernier[1], X. Vacheret[1], T. Laroche[1], A. Clairet[2]
[1]FEMTO-ST, UMR 6174, Frequency & Time Dept, 26 chemin de l'Epitaphe, 25000 Besançon, FRANCE
[2]RAKON-TEMEX, 2 rue Robert Keller, 10150 Pont-Sainte-Marie, FRANCE

Corresponding author : jjboy@ens2m.fr

Abstract— **High quality resonators for spatial and military applications are only made by unitary way and high speed directional etching of piezoelectric material is yet insufficiently developed to produce high aspect ratio microstructures.**
So, in this paper, we report on the theoretical definition and on the realization of BAW resonators, working at 20 and 40 MHz. Part of mechanical process is made by deep Reactive Ion Etching of AT- and SC-cut quartz crystal wafers. To avoid edge effects such as mechanical stresses induced by mounting structure or leakage of the vibration mode, we have to realize a good energy trapping of the selected resonant frequency. Several trapping methods can be used depending on the frequency, thereby changing the resonator design, such as mass loading by electrodes themselves, mesa forms (i.e. 1 to 3 µm circular or elliptical steps), or radius of curvature on one face of the resonator at least. Here, for question of manufacturing, we choose to trap the energy by a mesa form.
Fabrication of complete mesa architecture with bridges aperture (like in a bva structure) requires combining high depth (about 140 µm for a 40 MHz 3rd overtone resonator), high aspect ratios, good uniformity over the entire wafer (for about 40 resonators), vertical wall profiles and reasonable etching selectivity.
After describing different RIE processes, we analyze the quality of the realization through the surface roughness, the geometry, the homogeneity of the mesa-step, the wall profile.

I. INTRODUCTION

Within the framework of a french project aimed achieving quartz resonators for Time and Frequency applications, and entitled "*Crystal resonators modeling and innovative process for miniaturization*", we investigate several manufacturing processes compatible with 3 or 4" wafers. Indeed, innovative technological processes, such as micromachining, multi-layers assembly or wafer bonding can be successfully applied to quartz, in order to minimize packaged resonators while optimizing their performances. Our challenge is summarized in the following sentence:

"Smaller, cheaper and better performances"

Unfortunately, current manufacturing technology for BAW quartz resonators does not allow reducing the size and the cost. We therefore have to study collective processes, which are well-known in the Silicon industry.

II. PRESENTATION OF OUR PROCEDURE

In this paper, we present the manufacturing process of small resonators designed to vibrate at 20 MHz on their 3rd overtone (corresponding to a thickness of about 280 µm, which is a standard value for the purchased quartz wafers). Among different collective etching techniques (chemical etching, UltraSonic Machining or grinding and polishing…), we study the realization of resonators using the DRIE technique which is described in [1]. This technique has already proved that it is conceivable to realize mesa on the active part of the resonator and the apertures defining the links (or bridges) between active and dormant parts. Indeed, previously, we have fabricated few resonators in a 140 µm thick quartz wafer of 1.5 x 1.5 square inches using DRIE. The apertures were completely opened at least in the resonator made in the center as shown in the Fig. 1. We see here that the etching depth is not exactly the same in the entire surface of the wafer, but it is very difficult to measure it due to the strains (and fractures) induced by the thermal expansion created during the etching.

Fig. 1: DRIE of a small quartz wafer (1.5 x 1.5'' square)

So hereafter, we describe two different procedures in order to manufacture resonators with a "double step" on one face as shown in Fig. 2. This succession of steps on the surface is necessary if we want to discretize correctly the radius of curvature realized on resonators working at a few MHz in the way to correctly trap the energy of vibration. Therefore, we will first realize 2 steps with a size of few tens of µm each. After the machining of the apertures, we plan to implement the full procedure for manufacturing 37 resonators on a 3" AT-cut quartz wafer.

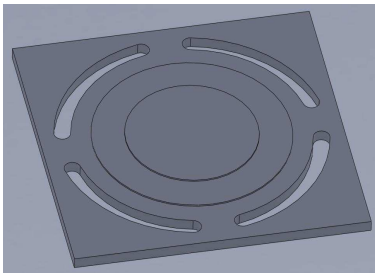


Fig. 2: design of the desired resonator

III. FIRST DRIE QUARTZ ETCHING PROCESS

In this first test procedure, we realize a succession of wet etchings in order to structure layers of Nickel. The total thickness has first been deposited to allow the total etching of about 140 µm. To etch 150 µm of quartz, we have to coat 8 µm of Nickel using the "electrolytic" method (selectivity 1:20). These two preliminary etching operations are separated by a coating of photoresist, exposed to light through a mask defining our pattern. Another etch of about 2 µm of Ni by *Ferric Chloride* is performed. The last etching operation is made to remove the last 4 µm of Ni at the level of the apertures.

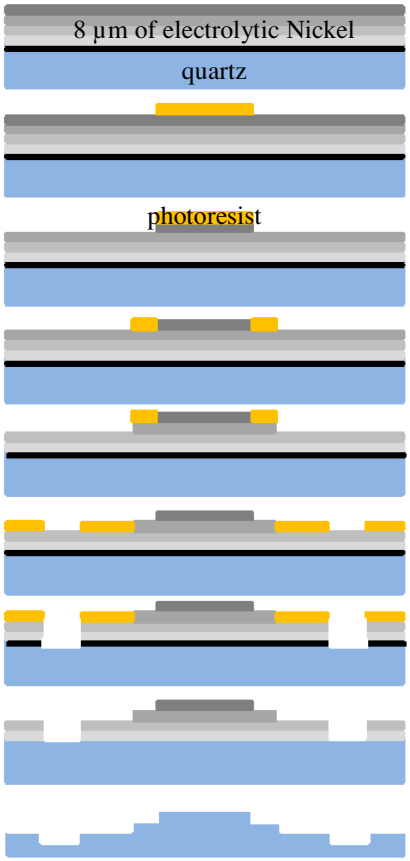


Fig. 3: procedure to structure layers of Nickel

Unfortunately, the chemical etching process of the Ni by Ferric Chloride is very inhomogeneous. Up to now, it is impossible to calibrate the thickness to be removed. Furthermore, the thickness of Ni on the quartz wafer is too thick and cracks appear systematically during the ionic etching which generates very high stresses.

IV. SECOND PROCEDURE

We have then tested a second procedure, which is more "classical" and consists in 3 successive phases with 3 different masks (Fig. 4): the first one realizes the apertures (first mask on the left) with the coating of 4 µm of electrolytic Ni allowing the dry etching of 80 µm of quartz.

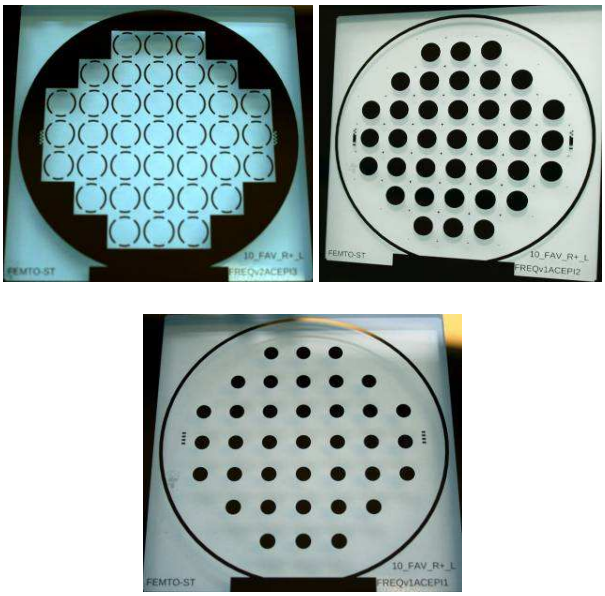


Fig. 4: Three successive masks for our second procedure

Then, coating of a new layer (about 2µm) corresponding to the pattern defining the biggest mesa and dry etching for a 40 µm step is performed. Finally, the third coating defines the small mesa with also 2 µm of Ni. Unfortunately, here also, the mechanical stresses induced by the layer of electrolytic Ni submitted to high temperatures during the DRIE process deform the quartz wafer up to cracks (see Fig. 5). Several tests have been realized in changing the temperature in the back of the substrate, but the problem remains the same: the wafers still break down.

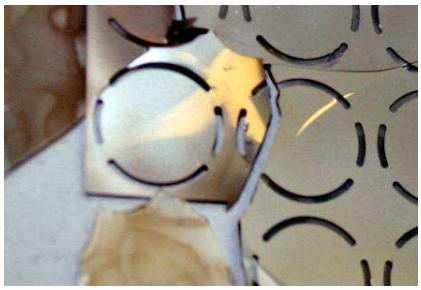


Fig. 5: photography of an example of crack after DRIE

Nevertheless, we have observed MESAs on the AT-cut quartz wafer and we have measured the height of some steps.

Classically, on our STS machine, the etching process is faster on the edges than on the center. Here, we have etched 27 µm in the edges and only 17 µm in the middle.

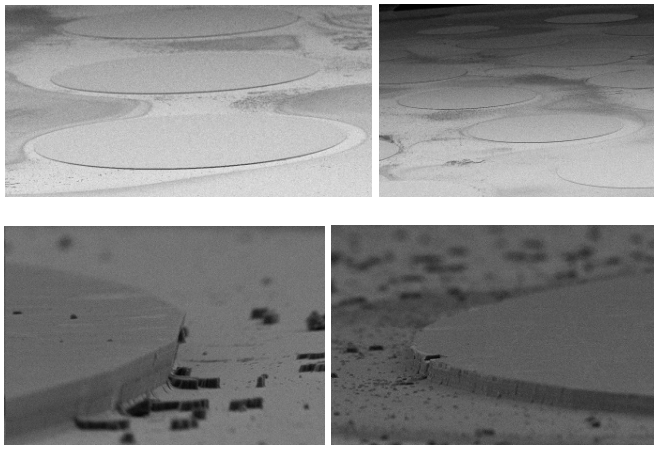


Fig. 6: MESAs observed by BEM… and details of the steps

V. THIRD PROCEDURE

The last procedure described here aims to prove that it is possible to chain 2 realizations of mesa one after another; the apertures (defining the bridges) may be made, at the end, using Ultrasonic Machining [2].

Considering the small heights of the steps to realize (see Fig. 7 below), we have finally chained here 2 Reactive Ion Etching processes, allowing etching 2 to 2.5 µm thick on quartz. Indeed, our wafer is 280 µm thick, processed to realize resonators working at about 20 MHz on their 3rd overtones. So, a 1000 mm radius of curvature seems satisfactory to trap correctly the vibrating energy and the 2 corresponding steps have to be of about 2 to 2.5 µm height.

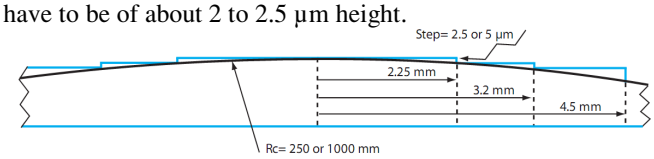


Fig. 7: example of discretization of a given radius of curvature

If the photoresist has been deposited by spin coating for the first step, we have coated it by spray coating for the second one, the relief of the first deposit being too high. We see in Fig. 8 that the quality of the pattern is the same than previously.

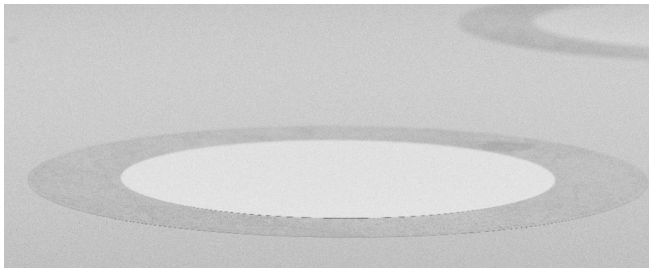
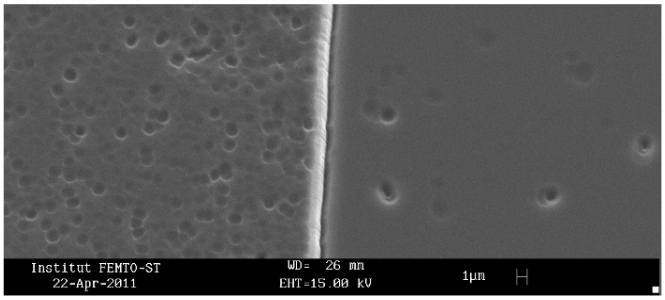


Fig. 8: The 2 MESAs with 4.5 and 6.4 mm of diameters

But, if the surface roughness does not depend on the number of the manufacturing process, the quality of the edge of the pattern is highly dependent on how the photoresist is deposited. Indeed, a thickness of the resin deposited by spin coating is uniform only if the surface of the wafer is uniform, a pattern higher than 2 µm creating a kind of bulge in the thickness of the coating. And consequently, the edge of the pattern will have the quality of the deposition, this defined by the spray coating being linked to the microbubble size of the spray. The 2 Fig. 9 below illustrate this difference.



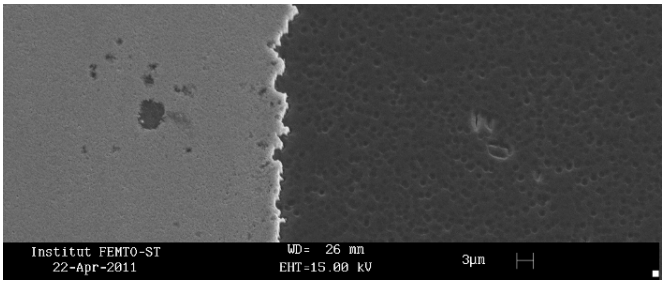


Fig. 9a and 9b: definition of the edges of the first and second MESA, linked to the quality of the deposition to the Nickel (spin coating for the first one and spray coating after).

And so, the height of the step is not uniform as shown in the Fig. 10 where we indicate 2 different values of steps: 2 and 3 µm which are the minimum and the maximum of the step heights.

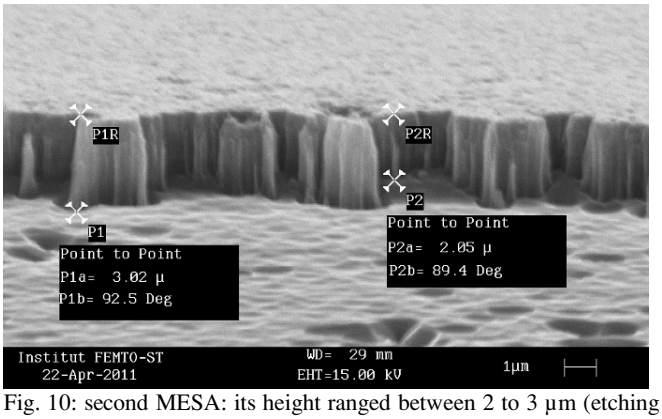


Fig. 10: second MESA: its height ranged between 2 to 3 µm (etching time: 70 min)

We note, at least, the second pattern has been obtained with the following conditions of Reactive Ion Etching:

- deposition of 320 nm of Ni/Cr
- spray coating of S1813 photoresist (4.2 µm)
- photolithography with the given mask,
- annealing at 120°C during 1 min,
- Cr etch / Ni etch and finally
- 7 steps of 10 min with SF_6 gas at 2mT and 220 W.

This work will be followed by the deposition of the metallization allowed to excite the chosen mode of vibration, which is the C-mode 3rd overtone, working at about 20 MHz for a 280 µm thick wafer (in fact 17.68 MHz, the required value of thickness being not reached). Then, we will open the apertures by UltraSonic Machining (USM) before bonding the covers on each side.

VI. FEM SIMULATION OF THE RESONATORS

Previously to this technical work, we have initiated efforts in the calculation of the best design to well trap the vibration energy of the C-mode, 3rd overtone. For that, we use the Finite Element Method allowing calculating resonant frequency and Q-factor for a single-step-mesa resonator and a stepped bi-mesa structure. Below, we compare these calculated values with the output frequency characteristics of a beveled "standard" quartz crystal resonator working at about 20 MHz. Finally, this tool should help us to simulate the complete behavior of any resonator, including the mounting structure and its influence on the resonant frequency.

Few works have been done in this area [3-4], completed by the realization of QCM resonators (Quartz Crystal Microbalance) for which the goal is not to reach very high Q-factors, dedicated to frequency and time applications [5], but only to measure mass loading effect.

So, to evaluate our simulation, we have modeled a resonator with an external diameter of 9 mm, a diameter of 3 mm for the electrodes (constituted by 200 nm of gold) and 6 layers in the thickness of quadratic elements (see Fig. 11).

At least, to calculate the acoustic loss of the desired resonant frequency, we have introduced the tensor of the viscosity constant measured by Lamb and Richter [6]. And so, the displacement along the X-axis for different designs is as indicated in the following figures (Fig. 12a to Fig. 12d).

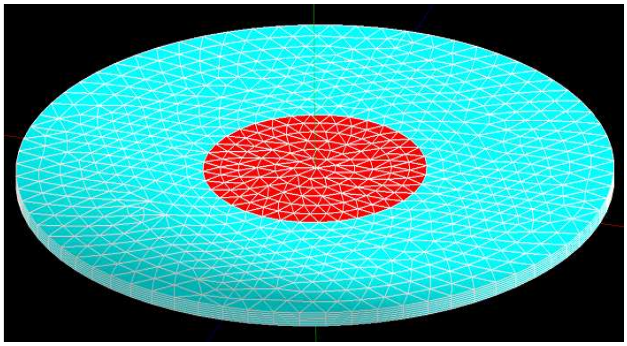


Fig. 11: finite elements model of the resonator and its electrode

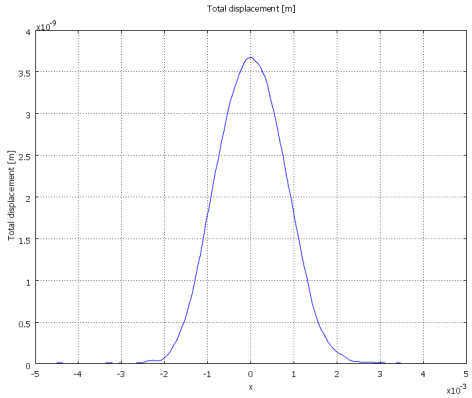


Fig. 12a: displacement along X-axis for a plano-convex resonator

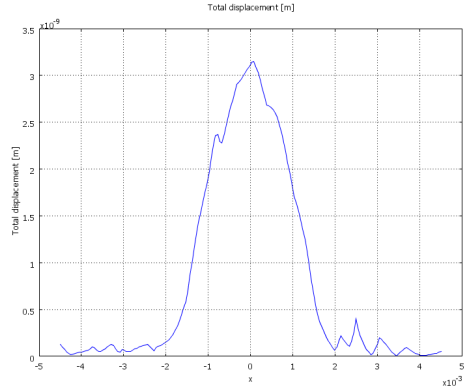


Fig. 12b: displacement along X-axis for a plano-plano resonator

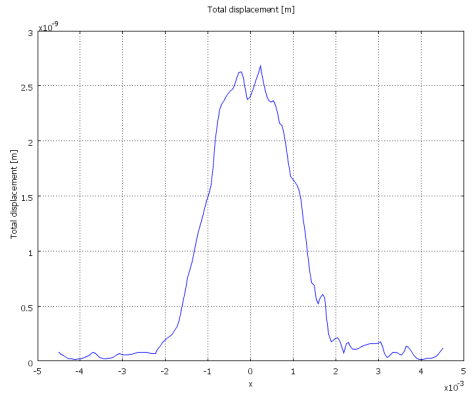


Fig. 12c: displacement along X-axis for a "1 MESA" resonator

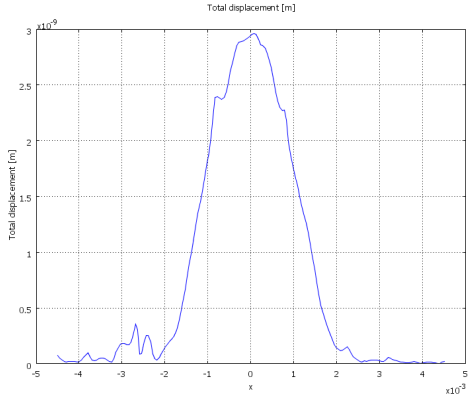


Fig. 12d: displacement along X-axis for a "2 MESAs" resonator

The Q-factor of the C-mode 3rd overtone has been calculated for each design. They are indicated in the following table:

design	Q-factor
Flat resonator	771,000
with Rc = 1000 mm	766,000
with 1 MESA	767,500
with 2 MESA	767,200

Table 1: Q-factor calculated by Finite Element Modeling

We observe that the design has a very little influence on Q values. It is not surprising because the working frequency of 20 MHz corresponds to the point for which a curvature or the mass loading have almost the same influence on the energy trapping.

VII. CONCLUSION

Our test vehicle is not achieved and we have to develop:

Metallization, wafer bonding, dicing,
Chemical etching,
ultrasonic machining of the covers,
…

and to realize 3 and 4'' SC-cut quartz wafers.

Nevertheless, we have demonstrated several points:

- ✓ DRIE shall be an efficient tool to realize deep patterns if we anneal the thermal stresses induced by the electrolytic Ni layer (by deposition of new Ni layer on the other side of the wafer, for example). So, we hope to etch more than 80 µm deep currently,
- ✓ we are able to chain 2 or more RIE processes on the same wafer,
- ✓ …

and we own an efficient tool to calculate the motional parameters of a given resonant frequency of a resonator with particular design.

ACKNOWLEDGMENT

We thank Aurelie Lecestre and Artur Zarzycki who have performed the different etching processes.

REFERENCES

[1] J.J. Boy, S. Queste: "deep reactive ion etching of quartz crystal for frequency metrology" – Proc EUSPEN 09

[2] Andrey E., Boy J. J., Khan Malek C.: "Tool wear for micro-ultrasonic machining (MUSM)" - 5th Int. Conf. European Society for precision Engineering and Nanotechnology, Montpellier, 8-11 mai 2005, Proc. EUSPEN 05, pp. 345-348.

[3] M. Tanaka, T. Imai and Y.K. Yong: "3-D FEM Eigenvalue Analysis od Relative Impedance and Energy Trapping of Resonant Modes in AT-cut resonators" –IEEE Freq. Control. Symp. 2003, pp. 709, 714.

[4] F. Lu, H.P. Lee, S.P. Lim: "Energy trapping analysis for the bi-stepped mesa quartz crystal microbalance using the Finite Element Method" – Smart Mater. Struct. 14, 2005, pp. 272, 280.

[5] T. Abe, V.N. Hung, M. Esashi: "Inverted mesa-type quartz crystal resonators fabricated by Deep Reactive Ion Etching" – IEEE Trans. of UFFC, vol. 53, n° 7, July 2006, pp.1234, 1236

[6] J. Lamb, J. Richter: "Acoustic Attenuation with New Measurements for Quartz at Room Temperatures" - Proceedings of the Royal Society of London Vol. 293, No. 1435, 1966, pp. 479-492

Quartz Crystal Industry of China at Crossroads

Ji Wang[1], *Liansheng Jiang*[2], *Min-Chiang Chao*[3], *Xuming Chi*[4], *Jianwei Hu*[5], *Zhuzhi Ye*[6], *Lihu Pan*[7], *and Weiqiu Chen*[8]
[1]Piezoelectric Device Lab, School of Mech. Eng. & Mechanics, Ningbo University, Ningbo, Zhejiang 315211, CHINA
[2]Piezoelectric Crystal Association of China, P. O. Box 120, Beijing 100015, CHINA
[3]TXC (Ningbo) Corporation, 189 West Huangshan Rd., Ningbo, Zhejiang 315800, CHINA
[4]Zhejiang East Crystal Electronic Co., Ltd., 555 Binhong Rd., Jinhua, Zhejiang, 321017, CHINA
[5]Ningbo Hiking Electronics Tech. Co., Ltd., 11 Qixin Rd., Ningbo, Zhejiang 315051, CHINA
[6]Timemaker Crystal Tech. Co., Ltd., Bldg. 63/6F, Majialong Industrial Zone, Shenzhen, Guangdong 518000, CHINA
[7]Beijing Institute of Radio Metrology & Measurements, P. O. Box 3930, Beijing 100854, CHINA
[8]Department of Eng. Mechanics, Zhejiang University, 38 Zheda Rd., Hangzhou, Zhejiang 310027, CHINA
E-mail: wangji@nbu.edu.cn

Abstract—**Quartz crystal devices have been playing important roles as enablers in modern electronics, and China also has a strong industry for circuit applications in the beginning and as a leading global supplier and end-user lately. The transformation of the quartz crystal industry in the US and Europe has led to the globalization and reorganization of the production of quartz crystal devices in China with significant impact in the market and technology. It has also been known that as a producer of quartz crystal products, leading companies in China have been collaborating with partners outside China for technology and marketing to promote trade and production. In the meantime, many manufacturing hubs for leading OEM companies of consumer electronics industry, with telecommunication and appliances in particular, have been using many quartz crystal products from Japan and Korea. It is possible that large producers in Japan and Korea will continue their production through technology leadership, but the market share can be significantly weakened by Chinese manufacturers. As an initiative to foster the technology development and innovation, Chinese companies have been encouraged and assisted through various partnerships such as joint ventures with foreign companies, research and development collaboration with universities, and governmental subsidy for innovative and quality products, to upgrade technology and serve the high-end application in China.**

I. INTRODUCTION

The quartz crystal industry in China has been making great impact on the global market and industry landscape through supplying vast quantity of lower-end products and force many manufacturers in the United States and other countries to move their production facilities to China and other low-cost nations. This has been noted in the industry for a while and many strategies to counter change have been developed and implemented. For instance, many manufacturing operations of US companies have been relocated to China or even sold to other operators as major indicators of industry-wide changes. On the other hand, the Chinese industry has also been greatly influenced by changes in the US, notably the emergence of newer technology such as the micro-electrical-mechanical systems (MEMS) and film bulk acoustic resonators (FBAR) products which are already invading certain portion of traditional market of acoustic wave resonators in electronic products and new applications. One frequent question we encountered in various occasion is that how the new and emerging technologies in the US will impact the existing investments in the quartz crystal and larger acoustic wave device industry? Then an immediate question followed is that do we need to invest in quartz crystal technology while the emerging products are imminent? The answers to these questions will not only impact the individual company looking for advices but also the industry in China and the global market.

Fig. 1 Major production clusters of quartz crystal products in China (Quartz crystals; Resonators and oscillators)

As an attempt to address these and similar questions and answering more outside of China, we have already had many discussions in China in various occasions with business and industrial leaders. Our general impression is that while the quartz crystal industry in China is growing at a much faster

This industry surveying project is supported by Zhejiang Province through the Key Team of Excellence for Technological Innovation under Grant 2011R09025.

pace than the overall economy because the electronics sector has been the leading force of the recent fast growth and it is specially encouraged by central and local government under many incentives, which often are on competitive basis in coastal regions. In addition, the quartz crystal and acoustic wave device industry is China is still one of the many high-tech sectors and product categories under special promotion in economic plans and policy. Of course, as many of us already know, there are also generous incentives on tax-relief from local governments and export tariff refund for selling products to international market. Such coupled efforts in promotion of quartz crystal industry and other sectors have significant effect on the fast growth, since the production of quartz crystal products have been growing over 10% a year as shown in Table 1. Besides, quartz crystal manufactures outside China, Japan in particular, have been under intensive pressure in terms of the change of market share and fast advances in product quality and manufacturing efficiency. In other words, the gap between Chinese manufacturers and the leaders in Japan has been shrinking and it is always a question of how the advantage can be kept or changed for different competitors.

Table 1 Yearly production of quartz crystals in China

Year	Product (billion pieces)
2002	1.08
2003	1.21
2004	1.39
2005	1.54
2006	1.77
2008	2.25
2009	2.41

For a snapshot of quartz crystal industry in China, here are some basic figures: there are about 500 hundred companies with annual production of 3 billion pieces of quartz crystal products in the Year 2010. The total annual sales are about 1 billion US dollars, about a quarter or less of the international total sales. Not surprisingly, also contrary to our impression, China buys more quartz crystal components than it produces (see Table 2), obviously leading manufacturers of electronic products like computers and mobile phones use more components made outside of China. Most lower-end products are made with everything, production equipments and raw materials included, supplied in China while the high-end products are made with Japanese equipments and even materials. Also there are Japanese engineers, some retired, working in Chinese companies for the production and design.

Table 2 Quartz crystal export and import in China

Year	Export (pieces, billions)	Import (pieces, billions)
2008	0.94	1.43
2009	1.05 (USD 0.8 billion)	1.61 (USD 2 billion)
2010	1.18	1.80

As an indicator of slowing moving up of business ladder, two major producers of quartz crystal products have been listed in stock changes in China so far: Jingyuan Electronic, Tangshan, and East China Electronic Crystal Company (ECEC). It is a measure of maturity of these leading companies in the business world in China. Research and development is inadequate in this industry in China because many companies cannot afford the talents and related costs.

II. Industrial Resources and Advantages

In the international market, the competitive edge of Chinese quartz crystal products is the price, as the majority consumer products made in China. This is the outcome of the abundance of low-cost labors available to industry. Of course, as said before, the preferential policy for electronics components including quartz crystal products also plays an important role in the fast growth in recent years. The availability of land, raw materials, and electricity in certain regions are also major factors of the continuing growth.

If we look into the quartz crystal industry further, we can find more factors which contributed to the fast growth of the sector in China. One major reason is the availability of quartz crystal resonator production equipments from Japan, which enable many companies to make competitive products without invest too much efforts in making and improving production lines as did by major producers in Japan and the US before. Actually, there are also manufacturers in Hong Kong and China to accelerate the growth through even lower costs in obtaining a quality production line. Of course, this also contributes to the fact that many products from Japan are almost same because they do not differentiate from the production equipments and accessories needed for the packaging of resonators and oscillators. With the production line installed, the manufacturing is a straightforward process as we know, and it can churn out products with the standard supplies of all elements and accessories. In other words, the entry barrier of quartz crystal products is relatively low and this also contributes to the fast spread of companies making the same products in China. In fact, companies are clustered after one demonstrated the technology and production process.

We have mentioned that even leading producers of quartz crystal resonators lack technical capabilities in research and development, thus keeping the products and process in a much lower level in the global competition. There are a few focused research institutes in government laboratory systems, such the few in the China Electronics Technology Group Corporation (CETC), and BIRMM, and the research groups and laboratories in leading companies such as the Beijing Chengjing Electronic Technology Company (China 707), which have relatively strong technical team to engage in research activities for the design and fabrication. It is generally expected that these work should contribute to advances of technology of the quartz crystal industry in China, but in fact this did not happen. The main reason, as we can guess, is that most such research oriented organizations are also production companies making devices for specific applications of their parental industry. For instance, BIRMM is an affiliate of the space industry in China and their products are being used primarily in the space programs. As a result, it is regrettable that these research laboratories, with considerable resources in technical talents and equipments, cannot be utilized effectively to lead the industry. Earlier attempts to improve the overall maturity of quartz crystal products in China through research led process did not succeed as the mass production seems not always correlated

with the technology supremacy they have. This reminds us how to use the full potential of research laboratories in the quartz crystal industry in China in the near future is still a challenge to be solved. However, recently increased participation of the annual Symposium on Piezoelectricity, Acoustic Waves, and Device Applications (SPAWDA) by industry is a good sign of embrace to technology and research through interactions with academic institutions.

On the other hand, there are many manufacturing subsidiaries of leading international quartz crystal producers in China today. Their core developments activities are largely kept in the home office like Japan and the US, as we always meet engineers come to China to discuss potential products for Chinese markets. However, more or less, as we have seen in past few years, many engineers and managers from such companies are also migrating to Chinese companies and this will inevitably add momentums to accelerate changes. Many such multinational companies, particularly manufactures from Japan, are concentrated in Yangtze Delta (including Suzhou, Nanjing, and Zhejiang), while a few others particularly Korean companies are in Yantai area and a few others are in Shenzhen area. Some of them are joint ventures with local partner companies in China and this will contribute to the growth of Chinese companies with technology and management advantages. It is clear that the globalization of this industry will also give the advantage, although limited, to the Chinese industry.

It has been clear that the economic growth and accumulation of wealth in China in recent decades add the power to invest in the quartz crystal industry also, as it is viewed as one of the high-tech sector with relatively lower entry barrier. Indeed we have seen new companies of various sizes in almost every process of quartz crystal products through crystal growth, blanks, holders, and equipments. The global financial crisis has some effects on the growth and investments in this traditional technology and the further relocation to China is expected. On the other hand, it is generally sensed that newer technologies in competition with traditional frequency control devices have some impact on the sentiment of investment, because many investors are not sure how and when the alternative technology will phase out the current products. This concern also reflects the lack of research and understanding of the technology trend in this industry and applications. It is generally recognized now that new technologies will replace some products and applications with a much lower pace and it is not the situation that the technology will end abruptly soon. The recent emergence of a few more joint ventures and some expansions of major companies in China may prove that concerns caused by MEMS and FBAR technologies have been smoothed and there are more interests in further investment and improvement of development and production in the quartz crystal industry in China. To speed up the process, again, a short cut is being taken by many companies through increases of investments in buying existing design and technology from international companies. Particularly, many companies in China have expressed interests to buy some small quartz crystal companies to acquire the technology in product development they need badly. Also it is suggested that many companies in the US and Japan may be in the position to sell some of their know-how and technology while their value are at the peak when there are great demands from China. The time window for such buying mood may be limnited and the progress in technology in China could cause depreciation quickly.

III. MAJOR PLAYERS AND GROWTH POTENTIALS

Most Chinese manufactures of quartz crystal products are less well-known even in this global industry, because we do not see engineers and business leaders from China in the IEEE International Frequency Control Symposium and other important occasions. Their appearance in international trade shows is also limited, although we sometimes see Chinese companies in a few events like in Las Vegas and Munich. The reason of their absence in international events is that their products are usually not marketed with their own brand, as most other industries in China. Or, this is the true meaning of "Made in China", while most product today are created outside of China. This is the business strategy of a country and industry while the technical capability is limited and low-cost labor supply is abundant. Such a business practice has its disadvantages, but it will take a long and slow path to change giving the economic and technological reality in China. Then it is easy to understand the choices of Chinese companies and industry, and it also offers the strategy for outsiders to choose the right road to enter the seemingly prosperous and strong economy.

There is a relatively fully fledged quartz crystal industry in China with manufactures for quartz crystal material, quartz blanks and chips, packaging and accessories, and all equipments for the making of resonators. This means the production of quartz crystal devices can be done with all essential materials, accessories, and services in Chinese companies for lower-end products. For high-end products, we know that many elements, especially the holders and high performance materials have to be imported from Japan and the US. The size of market and amount of demands in China is significant and all the major equipment makers and material suppliers have opened offices and branches in China and have been enjoying fast growth of businesses in recent years even with a bright future.

As the industry grows in the technology and market, an industrial association of piezoelectric devices, **Piezoelectric Components Association of China** (PCAC, http://www.chinapcac.org), has been established by the major producers. It is the only official trade association of quartz crystal and acoustic wave device industry in China, and has been enjoying the support nationwide with intensive contacts with global major producers of the industry and joint activities with similar groups. The PCAC has a major mission of promoting the industry and support members through sharing information on government policy and regulation, market, technology, and trade. It is also leading the task to maintain national standards on quartz crystal devices as the IEC TC-49 counterpart. It also holds annual trade shows and meetings of leading producers to review market and technology trends.

Zhejiang East Crystal Electronic Co., Ltd. (ECEC, http://www.ecec.com.cn) is one of the two listed company

with top annual sales (about RMB 300 million / USD 45 million). ECEC has been working with TXC for distribution and improvement. The ECEC has manufacturing equipments and support expertise from Japan. The company is actively seeking growth opportunities with emphasis on engineering and technical leads.

Beijing Chengjing Electronics Technology Company, Ltd. (China 707, http://www.china707crystal.com) is the first quartz crystal resonator manufacturer in China with many years in supporting the nation's needs of technology. After many years and extensive investments by government, Chengjing has also maintained an up-to-date production line for full line of products. It also has invested in research and engineer training with leading academic institutions. Chengjing has also been actively engaged in global marketing of products with its own brand and sales force.

Beijing Institute of Radio Metrology and Measurements (BIRMM, http://www.casic203.com) is a research laboratory of the space engineering company for time standards and metrology in space navigation and communication. To satisfy the precision needs of its space programs, it maintains a division with technical capabilities in the design and manufacrurngof precision quartz crystal products. This is a research laboratory turned production company with many lines of products. BIRMM has a strong presence and interaction with the international frequency control technology community through participation of international conferences and invited visits by many leading experts in Europe and the US. To enable the technology transfer, it is seeking to expand the production capacity and new product development partner outside China. BIRMM also maintains a subsidiary for mass production of quartz crystal products for consumer products.

One fast growing company in the quartz crystal industry is the TXC Corporation, which is the major producer of quartz crystal products and has its major outpost in Ningbo, China as the **TXC (Ningbo) Corporation** (http://www.txccorp.cn). TXC Ningbo is gradually grows its own production and technical capabilities to enable autonomous business in China. TXC (Ningbo) has been helping many quartz crystal companies in China through partnership in quality management, equipment, and engineering capabilities and improvements are clear and significant. TXC (Ningbo) has also been forming close relationships with universities and laboratories for technical capability building and new products.

As a domestic grown company in China with relatively short history, **Timemaker Crystal Technology Company, Ltd.** (http://www.timemaker.com) started as a manufacturer of crystal blanks, but the business has been expanding to resonators through the partnership with Rakon of New Zealand. There are investments in engineering capabilities through collaboration and in-house improvements. The engineering team of Rakon will provide design and support to Timemaker and the joint venture in Chengdu.

Other major producers of quartz crystal products include Jingyuan in Tangshan, close to Beijing, another listed company in the quartz crystal industry with annual sales about RMB 300 million (USD 45 million). Hualianxing (reorganized as part of China Electronics Panda Crystal Technology Corporation, CEC Xtal) is another major producer of quartz crystal resonators in Nanjing. Major clusters of quartz crystal companies are in Beijing, Yantai, Shanghai, and Shenzhen. Production centers of quartz crystal materials are in Lianyongang, Hohhot, and western regions.

For scientific and engineering research related to quartz crystal industry, the internationally renowned research work on crystals are being done at Shanghai Institute of Ceramics, CAS and Institute of Crystal Materials, Shandong University. There are research works on acoustic wave devices in the Institute of Acoustics, CAS and Nanjing University. The Piezoelectric Device Laboratory, Ningbo University, has been focused on the analysis of quartz crystal resonators through analytical and experimental approaches. Electrical circuit analysis and design are done in University of Electronic Science and Technology in Chengdu and Xi'an.

The quartz crystal industry in China also faces challenges in the road ahead. First of all, to supply quartz crystal devices in coming decades with the latest trend of globalization, rising of labor cost, and technology advancement, quartz crystal companies in China have to upgrade the technical capabilities in every aspect: design, process, equipment, and applications. With the increasing of competition, professional engineers with both analytical and experimental skills are needed in product development and production lines. Most companies need to improve the marketing and sales capability in global scale to improve the profitability and product development pace. Investment on technology is essentials and possible through government support and long term planning. The production capability of quartz crystal industry has the weight and size to make a global impact, and industrial relocation and manufacturing concentration in China provide unique opportunity to accelerate the change. Demands of technology and know-hows in China offer opportunities for Western and Japanese companies in the near future. Chinese industry needs to invest on technology through well-trained engineers, management efficiency, and marketing and sales skills. Chinese industry should also invest in the new and alternative technology to satisfy the international market as it emerges as a strong industrial player in the near future.

IV. SUMMARY

There are heavily invested production facilities and adequate material supply for quartz crystal products in China. Production equipment are on par with international standard and production management are also up to the need. The major hurdles are on the engineering capabilities of design and precision manufacturing, but the rising cost of labor, land, and energy should also be kept in mind. Quartz crystal and frequency control products are still encouraged with government incentives on investment and aid, but the true advantage has to be maintained for the future growth and included in new investments. Globalization should also be expanded to include technical talents and intellectual properties for balanced and sustainable growth.

Straightforward estimations of GNSS on-board clocks

Jérôme Delporte, Cyrille Boulanger, Flavien Mercier

CNES, French Space Agency
18, avenue Edouard Belin
31401 Toulouse cedex 9 – France
jerome.delporte@cnes.fr

***Abstract*— In this paper, we present two techniques to estimate GNSS on-board clocks. These techniques only require a single GNSS receiver and allow to evaluate the stability on a given pass of the time difference between the GNSS on-board clock and the clock that drives the GNSS receiver. These methods have been validated using IGS clock products for GPS on-board clocks. We also show that they can be used to characterize ground clocks provided a better space clock is available.**

I. INTRODUCTION

Global Navigation Satellite Systems embark atomic frequency standards, the characteristics of which have a direct impact on the overall positioning performance. Therefore the general monitoring and the performance assessment of these on-board frequency standards are of utmost importance. For instance, the performance assessment of the GPS on-board clocks is carried out on a regular basis [1,2].

The estimation of GNSS on-board clocks behaviour is generally a complex task requiring a large ground infrastructure (a global network of ground stations) and an intense computation (the ODTS, Orbit Determination and Time Synchronization). This task consists first in collecting the GNSS measurements on a continuous basis. These stations measurements are then merged through an infrastructure and processed together by a dedicated facility. This facility computes the clock of each ground station and the orbits and clocks for each GNSS space vehicle. The clocks are estimated with respect to the GNSS reference time scale.

The aim of this work is to propose two alternative methods that allow to estimate the behaviour of GNSS on-board clocks with respect to a given ground station. This estimation is therefore limited to the pass duration, but on the other hand provides data at the rate of the station measurements, thus allowing to assess the short term stability of the GNSS on-board clock. Obviously this requires that the stability of the ground station clock is better than the stability of the GNSS on-board clock that we want to estimate. These methods are therefore considered as complementary to the usual method, while being much simpler and, for at least the polynomial method, available in nearly real time.

II. DESCRIPTION OF THE METHODS

The two alternative methods developed here use the phase measurements of a given ground station. The first method, called the polynomial method, consists in applying a high-order polynomial fitting to these measurements. The second method, called the residuals method, consists in carrying out a residuals computation on a given pass using precise ephemerides. This residuals method is somewhat similar to the One-Way Carrier Phase developed in [7]. These methods are described hereafter.

A. The Polynomial Method

This straightforward method is based on the use of raw phase measurements (in RINEX format) coming from a GNSS receiver connected to an atomic clock. Basically the GNSS measurements represent the clock difference between the transmitter and the receiver. The idea is to exploit that feature in a very simple way. The raw phase measurements and their iono-free combination for a given satellite pass is merely adjusted by a polynomial of high order (typically 24). Thus we keep the high-frequency part of the measurement which is expected to be representative of the clock difference. We then compute the Allan deviation excluding end segments that are generally noisier.

B. The Residuals Method

For this method, we carry out a residuals computation on a given pass using precise ephemerides. The computation is carried out as follows:

- the transmission dates are computed using the RINEX pseudo-ranges measurements
- these transmission dates are corrected using a rough satellite clock offset estimation
- the satellites' positions are computed by interpolation of the satellite ephemeris at these corrected dates

978-1-61284-110-6/11/$26.00 ©2011 IEEE

- the geometry is computed using the satellites' position and the measurements
- the corresponding phase residuals are estimated together with vertical troposphere delay adjustment

We then compute the stability of the clock time difference between the on-board clock and the receiver clock for each phase observable and for the iono-free phase combination.

III. RESULTS ON GPS ON-BOARD CLOCKS

For GPS on-board clocks, we compare our estimations to the IGS products [4] of the CODE Analysis Center. COD clock products have a sampling rate of 5 seconds. This can be considered as a validation of the two methods. The data used in this section have been obtained with receivers driven by Hydrogen Masers.

A. GPS PRN 3 Caesium Clock

Fig. 1 shows the stability of GPS PRN 3 (SVN 33) on-board clock (Cs II/IIA) obtained with both methods and compared to the COD clock solution in the very same period. The stability obtained with the polynomial method (in blue) is very close to the one obtained with the residuals method (in red).

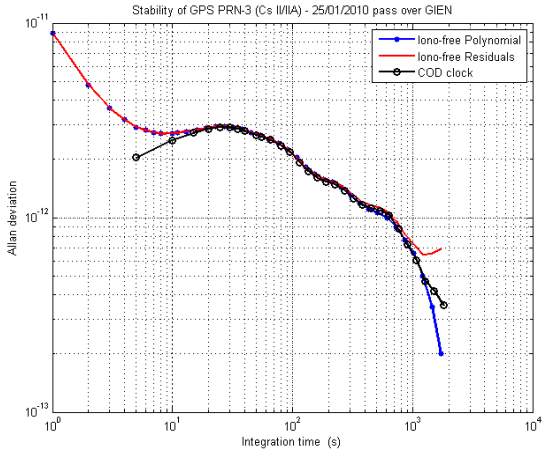


Figure 1. Stability of GPS PRN-3 Cs II/IIA using both methods on the iono-free combination

Both results are consistent with the COD clock solution between 10 seconds and 1000 seconds. It is clear that the stability is limited in the very short term (between 1 and 10 seconds) by the noise of the iono-free phase combination.

Fig. 2 shows the same stabilities computed with the polynomial method only with each phase observable (L1 and L2). Similar results are obtained with the residuals method (the short term stability is the same for a given observable in the 2 methods).

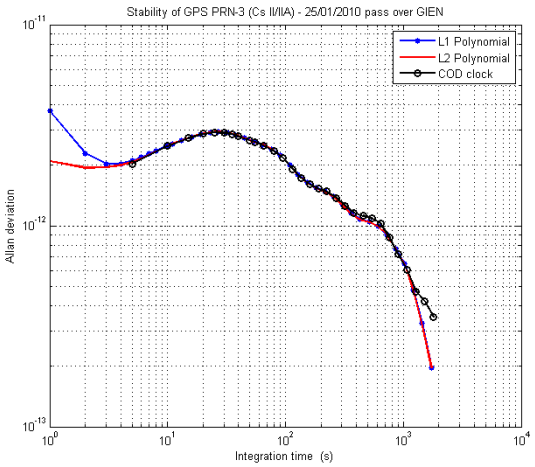


Figure 2. Stability of GPS PRN-3 Caesium II/IIA using the polynomial method on L1 and L2

We notice that the L2 observable is less noisy than L1 in the very short term. We get in this case a better short term noise because the iono-free combination increases the noise. Comparing to the previous figure, we see that the stability obtained with L1 or L2 is consistent with the stability obtained with COD between 5 and 10 seconds. We also observe here a "bump" probably due to the servo of the quartz crystal oscillator on the atomic transition that seems to have a time constant of about 30 seconds for this type of clock. For integration times between 30 and a few hundreds of seconds, we observe the expected white frequency noise.

B. GPS PRN 4 Rubidium II/IIA clock

Fig. 3 is another example with the GPS PRN 4 (SVN 34) which is a Rb of Block II/IIA. The stability obtained with the polynomial method (in blue) and with the residuals method (in red) are in good agreement with the COD clock solution from 5 to up to almost 1000 seconds. We also observe that the polynomial method can sometimes be slightly too optimistic for integration times higher than 1000 seconds.

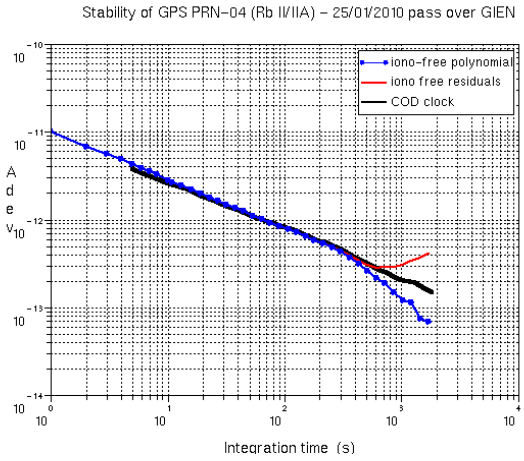


Figure 3. Stability of GPS PRN-4 Rb II/IIA using both methods on the iono-free combination

C. GPS PRN 20 Rubidium IIR

Fig. 4 is another example with the GPS PRN 20 (SVN 51) which is a Rb of Block IIR.

We also notice a very good consistency between 20 and 1000 seconds. Fig. 5 shows the stabilities computed with the polynomial method only with each phase observable (L1 and L2).

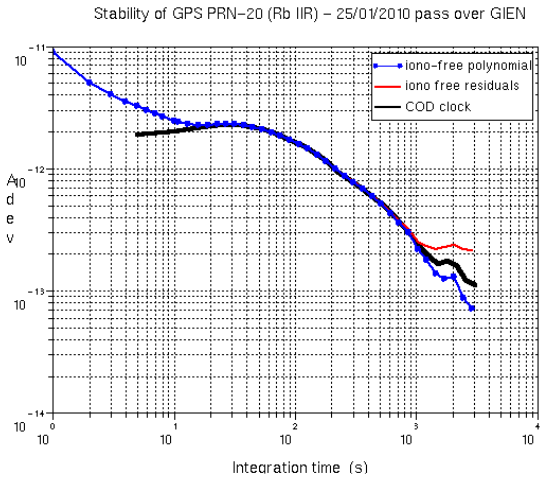


Figure 4. Stability of GPS PRN-20 Rb IIR using both methods on the iono-free combination

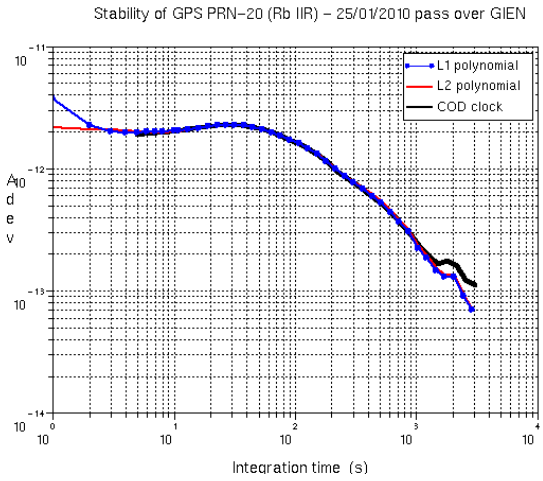


Figure 5. Stability of GPS PRN-20 Rb IIR using the polynomial method on L1 and L2

The same conclusions can be drawn with that type of clock. Similar results have been obtained with the other GPS on-board clocks and also with the measurements of other ground stations. This allows to validate these 2 methods.

D. GPS PRN 25 Rubidium IIF

The GPS PRN 25 (SVN 62) is the first Block IIF satellite, it was launched in May 2010 and embarks two Rubidium and one Caesium clocks. At the time of this experiment, the payload was driven by one of the Rubidium clocks.

Fig. 6 shows the stabilities obtained with both methods using the L1/L2 iono-free combination and compares to the COD clock solution. A white frequency noise of $1.10^{-12}.\tau^{-1/2}$ has also been added.

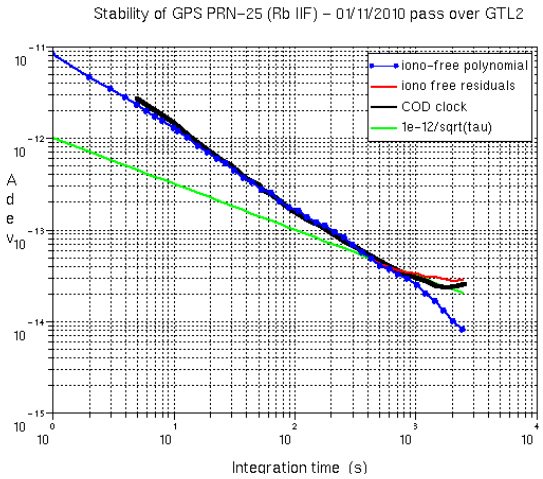


Figure 6. Stability of GPS PRN-25 Rb IIF using both methods on the iono-free combination

The iono-free combination does not allow to fully characterize the short term stability of this clock .Fig. 7 shows the results for the polynomial method on L1 and L2.

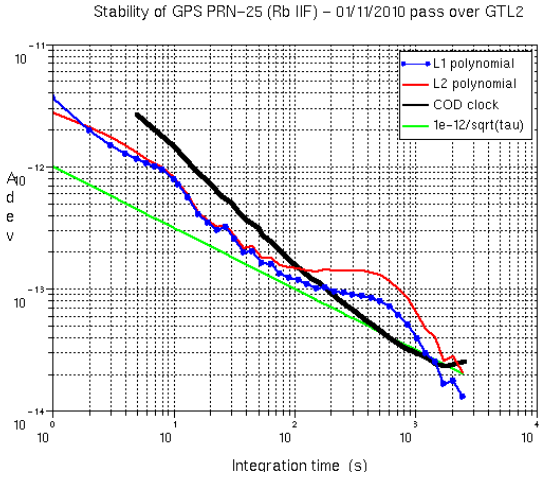


Figure 7. Stability of GPS PRN-25 Rb IIF using the polynomial method on L1 and L2

The estimated stability of this clock is 1 to $2.10^{-12}.\tau^{-1/2}$, this result is consistent with [6]. The sinusoidal perturbations identified in [6] can not be seen with these methods because they have too long a period with respect to the pass duration. However these methods allow to identify sinusoidal perturbations with shorter periods like on Figure 7 where a harmonic at about 20 seconds can be seen.

IV. RESULTS FOR GROUND CLOCKS

We set up another validation experiment in which a GPS receiver (Septentrio PolaRx2) is driven by a high-performance Agilent 5071A-001 Caesium beam clock. In the same time, this clock is measured against an active Hydrogen Maser as reference. We then compare the estimation of the Cs clock by the polynomial method (using GPS PRN 25) to the local measurement.

Fig. 15 shows the stability obtained with the polynomial method on L1 and L2 for a GPS PRN 25 pass and the stability obtained by the local measurement against the H Maser in the very same period.

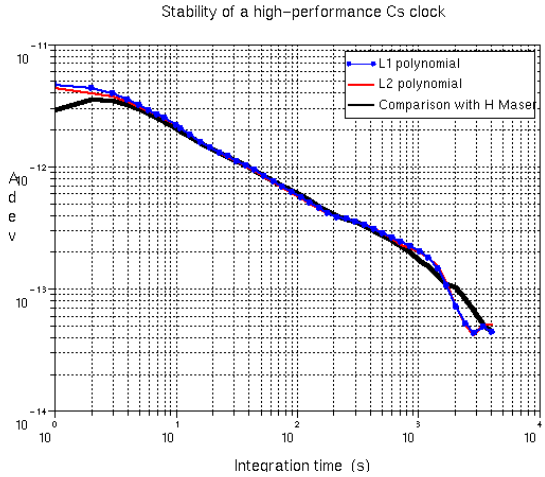


Figure 8. Stability of a ground high-performance commercial Cs beam clock using the polynomial method on L1 and L2

This figure shows an excellent agreement up to at least 1000 seconds between the polynomial results and the local measurements. This highlights the possibility to easily characterize the performance of atomic clocks such as high performance Cs beam clock with these simple techniques, i.e. without the need for a local clock with a better stability. This amounts to having a reference clock in space and therefore it works only with a GNSS on-board clock with better characteristics than the ground clock under test.

CONCLUSION

The methods developed here represent alternative ways to characterize the short term stability of GNSS on-board clocks. These methods have been validated using GPS on-board clocks and COD clock products. The results provided by the 2 methods are completely consistent with one another and with the COD clock solution.

The outstanding advantage of the polynomial method is that all you need is a ground receiver connected to a clock, the performances of which are better than the space clock. This feature is obviously extremely interesting when a global network of stations is not yet available (e.g. in a GNSS early development phase) or not accessible, or when the ODTS (Orbit Determination and Time Synchronization) process is not available. Moreover, the ODTS often provides a clock solution with a rate of 300 seconds that does not allow to estimate the short term behavior of the GNSS on-board clocks. The methods developed here provide an estimation of the on-board clock with the rate of the used RINEX observation files.

The polynomial method is considered to provide accurate estimations of the GNSS on-board clock stability up to at least some hundred seconds and then might be slightly too optimistic for longer integration times. The residuals method, which is somewhat similar to the OWCP (One-Way Carrier Phase) [7], is thought to provide accurate estimations of the GNSS on-board clocks stability up to at least one thousand seconds (provided the troposphere delay effects are taken into account).

Giving access to the high rate behaviour of the GNSS on-board clocks, these methods allow to identify potential sinusoidal patterns that may not be clearly visible with the ODTS clock solution. These methods are considered as simple, fast and efficient ways to characterize the behaviour of GNSS on-board clocks on a given pass. They are however limited because, unlike the global approaches, they do not provide continuous estimations of on-board clocks. This limitation might be overcome if the possibility to connect successive passes is demonstrated.

Moreover, these methods can be used also to characterize ground clocks with respect to better GNSS on-board clocks.

ACKNOWLEDGMENT

The authors wish to thank Francisco Gonzalez for helpful discussions.

REFERENCES

[1] J. Oaks, M. Largay, W. Reid, J. Buisson, "*Global Positioning System Constellation Clock Performance*", Proc. of the 34th PTTI, 2002.

[2] M. Epstein, T. Dass, J. Rajan, P. Gilmou, "Long Term Clock Behavior of GPS IIR Satellites", Proc. of the 39th PTTI, 2007.

[3] P. Waller, *"In-orbit Performance Assessment of GIOVE Clocks"*, Proc. of the 40th PTTI, 2008.

[4] J. Dow, R. Neiland, C. Rizos, "*The International GNSS Service in a changing landscape of Global Navigation Satellite Systems*", Journal of Geodesy (2009) 83:191–198, DOI: 10.1007/s00190-008-0300-3.

[5] W. Gurtner, "*RINEX - The Receiver Independent Exchange Format - Version 3.00*", Journal of Geodesy (2009) 83:191–198, DOI: 10.1007/s00190-008-0300-3

[6] F. Vannicola, R. Beard, J. White, K. Senior, M. Largay, J. Buisson, "GPS Block IIF Atomic Frequency Standard Analysis", Proc. of PTTI 2010.

[7] F. Gonzalez, P. Waller,"Short term GNSS clock characterization using one-way carrier phase", Proc. of the IEEE Frequency Control Symposium- 21st European Frequency and Time Forum. Joint Meeting, 2007.

Statistical biases and very long term time stability analysis

François Vernotte
UTINAM
Observatory THETA of Franche-Comté
University of Franche-Comté/CNRS
41 bis avenue de l'observatoire - B.P. 1615
25010 Besançon Cedex
Email: francois.vernotte@obs-besancon.fr

Éric Lantz
Department of Optics
Femto-ST
University of Franche-Comté/CNRS
UFR Sciences - Route de Gray
25030 Besançon Cedex

***Abstract*— The prediction of very long term time stability is a key issue in various fields, such as time keeping, obviously, but also navigation and spatial applications. This is usually performed by extrapolating the measures obtained by estimators such as the Allan deviation, modified Allan deviation, Hadamard deviation, etc.**

Unlike the variance estimates, the deviation estimates are severely biased and it is of importance to perform this extrapolation from a fit over the variance estimates, and not over the deviation estimates. However, the fit should be performed on the log-log graph of the estimates, which corresponds to a least-squares minimization of the relative difference between the variance estimates and the fitting curve. But also in this case, a bias exists between the average of the log of the estimates and the log of the true value of the estimated variance.

This paper presents the theoretical calculation of this "log-log bias" based on the number of equivalent degrees of freedom of the estimates, shows simulations over a large number of realizations, and provides a reliable method of unbiased logarithmic fit. Extrapolating this fit yields a more confident assessment of the very long term time stability.

I. Introduction

The prediction of very long term time stability is usually performed by fitting and extrapolating an Allan variance (AVAR) curve. Such a curve may be modeled as:

$$\sigma_y^2(\tau) = \sum_{i=0}^{4} C_i \Phi_i(\tau) \quad \text{with} \quad \Phi_i(\tau) = \tau^{i-2}. \tag{1}$$

An extrapolation may be achieved by extending the asymptotes of such a curve beyond the largest τ value (see figure 1).

Thanks to table I, the C_i coefficient are related to the noise levels h_α of the power law noise model:

$$S_y(f) = \sum_{\alpha=-2}^{+2} h_\alpha f^\alpha \tag{2}$$

and to the linear frequency drift D_1:

$$y(t) = D_0 + D_1 t + \eta(t) \tag{3}$$

where $\eta(t)$ represents the pure random component of the frequency deviation $y(t)$.

However such an extrapolation is very sensitive to estimation error. It is then of importance to perform a reliable

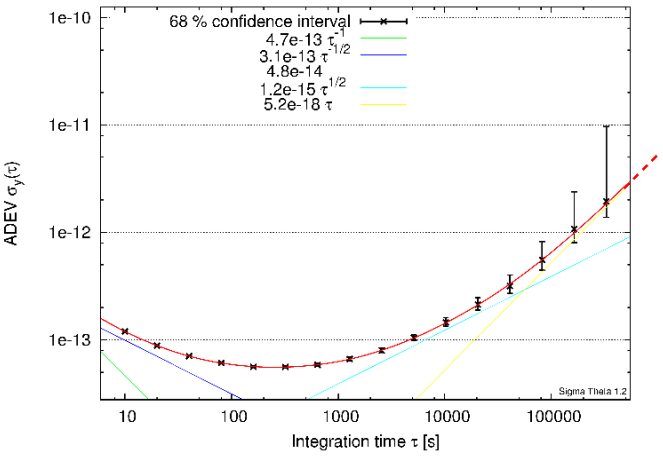


Fig. 1. Extrapolation of a least square fit.

estimation of the C_i coefficients. The weighted relative least squares are an effective method [2]:

- **weighted** in order to take into account the uncertainty Δ_j over each estimate $\hat{\sigma}_y^2(\tau_j)$,
- **relative** in order to be equivalent to a classical least square fit on the log-log plot by minimizing the relative distance between the curve and the estimate $\frac{\text{estimate}-\text{curve}}{\text{estimate}}$,

$$\sum_{j=1}^{N} \left[\frac{1}{\Delta_j} \times \frac{\hat{\sigma}_y^2(\tau_j) - \sum_{i=0}^{4} C_i \Phi_i(\tau_j)}{\hat{\sigma}_y^2(\tau_j)} \right]^2 \text{minimum.} \tag{4}$$

We know that the estimate $\hat{\sigma}_y^2(\tau)$ is an unbiased estimator of $\sigma_y^2(\tau)$, but is $\ln\left(\hat{\sigma}_y^2(\tau)\right)$ an unbiased estimator of $\ln\left(\sigma_y^2(\tau)\right)$?

II. Statistical approach

A. Statistics of the Allan variance and the Allan deviation

The definition of the Allan variance is [3], [4]:

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle (\bar{y}_2 - \bar{y}_1)^2 \right\rangle. \tag{5}$$

$S_y(f)$	$h_{+2}f^{+2}$	$h_{+1}f^{+1}$	$h_0 f^0$	$h_{-1}f^{-1}$	$h_{-2}f^{-2}$	$y(t) = D_1 t$
$\sigma_y^2(\tau)$	$\frac{3h_{+2}f_h}{4\pi^2\tau^2}$	$\frac{[1.038 + 3\ln(2\pi f_h \tau)]\, h_{+1}}{4\pi^2\tau^2}$	$\frac{h_0}{2\tau}$	$2\ln(2)h_{-1}$	$\frac{2\pi^2 h_{-2}\tau}{3}$	$\frac{D_1^2\tau^2}{2}$

TABLE I

THEORETICAL RESPONSES OF AVAR FOR THE DIFFERENT NOISE TYPES OF THE POWER LAW MODEL AND FOR A LINEAR FREQUENCY DRIFT. FOR THE HIGH FREQUENCY NOISES (f^{+2} FM AND f^{+1} FM), IT IS NECESSARY TO INTRODUCE A HIGH CUT-OFF FREQUENCY f_h [1].

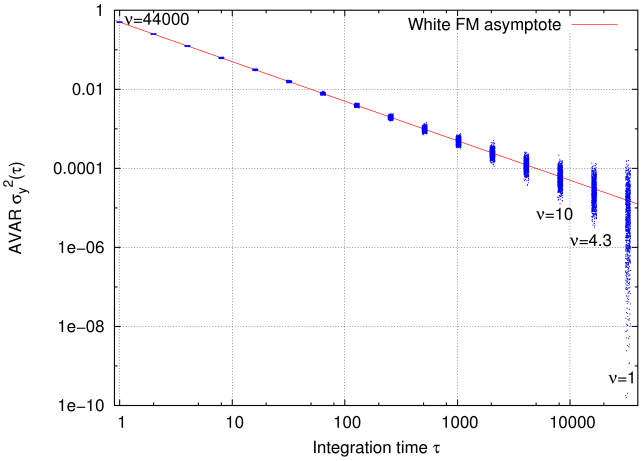


Fig. 2. Dispersion of Allan variance estimates for a pure white frequency noise and their edf ν (65536 samples @ 1 s).

The estimate of the Allan variance may be calculated as:

$$\hat{\sigma}_y^2(\tau) = \frac{1}{2N}\sum_{i=1}^{N}(\bar{y}_2 - \bar{y}_1)^2 . \tag{6}$$

- Since $\bar{y}_2$ and $\bar{y}_1$ are Gaussian centered values,
- then $\bar{y}_2 - \bar{y}_1$ is a Gaussian centered value,
- so $(\bar{y}_2 - \bar{y}_1)^2$ obeys a χ_1^2 distribution,
- therefore $\frac{1}{2N}\sum_{i=1}^{N}(\bar{y}_2 - \bar{y}_1)^2$ is a χ_ν^2 distribution with $\nu = N$ if the $\bar{y}_i$ samples are independent, with $\nu < N$ otherwise.

ν is the number of Equivalent Degrees of Freedom (EDF).

The χ_ν^2 distributions have the following properties:

- $E(\chi_\nu^2) = \nu$
- $\mathrm{Var}(\chi_\nu^2) = 2\nu$ ($\rightarrow \mathrm{Std}(\chi_\nu^2) = \sqrt{2\nu}$).

As shown by figure 2, the number of EDF reflects the AVAR estimates dispersion. A method for assessing the EDF may be found in [5].

B. Model world and measure world

Let us define:

- θ as the model parameter
- ξ as a measure of the parameter

(see figure 3.A).

As an example, we could consider the true value $\sigma_y^2(\tau = 10\text{ s}) = h_0/20$ (where h_0 is the white FM level) as a parameter; $\hat{\sigma}_y^2(\tau)$ is a measure of $\sigma_y^2(\tau = 10\text{ s})$.

1) Model world (direct problem): In the world of the model, the question is *"Knowing the parameter θ_0, how is the measure ξ distributed?"* (see figure 3.B).

However, the parameter θ_0 is precisely the unknown quantity that we want to estimate. Supposing this parameter is known has sense only in theory and simulations.

2) Measure world (inverse problem): In the world of the measures, the question is *"Knowing the measure(s) ξ_0, how to estimate a confidence interval over θ?"* (see figure 3.C).

This is the right question of the metrologist!

C. Study of a χ^2 distribution with $\nu = 1$ degree of freedom

1) Properties: A χ_1^2 distribution has the following properties [6]:

- Probability density function:

$$p(X) = \frac{e^{-X/2}}{\sqrt{2\pi X}} \tag{7}$$

- The pdf is normalized:

$$\int_0^\infty p(X)dX = 1 \tag{8}$$

- Mathematical expectation:

$$\int_0^\infty X \cdot p(X)dX = \nu = 1 \tag{9}$$

- Cumulative distribution function:

$$P(X) = \int_0^X p(y)dy = 1 - \frac{\Gamma_{\mathrm{inc}}(1/2, X/2)}{\sqrt{\pi}} \tag{10}$$

- Inverse cdf:

$$P^{-1}(\alpha) = 2\Gamma_{\mathrm{inc}}^{-1}\left(1/2, \sqrt{\pi}(1-\alpha)\right) \tag{11}$$

This relationship is useful for getting the bounds of a confidence interval.

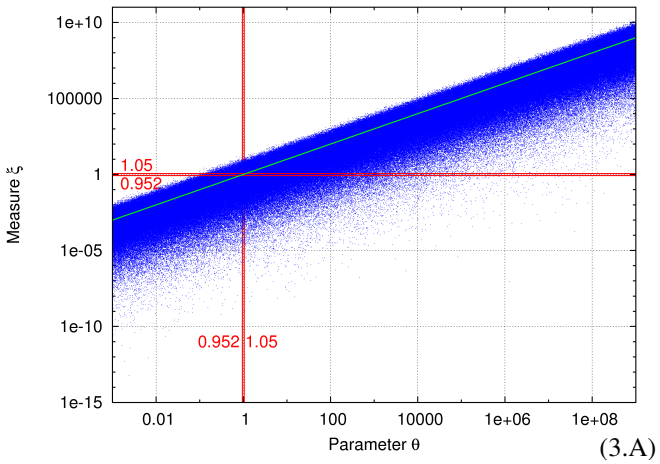
How is the reduced variable X of the relationships (7) to (10) related to the measure ξ and the model parameter θ?

2) Parameter, measure and reduced variable: Let us consider the standard χ_1^2 variable $X = x^2$ where x is a Gaussian centered standard random variable

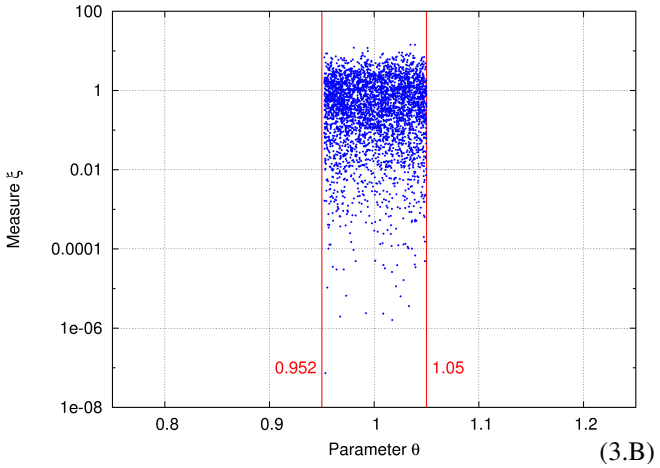
$$E(X) = \nu = 1 \quad \Rightarrow \quad E\left(\frac{X}{\nu}\right) = 1. \tag{12}$$

It is well known that:

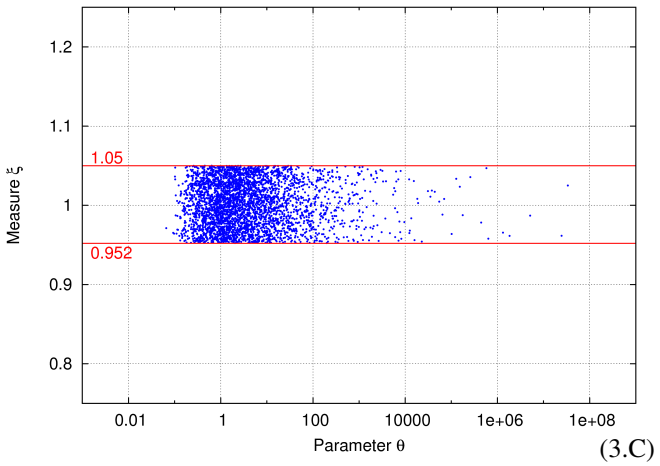
1) $\hat{\sigma}_y^2(\tau) = \xi$ is χ_1^2 distributed
2) ξ is an *unbiased* estimator of the parameter $\sigma_y^2(\tau) = \theta$.



(3.A)



(3.B)



(3.C)

Fig. 3. A: Model parameter and measure for a χ_1^2 distribution (10^6 realizations of $\theta \cdot \chi_1^2$ for $\theta = 10^k$ and $-3 < k < +9$). B: Measure dispersion for a parameter $\theta_0 \approx 1$. C: Model parameter values for a measure $\xi_0 \approx 1$.

Therefore:

$$E\left(\frac{\xi}{\theta}\right) = 1. \tag{13}$$

We can then define the reduced variable X as:

$$X = \nu\frac{\xi}{\theta} = \frac{\xi}{\theta}. \tag{14}$$

The differential dX is then:

$$dX = \frac{\partial X}{\partial \xi}d\xi + \frac{\partial X}{\partial \theta}d\theta. \tag{15}$$

3) Conditional probabilities in the model world (direct problem): The parameter θ is a constant. Let $\theta = \theta_0$, the differential dx is then:

$$dX = \frac{\nu}{\theta_0}d\xi = \frac{1}{\theta_0}d\xi. \tag{16}$$

From (16) and the relationships (7) to (9) it comes:

- Conditional probability density function:

$$p(\xi|\theta_0) = \frac{e^{-\xi/(2\theta_0)}}{\sqrt{2\pi\theta_0\xi}}. \tag{17}$$

- *ξ is unbiased:*

$$E(\xi) = \int_0^\infty \xi \cdot p(\xi|\theta_0)d\xi = \theta_0. \tag{18}$$

- Logarithmic mean:

$$\begin{aligned} e^{E[\ln(\xi)]} &= \exp\left[\int_0^\infty \ln(\xi) \cdot p(\xi|\theta_0)d\xi\right] \\ &= \frac{e^{-C}}{2}\theta_0 \approx 0.2807 \cdot \theta_0 \end{aligned} \tag{19}$$

where C is the Euler constant : $C \approx 0.57722$.

Thus, the mathematical expectation of the log of the measure doesn't converge to the log of the parameter. Therefore, there is a significant bias for log-log representation (and fit): let us call it the *"log-log bias"*.

4) Conditional probabilities in the measure world (inverse problem): The measure ξ is a constant. Let $\xi = \xi_0$, the differential dX is then:

$$dX = -\frac{\xi_0}{\theta^2}d\theta. \tag{20}$$

From (20) and the relationships (7) to (9) it comes:

- Conditional probability density function:

$$p(\theta|\xi_0) = \sqrt{\frac{\xi_0}{2\pi\theta^3}}e^{-\xi_0/(2\theta)}. \tag{21}$$

- There is **no mean value**:

$$E(\theta) = \int_0^\infty \theta \cdot p(\theta|\xi_0)d\theta \tag{22}$$

doesn't converge!

- Logarithmic mean:

$$\begin{aligned} e^{E[\ln(\theta)]} &= \exp\left[\int_0^\infty \ln(\theta) \cdot p(\theta|\xi_0)d\theta\right] \\ &= 2e^C\xi_0 \approx 3.562 \cdot \xi_0 \end{aligned} \tag{23}$$

In the measure world also, the mathematical expectation of the log of the parameter doesn't converge to the log of the measure. However, this bias is exactly the inverse of the log-log bias defined above. This implies that we just have to renormalize the measure by the log-log bias for getting both an **unbiased estimator** in the model world and a **mathematical expectation of the parameter equal to the measure** in the measure world.

D. Generalization to a χ^2_ν distribution

These results can be extended to any χ^2_ν distribution regardless of the number of degrees of freedom ν.

1) Model world:

- Reduced variable
$$X = \nu \frac{\xi}{\theta_0} \tag{24}$$
- Pdf:
$$p(\xi|\theta_0) = \frac{\nu^{\nu/2}\xi^{\nu/2-1}e^{-\frac{\nu\xi}{2\theta_0}}}{2^{\nu/2}\Gamma(\nu/2)\theta_0^{\nu/2}} \tag{25}$$
- Mean:
$$E(\xi) = \theta_0 \tag{26}$$
- Log mean:
$$e^{E[\ln(\xi)]} = \frac{2}{\nu}\exp\left[\psi^{(0)}\left(\frac{\nu}{2}\right)\right]\theta_0 \tag{27}$$
where $\psi^{(0)}(x)$ is the "digamma function" [7].

2) Measure world:

- Reduced variable:
$$X = \nu \frac{\xi_0}{\theta} \tag{28}$$
- Pdf:
$$p(\theta|\xi_0) = \frac{\nu^{\nu/2}\xi_0^{\nu/2}e^{-\frac{\nu\xi_0}{2\theta}}}{2^{\nu/2}\Gamma(\nu/2)\theta^{\nu/2+1}} \tag{29}$$
- Mean ($\nu > 2$):
$$E(\theta) = \frac{\nu}{\nu-2}\xi_0 \tag{30}$$
- Log mean:
$$e^{E[\ln(\theta)]} = \frac{\nu}{2}\exp\left[-\psi^{(0)}\left(\frac{\nu}{2}\right)\right]\xi_0 \tag{31}$$

Table II and figure 4 show the biases and the 95 % confidence interval versus edf.

III. Discussion

A. A "log-unbiased" estimator

We know that $\hat{\sigma}_y^2(\tau)$ is an unbiased estimator of $\sigma_y^2(\tau)$. However $E\left[\ln(\xi)\right] \neq \ln(\theta_0)$ and $e^{E[\ln(\xi)]} \neq \theta_0$. Similarly, $E\left[\ln(\theta)\right] \neq \ln(\xi_0)$ and $e^{E[\ln(\theta)]} \neq \xi_0$.

Let us define a new estimator $\hat{\mathcal{L}}_\theta$ as (with $\nu = 1$):

$$\hat{\mathcal{L}}_\theta = \ln(3.5625 \cdot \xi) = \ln(\xi) + 1.2705. \tag{32}$$

The mathematical expectation of $\hat{\mathcal{L}}_\theta$ is then:

$$< \hat{\mathcal{L}}_\theta >= \ln(\theta_0) \tag{33}$$

$\hat{\mathcal{L}}_\theta$ is then an **unbiased estimator** of $\ln(\theta_0)$.

ν	$E(\theta)/\xi_0$	$e^{E[\ln(\theta)]}/\xi_0$	$B_{2.5\%}$	$B_{97.5\%}$
1	∞	3.56213	0.19905	1018.26086
2	∞	1.78108	0.27108	39.49783
3	3.00000	1.44625	0.32091	13.90209
4	2.00000	1.31043	0.35896	8.25735
5	1.66667	1.23754	0.38964	6.01530
6	1.50000	1.19222	0.41524	4.84911
7	1.40000	1.16137	0.43715	4.14232
8	1.33333	1.13902	0.45624	3.67017
9	1.28571	1.12210	0.47312	3.33285
10	1.25000	1.10885	0.48821	3.07978

TABLE II
BIASES AND 95 % CONFIDENCE INTERVAL BOUNDS VERSUS EDF ν.

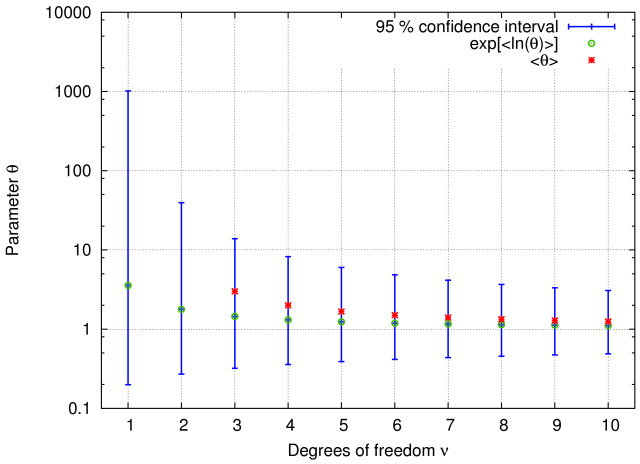


Fig. 4. Biases and 95 % confidence intervals versus edf ν.

On the other hand, for a measure ξ_0 (with $\nu = 1$):

$$E\left[\ln(\theta)\right] = \ln(3.5625 \cdot \xi_0) = \hat{\mathcal{L}}_\theta. \tag{34}$$

The mathematical expectation of $\ln(\theta)$ is equal to $\hat{\mathcal{L}}_\theta$.

Therefore, $e^{\hat{\mathcal{L}}_\theta}$ is a **log-unbiased estimator** of θ. For fitting a log-log plot, it's then better to use this "log-unbiased estimator" rather than $\ln(\xi_0)$.

B. Fitting an Allan variance curve

Figure 5 displays the AVAR estimates (black $\times$) as well as the log-unbiased estimates (green circles) of a simulated oscillator. Fitting the AVAR estimates rather than the log-unbiased estimates leads to a significant underestimation of the long term asymptotes (τ and τ^2 asymptotes) and therefore of the very long term stability of the oscillator.

C. Simulation results

1) Pure random walk FM: We simulated 3,500 sequences of pure random walk FM with a unit noise level:

- Model parameter: $h_{-2} = 1$ arbitrary unit
- Fit over classical estimates: $< \hat{h}_{-2} >= 0.911$ a.u. (-8.9 %)
- Fit over log-unbiased estimates: $< \hat{h}_{-2} >= 0.991$ a.u. (-0.9 %)

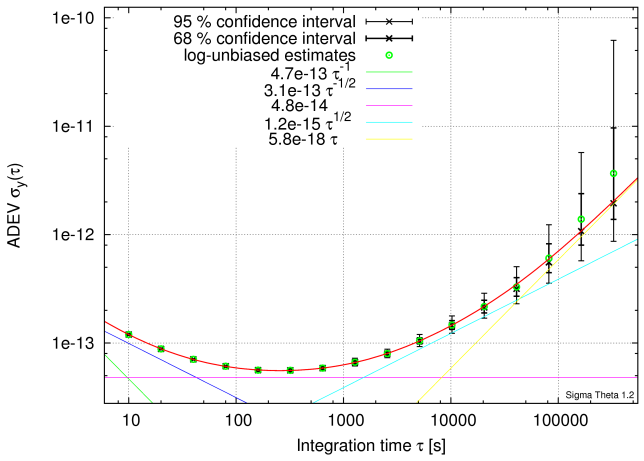


Fig. 5. Example of log-log fit (obtained with *SigmaTheta-1.2*).

The estimation of the random walk noise level is clearly better with the fit over the log-unbiased estimates, but the estimation of h_{-2} over the classical estimates is only slightly underestimated because it is based on all estimates, including those obtained for low tau values that are little biased.

2) White FM + random walk FM: We simulated 10,000 sequences of a combination of white FM and random walk FM in such a way that the random walk FM is the dominating noise over the last decade.

- Model parameter:
  - $h_0 = 1$ a.u.
  - $h_{-2} = 1.5 \cdot 10^{-5}$ a.u.
- Fit over classical estimates:
  - $< \hat{h}_0 >= 0.969$ a.u. (-3.1 %)
  - $< \hat{h}_{-2} >= 1.01 \cdot 10^{-5}$ a.u. (-33 %)
- Fit over log-unbiased estimates:
  - $< \hat{h}_0 >= 0.983$ a.u. (-1.7 %)
  - $< \hat{h}_{-2} >= 1.42 \cdot 10^{-5}$ *a.u. (-5.5 %)*

In this case, the estimation of the random walk FM noise level is drastically improved by the use of the log-unbiased estimates since the random walk asymptote is fitted over the very last tau values which are significantly affected by the log-log bias.

IV. Conclusion

AVAR estimates are not biased and are easily fitted by weighted relative least squares. However, ln(AVAR) estimates are biased, and this "log-log bias" affects the fits performed on log-log plots.

A **log-unbiased estimator** has been defined in such a way that its log is an unbiased estimator of the log of the model parameter. On a log-log AVAR plot, it's then recommended to fit the log-unbiased estimates rather than the classical AVAR estimates.

The *SigmaTheta* software package (including log-unbiased estimator) is available at the following url: `theta.obs-besancon.fr/spip.php?article103`. It is a free software developed under a GNU license for UNIX and windows systems (authors: F. Vernotte, P.Y. Bourgeois, F. Meyer).

References

[1] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources: fifteen years of progress," *Proceedings of the IEEE*, vol. 66, no. 9, pp. 1048–1075, Septembre 1978.

[2] F. Vernotte, E. Lantz, J. Groslambert, and J. Gagnepain, "Oscillator noise analysis: multivariance measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 42, no. 2, pp. 342–350, Avril 1993.

[3] D. W. Allan, "Statistics of atomic frequency standards," *Proceedings of the IEEE*, vol. 54, pp. 221–230, Février 1966.

[4] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Lesson, T. E. McCunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. Vessot, and G. M. R. Winkler, "Characterization of frequency stability," *IEEE Transactions on Instrumentation and Measurement*, vol. IM-20, pp. 105–120, Mai 1971.

[5] C. Greenhall and W. Riley, "Uncertainty of stability variances based on finite differences," in *35th annual Precise Time and Time Interval meeting*, San Diego (California, USA), December 2003, pp. 267–280.

[6] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, 4th ed., ser. Wiley Series in Probability and Statistics. Wiley & Sons, 2010.

[7] M. Abramowitz and I. Stegun, *Handbook of mathematical functions*. New York: Dover publications, 1964.

A new prediction algorithm for EAL

Gianna Panfilo, Aurelie Harmegnies, Laurent Tisserand
International Bureau for Weights and Measures (BIPM)
Pavillon de Breteuil 92312 Sevres France
Email: gpanfilo@bipm.org

***Abstract*— In this paper the new algorithm for the prediction of EAL (Echelle Atomique Libre) is presented. The effect of the application of the new model for the prediction term on the EAL frequency drift is quantified and presented together with its effect on TAI and on the atomic clock weights.**

I. Introduction

The EAL (Echelle Atomique Libre) time scale is computed at the BIPM as a weighted average of about 380 free running atomic clocks spread world wide. The weighting algorithm is designed to optimize long-term stability performance. At this stage the time and frequency continuity of the scale is provided by a time correction applied to each clock and calculated with a frequency prediction based on clock past data. In the present version of the algorithm the clock frequency prediction is taken constant over a month interval for all type of clocks. This simple approach does not take into account the frequency drift affecting the atomic clocks. Comparing EAL frequency to TT (Terrestrial Time, a time scale optimized for frequency accuracy) a frequency drift is clear. In a previous work [1] the role of the H-masers frequency drift as a possible source of the EAL frequency drift was studied and quantified. The result shown that the contribution of the H-masers to the drift is not negligible and detected a frequency drift (aging) for the cesium clocks at long term. In this paper we present a new model for the prediction term taking into account the frequency drift affecting the atomic clocks. This model is applied to all kind of atomic clocks. The use of TT as reference and the possibility to consider a longer past period to estimate the frequency drift affecting the atomic clocks are investigated. The effect of the application of the new model for the prediction term on the EAL frequency drift is quantified and presented together with the effect on the atomic clock weights. In section II a brief summary of the calculation of Coordinated Universal Time (UTC) and International Atomic Time (TAI) is presented; the present performance of EAL in term of stability with respect to TT is discussed in section III; in section IV the present and the new prediction algorithms are predented; particular attention is put on the parameter estimation given in section V. A discussion on the choice of TT and EAL as frequency references and on the length of the evaluation period is given respectively in sections VI and VII. The section VIII considers the effect of the new prediction algorithm on TAI and on the clock weights.

II. The calculation of UTC and TAI

The various time laboratories realize a stable local time scale using individual atomic clocks or a clock ensemble. Clock readings are then combined at the BIPM through an algorithm designed to optimize the frequency stability and accuracy and the reliability of the time scale beyond the level of performance that can be realized by any individual clock in the ensemble. An efficient algorithm is necessary for the statistical generation of the world reference time scale. The algorithm ALGOS [2]- [5] is used in the Time Department of the BIPM to generate, monthly, the international reference UTC. The calculation of UTC using ALGOS is carried out in three steps:

- The free atomic time scale EAL is computed as a weighted average of about 380 free-running atomic clocks spread world-wide. A clock weighting procedure has been designed to optimize the long-term frequency stability of the scale.
- The frequency of EAL is steered to maintain agreement with the definition of the SI second, and the resulting time scale is TAI. The steering correction is determined by comparing the EAL frequency with that of primary frequency standards (PFS).
- Leap seconds are inserted to maintain agreement with the time derived from the rotation of the Earth. The resulting time scale is UTC.

Different algorithms can be considered depending on the requirements on the scale; for an international reference such as UTC, the requirement is extreme reliability and long-term frequency stability. UTC therefore relies on the largest possible number of atomic clocks of different types, located in different parts of the world and connected in a network that allows precise time comparisons between remote sites. Each month the differences between the international time scale UTC and the local time scales UTC(k) maintained at the contributing time laboratories are reported at 5-day intervals in the official document called BIPM *Circular T* [6]. The original algorithm ALGOS for defining EAL gives the difference between EAL and each participant clock:

$$x_j(t) = EAL(t) - h_j(t) = \sum_{i=1}^{N} w_i \left[h'_i(t) - x_{i,j}(t)\right] \qquad (1)$$

where N is the number of participating clocks, w_i the relative weight of clock H_i, $h_i(t)$ is the reading of clock H_i at time t, and $h'_i(t)$ is the prediction of the reading of clock H_i that

978-1-61284-110-6/11/$26.00 ©2011 IEEE

serves to guarantee the continuity of the time scale. The data used by ALGOS and reported in (1) take the form of the time differences between readings of clocks, written as:

$$x_{i,j}(t) = h_j(t) - h_i(t) \tag{2}$$

The weight attributed to a clock reflects its long-term stability, since the objective is to obtain a weighted average that is more stable in the long term than any of the contributing elements [7]- [8]. Clock weights are determined from the variance of monthly average frequencies, constrained to a maximum value to avoid a few very stable clocks becoming dominant in the scale.

III. EAL FREQUENCY DRIFT ANALYSIS

TT provides a stable and accurate reference [9]- [10] for characterizing the performance of EAL frequency and for quantifying the frequency drift of the H-masers and the cesium clocks. TT is a time scale optimized for frequency accuracy, evaluated annually by making use of all available primary frequency standard data reported to the BIPM by national laboratories. By comparing the relative frequency of EAL with respect to TT a frequency drift of about -1.3x10^{-17}/day is clear. In a previous work [1] the possible reasons of the EAL frequency drift have been investigated and the role of the unmodelled drift of the H-masers has been quantified. Tests over a 3-year period (2006-2008) have been performed by calculating EAL without H-masers and the results seem to indicate that the H-masers could be responsible for 40% of the drift of EAL. We have used TT(BIPM) as the reference for characterizing the frequency drift of the H-masers and of the cesium clocks. The mean value for the H-masers frequency drift is -4x10^{-16}/day whereas for the cesium clocks it is -1x10^{-17}/day. Even if the frequency drift (or aging) observed for the cesium clocks is smaller than that for the H-masers, the cesiums represent about 80% of the weight of the clocks participating in the calculation of EAL, and consequently their impact on the frequency drift of EAL is more significant than that of the H-masers.

IV. PREDICTION ALGORITHM

In this section we will present the prediction algorithm used at present in ALGOS and the new proposed version. In the generation of a time scale, the prediction of the atomic clock behavior plays an important role, in fact the prediction is useful to avoid or minimize the time and frequency jumps of the time scale when a clock is added or removed from the ensemble or when its weight changes.

A. The prediction algorithm in ALGOS

Considering two successive intervals of TAI calculation $I_{i-1}(t_{i-1}, t_i)$ and $I_i(t_i, t_{i+1})$ we impose several conditions on the prediction term $h'_i(t)$ at time t_i to avoid or minimize time and frequency jumps in the resulting time scale. ALGOS operates in a post-processing mode, treating as a whole measurements taken over a basic period of T=30 or 35 days. For each one-month period, the results are the quantities $x_i(t_i + nT/6)$ (or $x_i(t_i + nT/7)$) with $n = 0, ..., 6$ (or $n = 0, ..., 7$) for each clock H_i. The least squares slope of these quantities is referred to as the frequency $B_{i,I_i}(t_i + T)$ of the clock H_i, relative to EAL, for the one-month interval $I_i = [t_i, t_i + T]$. The correction term $h'_i(t)$ for clock H_i is the sum of two terms:

$$h'_i(t) = a_{i,I_i}(t_i) + B_{ip,I_i}(t - t_i) \tag{3}$$

where $a_i(t_i)$ is the time correction relative to EAL of clock H_i at date t_i:

$$a_{i,I_i}(t_i) = EAL(t_i) - h_i(t_i) = x_i(t_i) \tag{4}$$

$B_{ip,I_i}(t - t_i)$ is the frequency of clock H_i, relative to EAL, predicted for the period $I_i = [t_i, t]$ and $(t - t_i) = nT/6$ (or $nT/7$) with $n = 0, ..., 6$ (or $n = 0, ..., 7$). At present, the predicted frequency $B_{ip,I_i}(t)$ for t included in $I_i = [t_i, t_i + T]$ is kept constant for the whole one-month interval. It is chosen to be the frequency computed for the previous one-month interval $I_{i-1} = [t_i - T, t_i]$, that is to say, the least squares slope of $x_i(t_i - T + nT/6)$ (or $x_i(t_i - T + nT/7)$) with $n = 0, ..., 6$ (or $n = 0, ..., 7$). This is a one-step linear prediction written as $B_{ip,I_i}(t) = B_i(t_i)$ for $t = t_i + nT/6$ (or $t = t_i + nT/7$), $n = 0, ..., 6$ (or $n = 0, ..., 7$). The assumption that clock H_i is most likely to behave over the coming one-month interval as it did on the previous one is linked to the choice of the sample duration T.

B. The new prediction algorithm

To evaluate the effect of the prediction in EAL frequency drift and to better describe the long term behavior of the H-masers and the cesium clocks we present here a new clock frequency prediction algorithm [11]- [12]. Considering two successive intervals of calculation of TAI $I_{i-1}(t_{i-1}, t_i)$ and $I_i(t_i, t_{i+1})$ we impose several conditions on the prediction term h'_i at time t_i to avoid or minimize time and frequency jumps in the resulting time scale. Instead of the linear term used in ALGOS we consider a quadratic behavior:

$$h'_i(t) = a_{i,I_i}(t_i) + B_{ip,I_i}(t - t_i) + \frac{1}{2} C_{ip,I_i}(t)(t - t_i)^2 \tag{5}$$

To the linear relation (3) we add the last term of (5) to describe the frequency drift affecting the H-masers and the long term behavior of cesium clocks. To evaluate the parameters in (5) we assume at time t_i the following conditions on h'_i:

1) no time steps by imposing the continuity of EAL;
2) no frequency steps by imposing the continuity to the first derivative of EAL;
3) no change in frequency drift by imposing the continuity to the second derivative of EAL.

We obtain a system of three equations with three variables and solving this system we obtain that the relation (5) can be expressed by the following relationship:

$$\begin{aligned} \widehat{h}'_i(t) = {} & \widehat{a}_{i,I_i}(t_i) + \widehat{B}_{ip,I_i}(t - t_i) \\ & + \frac{1}{2}\widehat{C}_{ip,I_{i-1}}(t)(t_i - t_{i-1})(t - t_i) + \frac{1}{2}\widehat{C}_{ip,I_i}(t)(t - t_i)^2 \end{aligned} \tag{6}$$

The physical meaning of the terms present in (6) is:

- $\widehat{a}_{i,I_i}$ is the estimation of the time correction relative to EAL of clock H_i at date t_i
- $\widehat{B}_{ip,I_i}$ is the estimation of the frequency of clock H_i, relative to EAL, predicted for the period $[t_i, t]$
- $\widehat{C}_{ip,I_i}$ is the estimation of the frequency drift of the clock H_i, relative to a frequency reference, predicted for the period $[t_i, t]$
- $\widehat{C}_{ip,I_{i-1}}$ is the estimation of the frequency drift of the clock H_i, relative to a frequency reference, predicted for the period $[t_{i-1}, t_i]$.

Two aspects are very important in this relation; the first is that we consider a constant drift during the calculation interval; the second is that with respect to the linear prediction the frequency does not remain constant during the interval. The parameters described above are considered constant during the calculation interval, but to avoid a frequency jump the frequency at the end of an interval and at the beginning of the next one should be the same. The frequency at the end of the interval is expressed by the addition of the second and the third term in the relation (6):

$$\widehat{B}_{ip,I_i}(t-t_i) + \frac{1}{2}\widehat{C}_{ip,I_{i-1}}(t)(t_i - t_{i-1})(t-t_i) \qquad (7)$$

and in particular $\widehat{B}_{ip,I_i}(t-t_i)$ is the estimated frequency at the middle of the interval.

V. The Parameter Estimation

The relationship (6) describing the correction applied to the atomic clocks is composed by three parameters. A very important point of this work is the optimization of the estimation procedure. In particular the estimation of the frequency drift needs particular attention. The estimation of the time correction $a_i(t_i)$, added to avoid time jumps, is given by the last known point of the difference between EAL and each clock and expressed by the following formula:

$$\widehat{a}_{i,I_i}(t_i) = EAL(t_i) - h_i(t_i) = x_i(t_i) \qquad (8)$$

From the theoretical point of view the best estimation of the mean frequency B_{ip,I_i} is the difference between the last and the first point on the interval duration of the clock data respect to EAL as expressed by the following relationship:

$$\widehat{B}_{ip,I_i} = \frac{x_i(t_{i+1}) - x_i(t_i)}{t_{i+1} - t_i} \qquad (9)$$

In order to estimate the frequency drift of atomic clocks under the hypothesis that it remains constant on one month interval the choice of the frequency reference is very important. In the following part of the paper we will discuss about the difference on the use of EAL and TT as frequency reference and we will show the improvement in EAL stability when TT is used. If the phase data $(EAL - h_i)$ have been used to estimate the time and frequency corrections the frequency data y_{TT-h_i} of the clock respect to TT are used to estimate the drift. The calculation of the drift is carried out by three steps:

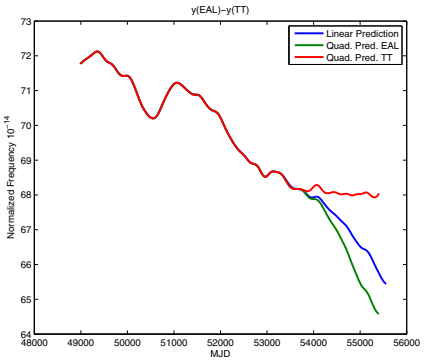


Fig. 1. y(EAL)-y(TT) calculated by applying the linear prediction (blue line), the quadratic prediction with EAL as frequency reference (green line) and the quadratic prediction with TT as frequency reference (red line).

- The frequency data of the clock respect to TT (y_{TT-h}) is filtered with a moving average of 15 days in order to remove the white frequency noise and better estimate the frequency drift. The obtained filtered data are called $\overline{y}_{TT-h_i}$
- The drift for all kind of clocks is estimated by using 4 months of $\overline{y}_{TT-h_i}$
- The relationship used to estimate the frequency drift is:

$$\widehat{C}_{ip,I_i} = \frac{\overline{y}_{TT-h_i,i}(t_{i+1}) - \overline{y}_{TT-h_i,i}(t_i)}{t_{i+1} - t_i} \qquad (10)$$

In the described procedure the biggest impact on the estimation of the drift is given by the choice of the frequency reference and by the choice of the length of the evaluation interval. The next part of this paper will be dedicated to discuss and develop these points.

VI. EAL and TT as Frequency Reference for the Drift Estimation

In this section the use of EAL an TT as frequency reference is discussed. TT-h_i is determinated by the sum of TT-EAL, given by the calculation of TT, and EAL-h_i. To evaluate the EAL performance, the EAL frequency (y(EAL)) is compared to TT frequency (y(TT)) which is a more accurate reference. In Fig. 1 the results of a test performed using 5 years of data from January 2006 to July 2010 are reported. EAL is obtained with the linear prediction (blue line) and quadratic predictions with the drift evaluated considering EAL (green line) and TT (red line) as frequency reference. The new prediction algorithm has been applied by using the described procedure with 6 months of past data to evaluate the frequency drift. The corresponding Allan deviations are reported in Fig.2. It is clear from the stability analysis that the use of EAL as frequency reference for the prediction has a negative impact on EAL performance while when TT is used as frequency reference the stability of EAL is clearly improved and the frequency drift affecting EAL is almost completely removed. To investigate the reasons of this result the comparison between the frequency drifts obtained by using EAL and TT as frequency references has

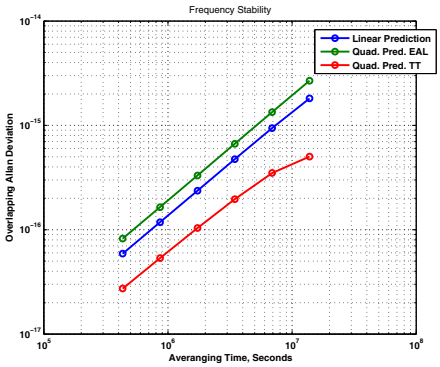


Fig. 2. Stability analysis of y(EAL)-y(TT) corresponding to the data presented in the Fig. 1.

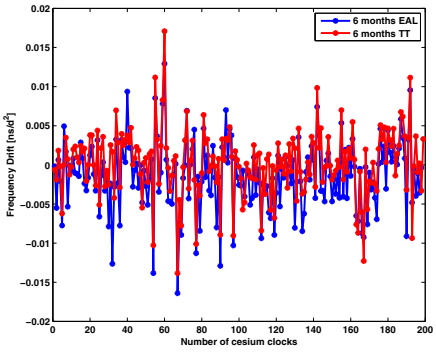


Fig. 3. With the red line the frequency drift calculated by using TT and with the blue line those obtained by using EAL are reported for the cesium clocks.

been made. In the next paragraph a summary of these studies will be presented. Considering the results obtained by using EAL and TT as frequency references it has been decided to choose TT as frequency reference to estimate the frequency drift of the H-Masers and the aging affecting the cesium clocks.

A. *Drift Analysis*

In this paragraph the difference between the drifts evaluated by using EAL and TT as frequency reference is investigated and explained. During the test period (2006- July 2010) the frequency drift has been calculated for all kind of clocks. Considering the mean value of the drift for about 200 cesium clocks (Fig. 3) and 35 H-Masers (Fig. 4) on the calculation period obtained by using EAL (blue line) and TT (red line) a systematic difference is observed. This is due to the fact that the freqency drift of EAL affects the estimation of the clock behaviour so invalidating it as a good reference.

VII. The Length of the Evaluation Period

In this section a study on the optimization of the length of the evaluation interval to estimate the frequency drift is presented. Considering the correlation at short term between TT and EAL it is not possible to use one month of data to

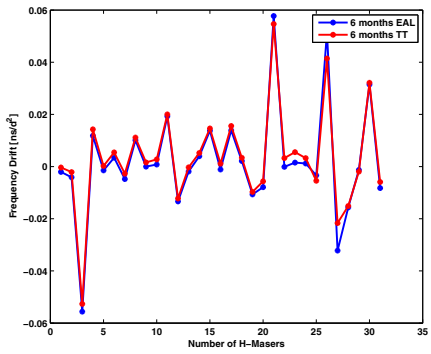


Fig. 4. With the red line the frequency drift calculated by using TT and with the blue line those obtained by using EAL are reported for the H-Masers.

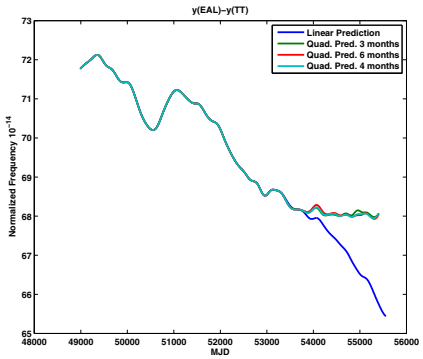


Fig. 5. y(EAL)-y(TT) obtained by using the quadratic algorithm with 3 (green line), 4 (light blue line), and 6 (red line) months of past data to evaluate the drift.

estimate the drift. A test was performed on 5 years of data by using 3, 4 and 6 months of past data to evaluate the drift. The results are compared using the stability of EAL as reference parameter. In Fig. 5 y(EAL)-y(TT) obtained by using the quadratic prediction algorithm with 3 (green line), 4 (light blue line), and 6 (red line) months of past data to evaluate the drift is reported. The zoom of this figure is reported in Fig. 6 in order to show the impact of the use of difference length of past period in the new prediction algorithm to evaluate the frequency drift during the test period.

From the results reported in Fig. 5 we can conclude that the EAL frequency drift is almost completely removed. If we compare the results obtained by using a different number of the months (Fig. 6) to evaluate the frequency drift it seems to have similar performance, in fact not big differences are observed among the results. To better evaluate this effect the stability analysis of EAL respect to TT in term of frequency is reported for the previous data. A substantial improvement of the EAL performance by using the quadratic prediction in term of stability is observed in Fig. 7. The Allan deviation of EAL respect to TT shows an aging or residual drift of about $1\text{x}10^{-18}$/day (1 order of magnitude less than the present case) not easily characterized and a possible random walk. The value

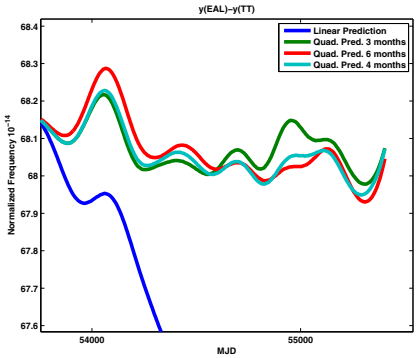


Fig. 6. A zoom of y(EAL)-y(TT) obtained by using the quadratic algorithm with 3 (green line), 4 (light blue line), and 6 (red line) months of past data to evaluate the drift.

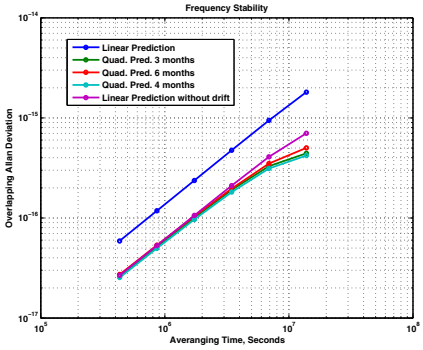


Fig. 7. The stability analysis of the data reported in Fig. 5.

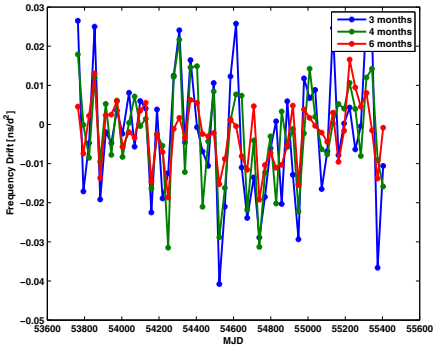


Fig. 8. The evolution of the drift obtained by using 3 (blue line), 4 (green line), and 6 (red line) months of past data for one cesium clock is reported.

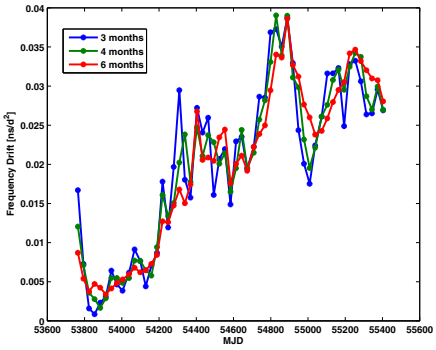


Fig. 9. The evolution of the drift obtained by using 3 (blue line), 4 (green line), and 6 (red line) months of past data for one H-maser is reported.

of the frequency stability at 5 days by using a linear prediction is between 5.9x10^{-17} and 4.8x10^{-16} at 40 days. With the new prediction algorithm the obtained values for the stability are 2.5x10^{-17} at 5 days and 1.8x10^{-16} at 40 days. In Fig. 7 with the violet line the stability analysis of the data obtained with the linear prediction algorithm and a linear frequency drift removed is reported. The obtained results are consistent with those obtained by using the quadratic prediction so we can conclude that the linear frequency drift affecting most of the clocks composing the ensemble is removed. A non linear component of the drift could be present but since it cannot be easily estimated it has not been considered for the moment. We can also conclude that the EAL performance does not change by using a different past period. In the following subsection the frequency drifts affecting the cesium clocks and the H-masers will be caracterized to show the difference between them and to make some conclusions about the optimization of the length of the past period to estimate the frequency drift.

A. Drift analysis of the cesium clocks and H-masers

During the test period (5 years) the frequency drift has been calculated in one month interval for cesium clocks and H-masers and time series are obtained. In Figs. 8 and 9 one example of the behavior of the frequency drift obtained by using 6 (red line), 3 (blue line) and 4 (green line) months of past data respectively for one cesium clock and one H-maser is presented. By analyzing the drift evolution of the cesium clocks (see Fig. 8) very stable drifts or drift carachterized by very small instabilities are observed. The situation is very homogeneous for most of the clocks. A very different situation is observed for the H-masers (see Fig.9) in fact the behavior of the drift can be characterized by important instabilities, and also important differences are present between different clocks. In the case of H-masers a long past period (6 months) is not recommended to estimate the frequency drift.

B. The length of the evaluation period

Considering that:

- the stability of EAL calculated with the new prediction algorithm does not change significantly by using 3, 4 or 6 months of past data to evaluate the frequency drift.
- the use of a long past period is not recommended (H-Masers)
- In the weight algorithm the clocks are studied during 4 months before being included in the calculation.

We conclude that the period equal to 4 months is adeguate for the drift prediction.

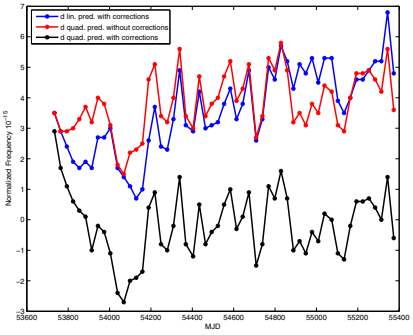


Fig. 10. The value of d as published in Circular T is reported with the blue line. The red line shows the values of d calculated by applying new prediction algorithm without frequency corrections and with the dark line the values of d obtained with corrections equal to -0.6x10^{-15} for the first 7 months.

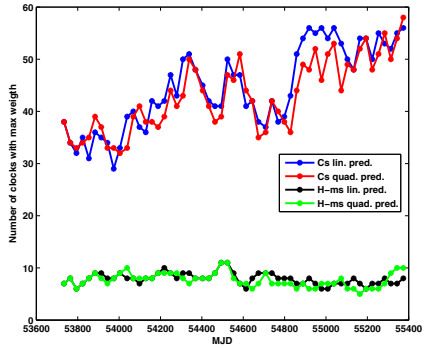


Fig. 11. The number of the clocks (H-masers and Cesium clocks) with maximum weight are reported as obtained by using the linear and the quadratic predictions. The results are reported respectively with the dark and green lines for the H-masers and with the blue and red lines for the Cesium clocks.

VIII. The effect of the new prediction algorithm on TAI and on the weights

In this paragraph we will present an analysis of the effects of the new prediction algorithm on TAI behavior and on the weights.

A. The effect on TAI

TAI is a realization of coordinate time TT. In Section 4 of *Circular T* [6] the fractional deviation d of the scale interval of TAI from that of TT based on all available PFS measurements is reported. In Fig. 10 the values of d as published in *Circular T* are plotted with the blue line. The red line shows the values of d calculated by applying the new prediction algorithm without frequency corrections; the dark line shows the values of d obtained again by using the new prediction algorithm with frequency corrections equal to -0.6x10^{-15} applied only for 7 months of the test period (5 years), after that no corrections are applied. We can conclude that no big differences are observed between TAI and EAL after the application of the new prediction algorithm and only small adjustments of the frequency of EAL are required to obtain TAI.

B. The effect of the weights

The weight algorithm has not been changed during the test period and it seems that the new prediction algorithm has a little effect on the weights. In Fig. 11 the number of the clocks (H-masers and Cesium clocks) with the maximum weight are reported as obtained by using the linear and the quadratic predictions. The results are reported respectively with the dark and green lines for the H-masers and with the blue and red lines for the Cesium clocks.

IX. Conclusion

A new mathematical expression for the prediction has been found including the treatment of the frequency drift. A test for the 5-year period has been done applying quadratic prediction to all atomic clocks and by using 4 months of past data to evaluate frequency drift. Different past periods to evaluate the frequency drift have been tested and the obtained results have been compared. The EAL perfomance has clearly improved and its frequency drift has been almost completely removed. EAL still shows a residual drift which is 1 order of magnitude smaller than before that it is not easily estimated and that is has not been removed. This residual drift can be due to a non-linear signature caracterizing the drift and not included in the model. The effect on TAI of the new prediction algorithm is discussed and now only small adjustments of the frequency of EAL could be necessary to obtain TAI. The new prediction algorithm has little impact on the weights; a revision of the weight algorithm is advisable to increase the role of the H-masers in the time scale ensemble for the future.

Acknowledgment

The authors would like to thank Felicitas Arias and Gerard Petit for the helpful discussions.

References

[1] G. Panfilo and E.F. Arias. "Studies and possible improvements on EAL algorithm". UFFC, Vol. 57, No.1, January 2010, pp. 154-160.
[2] G. Panfilo and E. F. Arias, "Algorithms for TAI". UFFC, Vol. 57, No.1, January 2010, pp. 140-150.
[3] C. Thomas, P. Wolf and P. Tavella, Time Scales, BIPM Monographie 94/1, 1994.
[4] B. Guinot and C. Thomas, "Establishment of international atomic time" Annual Report of the BIPM Time Section, vol. 1, 1988.
[5] P. Tavella and C. Thomas, "Comparative study of time scale algorithms" Metrologia, vol. 28(2), pp. 57-63, 1991.
[6] Circular T: http://www.bipm.org/jsp/en/TimeFtp.jsp?TypePub=publication
[7] C. Thomas and J. Azoubib "TAI computation: study of an alternative choice for implementing an upper limit of clock weights". Metrologia 33, pp. 227-240 (1996).
[8] B. Guinot "Some properties of Algorithms for atomic time scales". Metrologia 24, 195-198 (1987).
[9] J. Azoubib, M. Graveaud and B. Guinot. "Estimation of the scale unit duration of time scales". Metrologia, vol. 13, pp. 87-93 (1977).
[10] G. Petit, A new realization of Terrestrial Time, Proc. 35th PTTI, pp. 307-316 (2003).
[11] G. Panfilo. "The mathematical modelling of the atomic clock error with application to time scales and satellite systems". PhD Thesis in Metrology, Turin 2006.
[12] G. Panfilo and P. Tavella. "Atomic clock prediction based on stochastic differential equations". Metrologia, vol. 45(6), pp. S108-S116, 2008.

Simulation Of Future GPS Clock Scenarios With Two Composite Clocks

Matthias Suess (DLR)
German Aerospace Centre (DLR)
Oberpfaffenhofen, Germany

Demetrios Matsakis (USNO)
United States Naval Observatory (USNO),
Washington, D.C., USA

***Abstract*— The performance of the GPS constellation is modeled using four possible ensembles of monitor site clocks and two different versions of the composite clock algorithm. The analysis of our previous paper [1] is scrutinized; the conclusion remains that the improvements due to improved station clocks are minimal if the "Greenhall" algorithm is used for the composite clock. The response to a satellite and monitor station time step is investigated. Furthermore, the proportional process noise case is studied for different initial covariances.**

I. INTRODUCTION

The GPS control segment consists of 17 monitor stations. The monitor stations at USNO, Washington DC, and Schriever AFB are referenced to masers, while all other stations are equipped with HP 5071 cesium clocks **[7]**. USNO is planning to replace the maser references in Washington and Schriever with rubidium fountains. This paper uses a simplified model of the operational GPS Kalman Filter, limited to clock estimates, to investigate the impact of improved monitor station clocks on the GPS timing performance.

II. SIMULATION OF THE GPS WORKFLOW

Using the GPS toolkit software **[2]**, the positions of 31 satellites are modeled based on the GPS almanac file of week 531. The satellites are equipped with Rubidium Atomic Frequency Standards (RAFS). The measurements are simulated with a 15-minute spacing, a white Gaussian noise of 0.7 ns (1 sigma), and an elevation angle mask of 20°. The measurement noise models pseudorange measurements, which provide time information but are approximately 100 times noisier than carrier phase data.

The clock parameters computed by the Kalman Filter are uploaded to each satellite once a day as a broadcast model. In this simulation we assume the update occurs when the satellite is in view over Schriever AFB and the clock broadcast parameters are older than 1 day.
For this work, only clock parameters are solved for. In the operational filter, satellite orbit and other parameters are simultaneously included and the correlations among the full set of parameters would be expected to decrease the precision of the derived solutions. This would most likely decrease the differences we have computed between the clock models. However, double-differencing between satellite and station pairs allows one to solve for all other parameters without estimating clocks, and to solve for clocks in a second solution using the non-clock parameters as inputs.

A. Simulation of three Improved Monitor Station Clocks and Satellite Clocks

Each clock is modeled with a three-state vector X(t), which includes phase, frequency, and drift offset [3, 4].

$$\frac{dX}{dt}(t) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X(t) + W(t) \qquad (1)$$

$$EW(t)W^T(t) = \begin{pmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{pmatrix} = Q(0s) \qquad (2)$$

with a deterministic zero start value, $X(t_0)=0$, and W(t) is the clock process noise. The three q's measure the noise contributions from white frequency modulation (WFM), random walk frequency modulation (RWFM), and random run frequency modulation (RRFM).

Four different types of monitor station clocks are simulated: cesium, active hydrogen maser (AHM), fountain clock and optical fountain clock. Their stochastic components (q's) together with the RAFS are specified in Table 1, and were computed as in [1]. Typical Allan deviations are shown in Figure 1. The values for these q's should not be taken as authoritative by any means, nor should the associated Allan deviations be assumed as anything but a rough approximation of the clocks.

TABLE 1. CLOCK NOISE MODELS.

	q_1 WFM [s^2/ s]	q_2 RWFM [s^2/ s^3]	q_3 RRFM [s^2/ s^5]
Cesium	2.50e-23	4.44e-37	5e-53
AHM	2.8e-26	1.1e-35	4.4e-51
Fountain	2.5e-26	1.1e-37	1.1e-55
Optical Ftn	4.4e-27	1.1e-37	1.1e-55
RAFS	1.0e-24	1.1e-35	2.8e-46

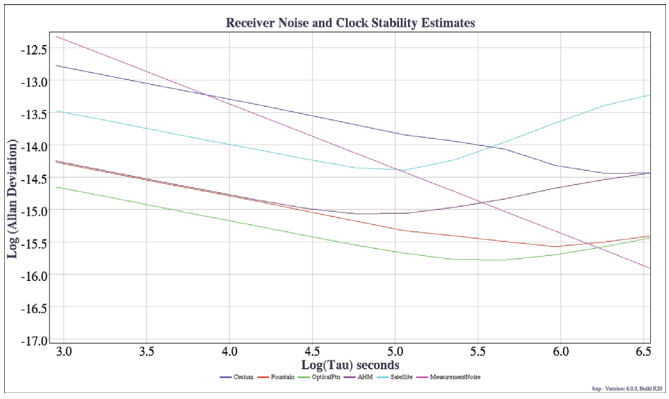


Figure 1. Allan deviation of one individual clock simulation of each modeled type. It is assumed that the necessary resources are expended to keep the masers and the optical fountains within their environmental specifications, and that the receiver's white phase pseudorange noise is diluted by 8 independent simultaneous satellite measurements.

For numerical issues it is found to be useful to change the units of the q values (Table 1), which are conventionally given in second, to nanoseconds and days:

$$\left(\frac{10^{18}}{86400^{-1}} q_1 \quad \frac{10^{18}}{86400^{-3}} q_2 \quad \frac{10^{18}}{86400^{-5}} q_3 \right) \left[\frac{ns^2}{d} \quad \frac{ns^2}{d^3} \quad \frac{ns^2}{d^5} \right] \qquad (3)$$

B. Composite Clock Algorithms – Brown and Greenhall

Two composite clock algorithms **[5,6]** are applied to estimate the three states of the satellite and control station clocks with respect to its implicitly defined time scales. Since N clocks involve maximal N-1 linearly independent measurements, each composite clock algorithm includes a method to prevent the formal covariance of the solution parameters from growing without bound.

1) The Brown Composite Clock

It can be shown **[5]** that no measured quantity is affected if the "Brown" Kalman Filter covariance estimates B(C(t)) are reduced at each iteration step as follows:

$$B(C(t)) = C(t) - \bar{H}(\bar{H}^T C^{-1}(t)\bar{H})^{-1}\bar{H}^T \qquad (4)$$

with

$$\bar{H} = \begin{pmatrix} 1_N & 0_N & 0_N \\ 0_N & 1_N & 0_N \\ 0_N & 0_N & 1_N \end{pmatrix} \in R^{3Nx3} \text{ and } 1_N = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in R^N \qquad (5)$$

0_N is a vector of N zeros. Each of the nine NxN submatrices C_{ij} with $i, j \in \{x, y, d\}$ is subtracted by a NxN matrix with a common element defined by c_{ij}

$$C = \begin{pmatrix} C_{xx} & C_{xy} & C_{xd} \\ C_{yx} & C_{yy} & C_{yd} \\ C_{dx} & C_{dy} & C_{dd} \end{pmatrix} \rightarrow B(C) = \begin{pmatrix} C_{xx} & C_{xy} & C_{xd} \\ C_{yx} & C_{yy} & C_{yd} \\ C_{dx} & C_{dy} & C_{dd} \end{pmatrix} - \begin{pmatrix} 1_{N,N}c_{xx} & 1_{N,N}c_{xy} & 1_{N,N}c_{xd} \\ 1_{N,N}c_{yx} & 1_{N,N}c_{yy} & 1_{N,N}c_{yd} \\ 1_{N,N}c_{dx} & 1_{N,N}c_{dy} & 1_{N,N}c_{dd} \end{pmatrix} \qquad (6)$$

with

$$(\bar{H}^T C^{-1}(t)\bar{H})^{-1} = \begin{pmatrix} c_{xx} & c_{xy} & c_{xd} \\ c_{yx} & c_{yy} & c_{yd} \\ c_{dx} & c_{dy} & c_{dd} \end{pmatrix} \text{ and}$$

$$1_{N,N} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in R^{NxN} \qquad (7)$$

The Brown reduction results in a series of reduced covariances which are called:

$$C_B(t_k) = B(F(C_B(t_{k-1}))) \qquad (8)$$

where the matrix function F models the Kalman filter iteration. The initial values are

$$\hat{X}_B(t_0) = 0 \text{ and } C_B(t_0) = 1e^{10}Q(\tau) \qquad (9)$$

The resulting ensemble estimate and estimation error (corrected ensemble) are denoted:

$$\hat{X}_B = \begin{pmatrix} \hat{x}_B \\ \hat{y}_B \\ \hat{d}_B \end{pmatrix} \text{ and } X_B = \begin{pmatrix} x \\ y \\ d \end{pmatrix} - \begin{pmatrix} \hat{x}_B \\ \hat{y}_B \\ \hat{d}_B \end{pmatrix} \qquad (10)$$

2) The Greenhall Composite Clock

The reduction method by Greenhall is similar; however, it works only on the submatrix C_{xx} of the phase states **[6]**. Using the auxiliary matrix

$$S = \begin{pmatrix} A & 0 & 0 \\ 0 & I_N & 0 \\ 0 & 0 & I_N \end{pmatrix} \qquad (11)$$

with

$$A = (I_N - 1_N 1_N^T (1_N^T C_{xx}^{-1} 1_N)^{-1}) \qquad (12)$$

the Greenhall reduction is defined by

$$G(C(t)) = SC(t)S^T \qquad (13)$$

The matrix operation only affects the submatrixes involving the phase-parts:

$$C = \begin{pmatrix} C_{xx} & C_{xy} & C_{xd} \\ C_{yx} & * & * \\ C_{dx} & * & * \end{pmatrix} \rightarrow G(C) = \begin{pmatrix} C_{xx}^* & AC_{xy} & AC_{xd} \\ C_{xy}A^T & * & * \\ C_{xd}A^T & * & * \end{pmatrix} \qquad (14)$$

with

$$C_{xx}^* = C_{xx} - 1_N (1_N^T C_{xx}^{-1} 1_N)^{-1} 1_N^T \qquad (15)$$

Running the Kalman filter with this reduction method changes the phase estimates, but not the frequency and drift estimates. The reduced covariance of the Greenhall composite clock is defined by:

$$C_G(t_k) = G(F(C_G(t_{k-1}))) \qquad (16)$$

Analogous to the Brown filter, the Greenhall estimates and estimation errors or corrected ensemble are denoted $\hat{X}_G$ and $\tilde{X}_G$.

C. Three Ground Segment Models

Table 2 outlines three different ground segment models for the stability of the monitor station clocks.

TABLE 2. MODEL DEFINITION.

Model C	Cesiums at every station except for AHMs at Schriever and USNO
Model M	AHMs at every station
Model F	Cesiums at every station except at Schriever and USNO, where USNO will maintain rubidium-based atomic fountains
Model O	As in Model F, except that the "optical fountains" have a factor of 2.4 less WFM noise

Model C describes the current situation in GPS, Model F represents the rubidium atomic fountains which the USNO is planning to install at its two facilities in the near future, and the other models represent possible future scenarios.

D. Time Scales

Both composite clocks define their time scales within the Kalman filter. In case of the Greenhall composite clock, the time scale weights are given by

$$w_G^T(t) = (1_N^T C_{G,xx}^{-1}(t) 1_N)^{-1} 1_N^T C_{G,xx}^{-1}(t) \qquad (17)$$

with

$$C_G(t) = \begin{pmatrix} C_{G,xx}(t) & C_{G,xy}(t) & C_{G,xd}(t) \\ C_{G,yx}(t) & C_{G,yy}(t) & C_{G,yd}(t) \\ C_{G,dx}(t) & C_{G,dy}(t) & C_{G,dd}(t) \end{pmatrix} \quad (18)$$

and its time scale is

$$sys_G(t) = w_G^T(t)(x(t) - \hat{x}_G(t)) \quad (19).$$

In contrast, the time scale of the Brown composite clock is the phase of the implicit ensemble mean (IEM). Using the Brown time state weight of all states

$$w_B^T(t) = ((\bar{H}^T C_B^{-1}(t)\bar{H})^{-1}\bar{H}^T C_B^{-1}(t))(1,:) \quad (20),$$

the IEM is defined by

$$sys_B(t) = w_B^T(t)\begin{pmatrix} x(t) - \hat{x}_B(t) \\ y(t) - \hat{y}_B(t) \\ d(t) - \hat{d}_B(t) \end{pmatrix} \quad (21)$$

Unlike the Greenhall model, the Brown time scale includes the corrected frequency and drift states.

III. IMPACT OF RECEIVER NOISE

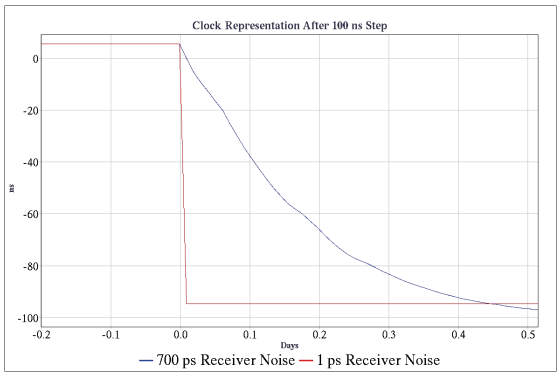


Figure 2. Impact of receiver noise on Kalman filter estimate.

It is well known that the Kalman Filter acts as a smoother to balance the noise in the data against the process noise in the parameters. The impact of this smoothing upon impulse response is shown in Figure 2.

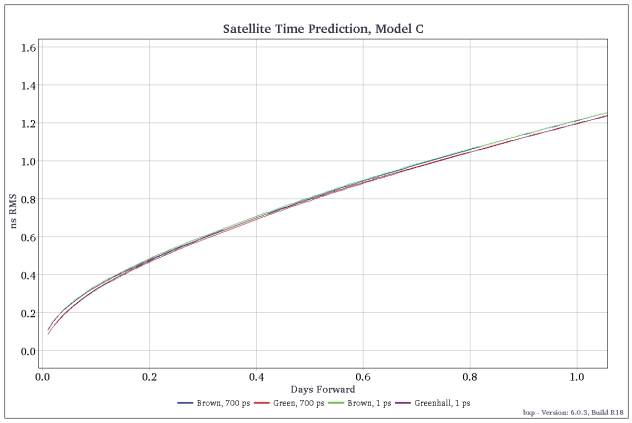


Figure 3. Impact of receiver noise on prediction error.

Figure 3 shows the accuracy of satellite prediction as a function of age of data. Lower curves are 700 ps receiver noise; upper are 1 ps noise case; the apparent discrepancy is due to the smoothing.

IV. TIMESCALE RESPONSE ON PHASE JUMPS

In order to investigate the impact of a phase step on the time scales a phase step test is performed:

- Monitor station step of 100 ns at 8 day (clock type depends on model)
- Satellite clock step of 100 ns at 10 day

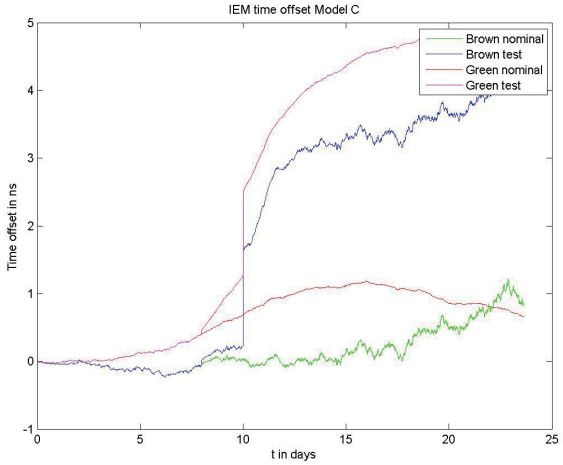


Figure 4. Phase step impacts on Model C.

Figure 4 compares the resulting IEMs of Model C to the nominal scenario. The impact of the monitor step at day 8, which is a cesium, is almost insignificant (~<0.5 ns) and at the receiver noise level. However, the result of satellite clock step is around (~2 ns). The different step size is due to the different weighting of the clock phases. Both algorithms act in a similar manner.

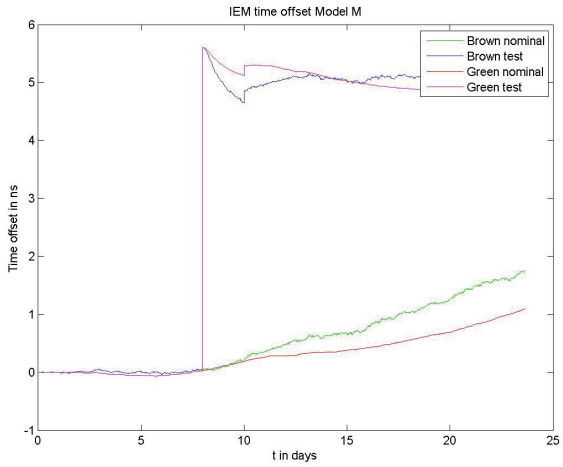


Figure 5. Phase step impact on both IEMs of Model M.

Figure 5 shows the phase step impact on Model M. The monitor station step (AHM) is significant and of size ~6 ns which is reflects the weighted step size 100ns/17=5.88ns using the fact that the 17 AHM carry all the weight. Vice versa, the satellite clock step is less, and ~0.5 ns. The Model M situation is switched relative to the Model C situation which again is due to the different weighting of the ensemble clocks. In the case of Model M the monitor stations carry the most weight, which is not the case for Model C.

V. Comparison of Brown And Greenhall assuming the PPN Case And Its initalizationi dependency

Our previous paper [1] empirically showed that the two covariance reduction methods of Brown and Greenhall can be commutatively combined without affecting the clock difference, frequency, or drift estimates. Within [5] Brown studied a special type of ensemble which is called the proportional process noise (PPN) case. Each clock noise covariance can be expressed by a scalar value and a 3x3 covariance:

$$Q_i = s_i Q_0 \tag{22}$$

It is shown [5] that the phase of the Brown IEM is identical to

$$sys_B(t) = w_B^T(t)\begin{pmatrix} x(t)-\hat{x}_B(t) \\ y(t)-\hat{y}_B(t) \\ d(t)-\hat{d}_B(t) \end{pmatrix} = \left(\sum_i \frac{1}{s_i}\right)^{-1} \sum_i \frac{1}{s_i} x_i \tag{23}$$

However, the Brown method weights corrected time, frequency and drift states as well:

$$w_B^T(t) = (w_{B,x}^T(t) \quad w_{B,y}^T(t) \quad w_{B,d}^T(t)) \tag{24}$$

In case of PPN, the frequency and drift weights are zero:

$$w_{B,y}^T(t) = w_{B,d}^T(t) = 0_N^T \tag{25}$$

By simulation the impact of the PPN case on the Greenhall and Brown algorithm is studied. Identical Q_i values of the clock process noises are used, but with a smaller system. The ensemble consists of seven clocks and two clocks of the ensemble observe the remaining five clocks at each measurement epoch. Different initialization values of the ensemble covariance are tested. The starting values are multiplies of the ensemble noise covariance

$$C(t_0) = kQ \tag{26}$$

with

$$k \in \{1e2, 1e4, 1e6, 1e8, 1e10, 1e12\} \tag{27}$$

Figures 6 and 7 show the time and frequency state weight of a observed clock 1: $w_{B,x,1}^T(t)$ and $w_{B,y,1}^T(t)$.

The time state weights fall on each other for k <=1e8. However, in case of k=1e10 and k=1e12 the weights converge to the same final value (Figure 6). Focusing on the frequency state weight (Figure 7), these values are close to zero for k<=1e8, however there are biases for k=1e10 and k=1e12. Recalling that the frequency weights are zero in the PPN case, we see large differences for k=1e10 and k=1e12. In contrast, the Greenhall time weight of observed clock 1 is not affected by the different initial covariances. Figure 8 shows the difference of an observed clock 1 between the time estimates of both algorithms. The time differences are numerically zero for k<=1e8, while the remaining two values (k=1e10 and k=1e12) result in biased differences.

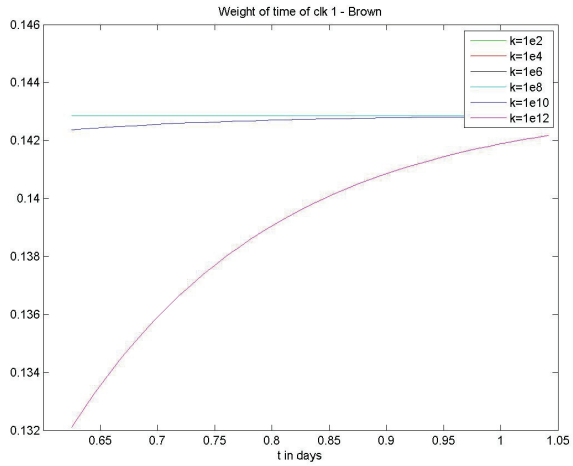


Figure 6. Weight of observed clock 1 time in Brown IEM phase, for 7-clock ensemble.

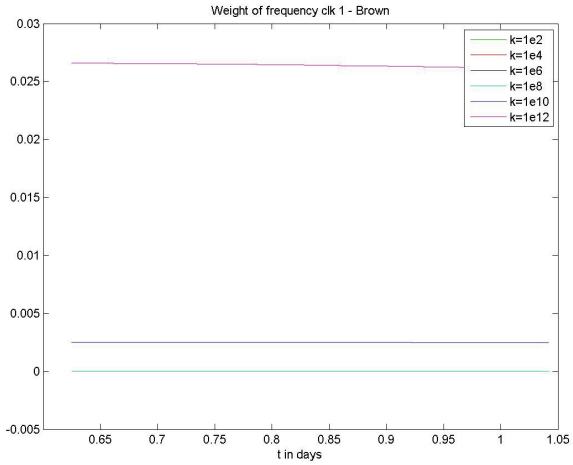


Figure 7. Weight of observed clock 1 frequency in Brown IEM phase, for 7-clock ensemble

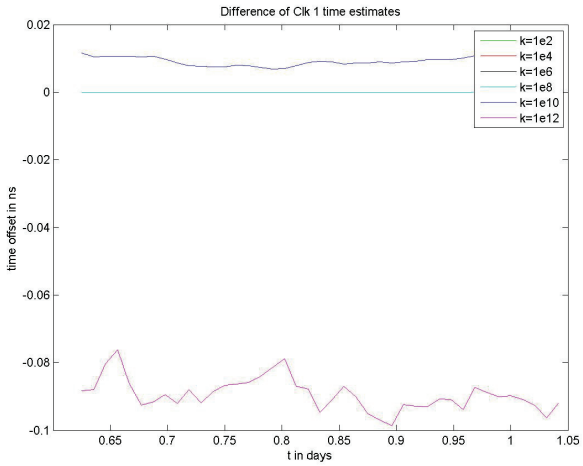


Figure 8. Differences between time estimates of Brown and Greenhall for observed clock 1 in 7-clock ensemble.

The results suggest that both algorithms generate identical estimates using proper initial values of the covariance (k<=1e8).

VI. Conclusion

The improvements due to improved station clocks are minimal if the "Greenhall" algorithm is used for the composite clock. Greenhall is generally superior, although the two algorithms respond differently to unanticipated clock variations. It is also noted that several other timescale algorithms have great merit, but have not been tested in this work. The study of the PPN case shows the sensitivity of the Brown algorithm to the initial values. It is suggested that the PPN case results in identical estimates of both algorithms. Simulations verify this behavior for a smaller ensemble using proper initial covariances.

Acknowledgment

We thank Benjamin Harris and the University of Texas for use of the GPS Toolkit. Dr. Greenhall contributed substantially to this work.

References

[1] M. Suess, D. Matsakis, and C. Greenhall, 2010, "Simulating Future GPS Clock Scenarios With Two Composite Clock Algorithms," 42nd Annual PTTI Meeting, Washington D.C., USA

[2] T. J. H. Craddock, R. J. Broderick, C. P. Petersen, and A. Hu, 2009, "The GPS Toolkit: Open Source Clock Tools." in Proceedings of the 40th Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 1-4 December 2008, Reston, Virginia, USA (U.S. Naval Observatory, Washington, D.C.), pp. 255-273.

[3] C. Zucca and P. Tavella, 2005, "The clock model and its relationship with the Allan and related variances," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, UFFC-52, 289-296.

[4] J. W. Chaffee, 1987, "Relating the Allan Variance to the Diffusion Coefficients of a Linear Stochastic Differential Equation Model for Precision Oscillators," IEEE Transactions on Ultrasonics, Ferroeletrics, and Frequency Control, UFFC-34, 655-658.

[5] K. R. Brown, 1991, "The theory of the GPS composite clock," in Proceedings of the ION GPS-91, 11-13 September 1991, Albuquerque, New Mexico, USA (Institute of Navigation, Alexandria, Va.), pp. 223-242.

[6] C. A. Greenhall, 2007, "A Kalman filter clock ensemble algorithm that admits measurement noise," Metrologia, 43, S311-S321.

[7] D. Manning, 2009, "AF/NGA GPS Monitor Station High-Performance Cesium Frequency Standard Stability 2007/2008: From NGA Kalman Filter Clock Estimates," in Proceedings of the 40th Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 1-4 December 2008, Reston, Virginia, USA (U.S. Naval Observatory, Washington, D.C.), pp. 335-348.

Supression of collisional frequency shifts in an optical lattice clock

(Invited Paper)

M. D. Swallows, M. Bishof, Y. Lin, S. Blatt, M. J. Martin, A. M. Rey and J. Ye
JILA, NIST, and the University of Colorado
Department of Physics, University of Colorado
Bouder, CO 80309-0440, USA
Email: ye@jila.colorado.edu

***Abstract*— By strongly confining atoms in a two-dimensional optical lattice, we have suppressed collisional frequency shifts in a ^{87}Sr optical lattice clock.**

I. Suppression of collisional frequency shifts

Optical clocks based on atoms confined in optical lattices provide a unique opportunity for precise study and measurement of quantum many-body systems. The state-of-the-art optical lattice clock has reached an overall fractional frequency uncertainty of 1×10^{-16} [1]. This uncertainty is dominated by two contributions – atomic collisions and blackbody radiation-induced frequency shifts. The density-dependent frequency shift arises from collisions between initially identical fermionic Sr atoms that are subject to slightly inhomogeneous optical excitations during clock operation [2]. We have recently implemented a seemingly paradoxical solution to the collision-shift problem: with a strong atomic confinement in one-dimensional tube-shaped optical traps, we dramatically increase the atomic interactions. Instead of a naively expected increase of collisional frequency shifts, these shifts are increasingly suppressed [3]. The large atomic interaction strength creates an effective energy gap in the system such that inhomogeneous excitations can no longer drive fermions into a pseudo-spin antisymmetric state, and hence their collisions and the corresponding frequency shifts are suppressed. We demonstrate the effectiveness of this approach by reducing the density-dependent frequency shift to the level of 10^{-17}, representing more than a factor of ten reduction from the previous record [2, 3]. In addition, we have observed well-resolved interaction sidebands separated from the main peak of the clock transition, giving direct evidence for the removal of the interaction energy from the clock carrier transition [4]. Control of atomic interactions at the level of 1×10^{-17} is a testimony to our understanding of a quantum many-body system and it removes an important obstacle for building a high accuracy optical atomic clock based on such systems.

Acknowledgment

The authors would like to thank C. Benko for assistance with the frequency shift measurements.

References

[1] A. D. Ludlow *et al.*, "Sr Lattice Clock at 1×10^{-16} Fractional Uncertainty by Remote Optical Evaluation with a Ca Clock," Science **319**, 1805-1808, (2008).
[2] G. K. Campbell *et al.*, "Probing Interactions Between Ultracold Fermions," Science **324**, 360-363 (2009).
[3] M. D. Swallows *et al.*, "Suppression of Collisional Shifts in a Strongly Interacting Lattice Clock," Science **331**, 1043-1046 (2011).
[4] M. Bishof *et al.*, "Resolved atomic interaction sidebands in an optical clock transition," arXiv:1102.1016v2 (to be published in Physical Review Letters) (2011).

978-1-61284-110-6/11/$26.00 ©2011 IEEE

Evidence of a fermionic collisional shift

Wilfried Maineult and Peter Rosenbusch
LNE-SYRTE, Observatoire de Paris, CNRS, UPMC, 61 av de l'observatoire, 75014 Paris, France
peter.rosenbusch@obspm.fr

Christian Deutsch,Jakob Reichel
Laboratoire Kastler Brossel, ENS, UPMC, CNRS, 24 rue Lhomond, 75005 Paris, France

Kurt Gibble
Department of Physics, The Pennsylvania State University, Pennsylvania 16802, USA

***Abstract*— We present here experimental result showing that an collisional shift that cannot be explained by mean-field effects. We demonstrate that the latter results from subtle interaction scheme in our cold atomic cloud, mathematically equivalent to a fermionic collisional shift. We show agreement between theory and experiment, and extract the relevant energy scales implied in the collision process.**

A new definition of the SI second will built on optical transitions. The interrogation of neutral fermionic atoms is performed to increase the clock stability while controlling the collisional shift. Theoretical efforts have been made in order to predict and model a fermion collisional shift [1] [2] [3]. For Ramsey interrogation, K. Gibble predicts a non-zero collisional shift due to an energy-dependent Rabi frequency of the driving pulse [2].

We apply his formalism to atoms in a harmonic trapping potential and extend the theory to the case of a position-dependent Rabi frequency. The case of an energy-dependent Rabi frequency can be viewed as the sideband-resolved limit of our calculation. We numerically solve the equations of motion for a complete Ramsey sequence. Just like in [2], we show that during the first Ramsey pulse, the inhomogeneous Rabi frequency coherently populates singlet and triplet states. Their populations beat at the frequency corresponding to the singlet state interaction energy. The time varying populations manifest themselves through the so-called spin waves [4]. The second Ramsey pulse reveals signature of multiple interfometers paths, whose beating is governed by this interaction energy alongside with the trap frequency and its multiples in the non-resolved sideband regime. This multiple path interference, direct consequence of the Rabi frequency inhomogeneity, results in a significant non-intuitive collisional shift when varying the second pulse area. Furthermore, the shift is highly non-linear at high atom density.

We verify our model by measurements of the $|F = 1, mF = -1\rangle$ to $|F = 2, mF = 1\rangle$ transition in magnetically trapped 87Rb atoms [5]. Because all scattering lengths coincide within 5%, this boson is mathematically equivalent to the fermion case. Spatial inhomogeneity of the Rabi frequency is obtained by using the evanescent field 150m above a microwave stripline. We indeed observe spin waves and a collisional shift that depends on the area of the second Ramsey pulse. We find excellent agreement between model and data. Thereby the data give evidence for the predicted fermion collision shift and suggest that it should be considered when designing future primary clocks.

References

[1] A. M. Rey, A. V. Gorshkov, and C. Rubbo, "Many-body treatment of the collisional frequency shift in fermionic atoms," *Phys. Rev. Lett.*, vol. 103, no. 26, p. 260402, Dec 2009.

[2] K. Gibble, "Decoherence and collisional frequency shifts of trapped bosons and fermions," *Phys. Rev. Lett.*, vol. 103, no. 11, p. 113202, Sep 2009.

[3] Z. Yu and C. J. Pethick, "Clock shifts of optical transitions in ultracold atomic gases," *Phys. Rev. Lett.*, vol. 104, no. 1, p. 010801, Jan 2010.

[4] J. M. McGuirk, H. J. Lewandowski, D. M. Harber, T. Nikuni, J. E. Williams, and E. A. Cornell, "Spatial resolution of spin waves in an ultracold gas," *Phys. Rev. Lett.*, vol. 89, no. 9, p. 090402, Aug 2002.

[5] C. Deutsch, F. Ramirez-Martinez, C. Lacroûte, F. Reinhard, T. Schneider, J. N. Fuchs, F. Piéchon, F. Laloë, J. Reichel, and P. Rosenbusch, "Spin self-rephasing and very long coherence times in a trapped atomic ensemble," *Phys. Rev. Lett.*, vol. 105, no. 2, p. 020401, Jul 2010.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

Ultrasonic Microparticle Trapping by Multi-Foci Fresnel Lens

Youngki Choe, Jonathan W. Kim, K. Kirk Shung[+], and Eun Sok Kim*
Department of Electrical Engineering-Electrophysics
University of Southern California
Los Angeles, CA 90089, USA
*Palos Verdes Peninsular High School
Rolling Hills Estate, CA 90275, USA
[+]Department of Biomedical Engineering
University of Southern California

***Abstract*—This paper describes an acoustic tweezers consisting of a multi-foci Fresnel lens on 127 μm thick PZT sheet, designed to capture micron-sized particles. The multi-foci Fresnel lens was designed to have similar working mechanism as that of an axicon lens to generate an acoustic Bessel beam and, correspondingly, to generate negative axial radiation force capable of trapping one or more microparticle(s). The fabricated acoustic tweezers successfully trapped lipid particles ranging in diameter from 50 to 200 μm and microspheres ranging in diameter from 70 to 90 μm at a distance of 2 to 5 mm from the tweezers without any contact between the transducer and microparticles.**

I. INTRODUCTION

Contactless particle trapping is an emerging research topic in biology, physical chemistry and bio-medical research. Among the various approaches of contactless tweezing techniques, optical tweezers is the most well developed and most commonly investigated technique [1]. However, particle trapping using optical manipulation has several limitations: first, the high energy of a highly focused laser can damage living cells. Secondly, optical tweezers cannot trap large particles due to the short wavelength of the light source. The size of the largest particle which can be trapped by an optical tweezers is limited to tens of micrometers or smaller. Finally, the use of optical tweezers is confined to optically transparent objects or shallow region of opaque medium.

Ultrasonic tweezers is a relatively young field, and seeks to provide an alternative to optical tweezers. Ultrasound is more attractive than light, because it imparts much higher energy (or impact force) and also offers greater spatial range, consequently providing the ability to capture larger particles over a wider range. However, ultrasonic tweezers is not well established or researched, possibly due to lack of good focusing technology. Control and manipulation of microparticles offer a wide range of applications ranging from the study of cells and molecules to the construction of labs on a chip, from cancer treatments to DNA studies.

Limited ultrasonic manipulation of microparticles has already been achieved. The systems used for the said manipulation, however, are not as versatile as are optical tweezers. One such system is the "acoustical tweezers" pioneered by Wu [2], a device capable of stably capturing latex particles of diameter 270 μm. The tweezers utilized two counteracting acoustic transducers to create a force potential well that could capture and manipulate different microparticles. Another device used for ultrasonic manipulation of microparticles used one highly focused transducer to hold a lipid droplet that was already confined to a mylar sheet in water [7].

This paper describes microparticle trapping in a 3-D space by a single ultrasonic transducer. We designed and fabricated transducers that could generate a Bessel beam containing a region of a negative axial radiation force [3]. In general, Bessel beams are generated by an axicon lens [4]. However, fabricating an axicon lens in small size can be challenging due to the steep lens angle required. Thus, we fabricated both an axicon lens and a Fresnel lens employing an air-reflector [5] that has wide tolerance to manufacturing imprecision. Both of the fabricated lenses were shown to generate a Bessel beam through particle trapping. The trapping by the Fresnel lens device was strong and stable, that there was no limit to moving range of the trapped particles in plane of the water surface.

II. THEORY

When a Fresnel lens is designed to have multiple focal points, the acoustic waves passing through the lens produces a Bessel beam with a micron sized region where the radiation force is negative, and microparticles can be trapped. The wave equation ψ_B for a scalar-wave Bessel beam is an axisymmetric solution of the free-space wave equation [3].

$$\psi_B(x,y,z) = \psi_0 \exp(i\kappa z) J_0(\mu\sqrt{x^2+y^2}) \quad (1)$$

where ψ_o, κ, J_o, and μ are the wave amplitude, axial wave number, zeroth-order Bessel function, and radial wave number, respectively.

Marston mathematically proved that negative axial radiation force exists in a Bessel beam under certain conditions [3]. To explain the condition for the negative axial radiation force, Marston defined parameter domain (k, a, β)

where k and a are the wave number and radius of trapping particle, respectively, while β is a cone angle. The cone angle β is an important parameter in the characterization of a Bessel beam, and linked to the parameters of Eq. (1) by

$$\beta = \arccos(\kappa/k) = \arcsin(\mu/k) \quad (2)$$

The square of the wave number k, i.e., k^2, is equal to $\kappa^2+\mu^2$, and also equal to $(\omega/c_0)^2$, where ω and c_0 are angular frequency and the phase velocity of the acoustic wave in the liquid, respectively.

III. Design and Fabrication

Observing the working mechanism of an axicon lens, we can see that the wave closer to the lens center is focused at a shorter distance from the lens. In the other end, the wave farther from the lens center is focused at longer distance from the lens as shown in Figure 2. A Bessel beam is formed in the region where the focused wave is uniformly distributed.

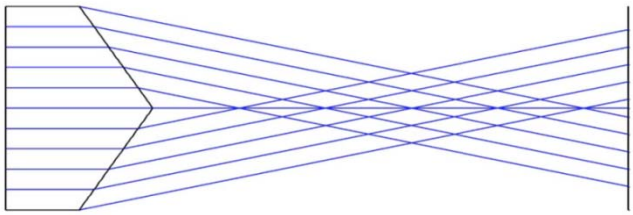


Figure 2. Ray trajectories of the acoustic waves going through an axicon lens.

To simulate the behavior of an axicon lens described above, a multi-foci Fresnel lens with an air-reflector was designed. The Fresnel lens has 7 rings in total. The 2 inner most rings, the next 2 rings, and remaining 3 rings have focal lengths of 830 µm, 860 µm, and 890 µm, respectively (Figure 4). As a control experiment, an acoustic tweezers with Fresnel lens having only one focal length of 800 µm and same number of rings was fabricated. When the control transducer was actuated, particles above the control transducer were immediately expelled from the region due to positive axial radiation force and the stirring effect, and no particle was trapped.

An acoustic tweezers with an aluminum alloy axicon lens also designed and fabricated to compare the performance. To have the cone angle β=60°, the axicon lens angle α was set to 63° (Figure 3)

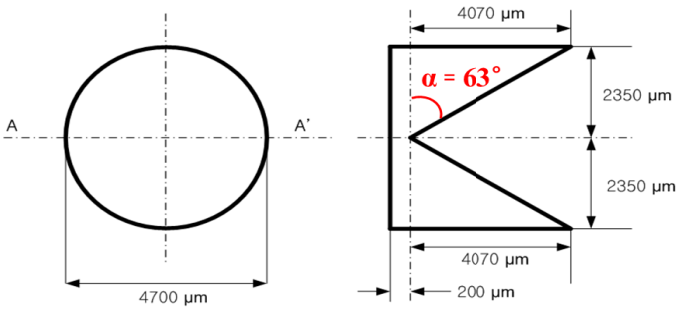


Figure 3. Top-view (*left*) and side-view (*right*) schematics of the Aluminum-Alloy Axicon Lens.

We fabricated the multi-foci Fresnel lens as follow. After delineating the top and bottom electrodes on a 127 µm-thick PZT sheet, we patterned Fresnel lens with photoresist as a sacrificial layer, on which 3.5 µm-thick parylene D layer was deposited as a lens structural layer. The sacrificial photoresist was then released through release holes to form air gaps, and additional 7 µm-thick parylene D layer was deposited to seal the release holes. Finally, a bulk micro-machined silicon was glued to the transducer as a structural support layer.

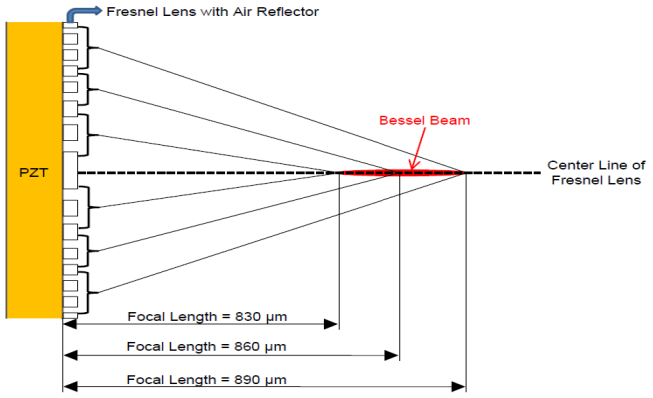


Figure 4. Side-view schematic of the Multi-Foci Fresnel Lens.

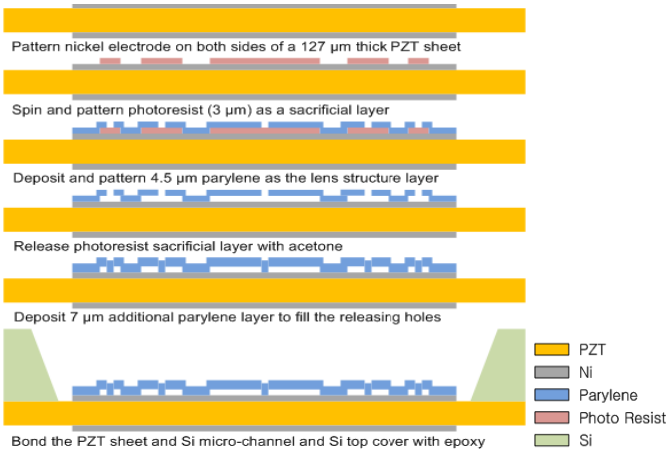


Figure 5. Brief fabrication steps of the acoustic tweezer built on a 127 µm-thick PZT sheet with Multi-Foci Fresnel Lens.

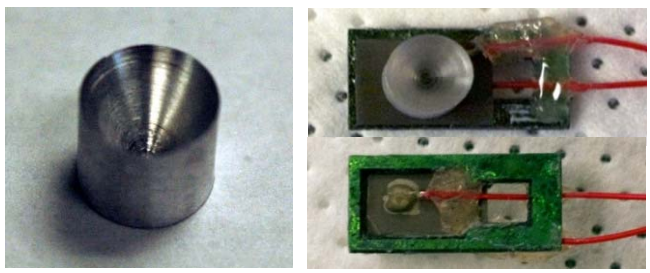


Figure 6. (Left) Photo of the fabricated aluminum-alloy axicon lens. (Right) Top-view and bottom-view photos of the acoustic tweezers built on the aluminum-alloy axicon lens.

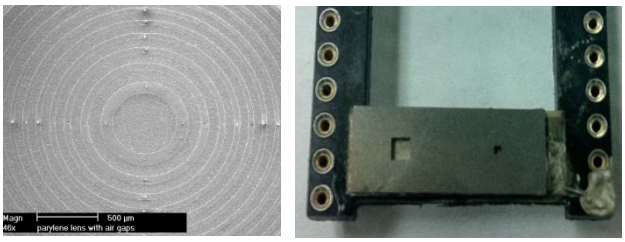


Figure 7. (Left) SEM photo of the fabricated multi-foci Fresnel lens on a PZT sheet. (Right) Photo of a packaged acoustic tweezers with the multi-foci Fresnel lens.

IV. Measurement Setup and Results

For a low energy in trapping microparticles, a pulsed 17.9 MHz sinusoidal signal, instead of continuous sinusoidal signal, was applied to the fabricated acoustic tweezers with 10 – 20 kHz pulse repetition frequency. The pulse width was 2 µsec, and the sinusoidal signal had 160 $V_{peak\text{-}to\text{-}peak}$ amplitude. The tweezers submerged in water was fixed to a manual XYZ moving stage, and lipid particles or microspheres were freely floating on and in the water. The movement of the microparticles was observed through a CCD attached a long range microscope, and the images and videos captured by the CCD were recorded with a computer (Figure 9).

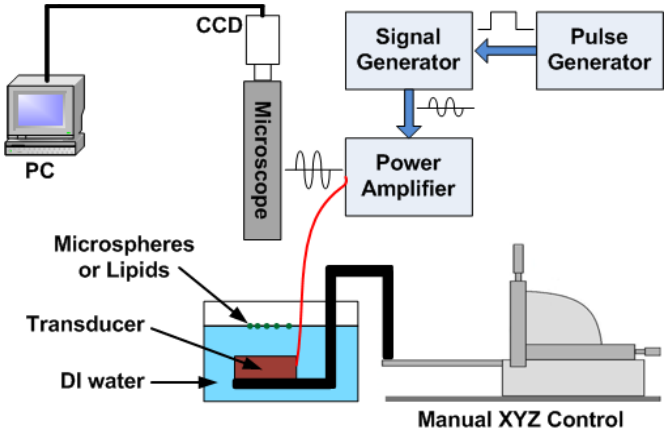


Figure 9. Schematic of the setup to observe the trapping of microparticles by the fabricated acoustic tweezers.

The fabricated tweezers, both the one based on the multi-foci Fresnel lens and the aluminum-alloy axicon lens were tested to trap lipid particles ranging from 50 – 200 µm in diameter and microspheres ranging from 70 – 90 µm in diameter. As the actuated tweezers produced acoustic waves and stirred the water as well as the microparticles in/on the water, the microparticles circled around the tweezers. Once a lipid particle hit the location where the Bessel beam was generated, the lipid particle firmly trapped to the spot and held there even when another lipid particle hit the trapped lipid particle (Figure 8). And we were able to move the trapped particle by moving the transducer over a wide range (limited only by the movable range of the transducer in the test apparatus).

Of the two devices tested, only the multi-foci Fresnel lens device demonstrated consistent success in capturing and manipulating both the lipid droplets (50 - 200 µm in diameter) and the polystyrene microspheres (70 - 90 µm in diameter) at a distance from of 2 to 5 mm from the lens surface without any contact between the transducer and microparticles.

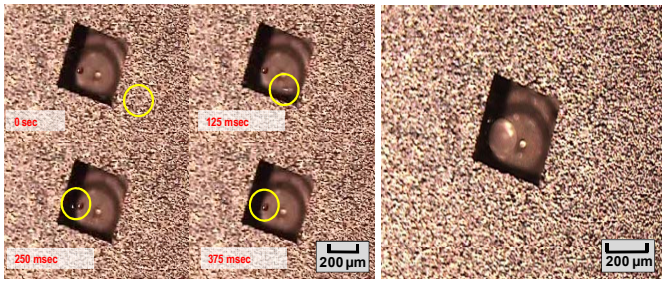


Figure 8. (Left) Photos taken at different times showing that a 72-µm-diameter lipid particle was being trapped by the multi-foci Fresnel transducer that was situated below the square opening. Another lipid particle (circled for clarification) that was being drifted by the stirring effect of the transducer hit the trapped lipid at 250 msec, but the trapped lipid was unmoved by the impact from the other lipid, as shown in the photo taken at 375 msec. (Right) Photo of a large 200-µm-diameter lipid particle being trapped by the same Fresnel transducer.

V. DISCUSSION

Axicon Lens Tweezers

The reason for the axicon lens transducer's ineffectiveness as a tweezers was the level of imprecision that was undeniably present in the making of the axicon lens. The machine shop which constructed the axicon lens was not able to build it to the high level of precision, and the performance of the axicon lens transducer was compromised. The interior surface of the lens had many imperfections and rough patches, which although very small, impacted the axicon lens transducer's performance. The axicon lens transducer was indeed able to capture one particle at the beginning of our tests, but quickly lost the particle and was not able to retain it for any significant period of time. With better fabrication at the machine shop, the axicon lens transducer should be able to produce consistent trapping and manipulation of microparticles. One major reason of going through this approach, and possibly continuing this approach, in spite of the success of the multi-foci Fresnel tweezers, is that the axicon lens is so much simpler in design and construction (than the multi-focus array) that the tests with it provides a very clear and solid understanding of principles behind the ultrasound-based particle trapping/cooling.

Multi-foci Fresnel Lens Tweezers

The multi-focus annular ring array was very successful in capturing particles, as it focused the waves very well over a multiple spots along the focal axis. This device merits further research, in an effort to possibly optimize the design and increase its capabilities in handling multiple particles on demand, moving the particles without involving the transducer movement, etc.

VI. CONCLUSION

A single ultrasonic transducer has been shown, for the first time, to capture microparticles (both lipid and polystyrene) in 3-D space. The microparticle trapping by a single focused-ultrasonic-beam transducer is very exciting, since the transducer (or an array of multiple such transducers) can pick up solid particles (ranging from a few microns to several hundreds in diameter), and move and place them at precise locations for diagnostics, construction, etc.

ACKNOWLEDGMENT

This paper is based on the work partially supported by National Institutes of Health under grant R21HG005118.

REFERENCES

[1] D. G. Grier, "A revolution in optical manipulation", Nature, vol 424, pp. 810-816, 2003

[2] J. Wu, "Acoustical Tweezers", Journal of Acoustical Society of America, vol. 89, issue 5, pp. 2140-2143, 1991.

[3] P. L. Marston, "Axial radiation force of a Bessel beam on a sphere and direction reversal of the force", Journal of Acoustical Society of America, vol. 120, issue 6, pp. 3518-3524. 2006.

[4] G. Milne, G. D. M. Deffries, and D. T. Chiu, "Tunable generation of Bessel beam with a fluidic axicon", Applied Physics Letters, Vol. 92, No. 26, pp. 261101-261101, 2008.

[5] C. Y. Lee, H. Yu, and E. S. Kim, "Acoustic Ejector with Novel Lens employing Air-Reflectors", Micro Electro Mechanical Systems, pp 170-173, 2006.

[6] J. Durnin, "Exact solutions for nondiffracting beams. I. The scalar theory", Journal of Optical Society of America, Vol. 4, No. 4, pp. 651-654, 1987.

[7] J. Lee, S. Y. Teh, A. Lee, H. H. Kim, C. Lee, and K. K. Shung, "Single beam acoustic trapping.", Applied Physics Letters, Vol. 95, No. 7, pp. 073701-073701, 2009

Ex Vivo Monitoring of Rat Heart Wall Motion Using Piezoelectric Cantilevers

Rui Zhang[1], Ya Chen[2], Wen H. Ko[1], David S. Rosenbaum[3,4], Xin Yu[2,3†], Philip X.-L. Feng[1†]
[1]Electrical Engineering, [2]Biomedical Engineering, Case School of Engineering
[3]Physiology and Biophysics, Case School of Medicine, [4]MetroHealth System
Case Western Reserve University, Cleveland, OH 44106, USA
[†]Corresponding Authors; Email: philip.feng@case.edu, xin.yu@case.edu

***Abstract* — We report on an experimental exploration of *ex vivo* measurements and *real-time* monitoring of the motions of heart wall for perfused rat hearts, by employing a surface-contact type of electromechanical probes based on piezoelectric transduction. In a hybrid experimental apparatus consisting of a conventional heart perfusion system and external electromechanical probing devices and circuitry, prototyped piezoelectric cantilever devices are calibrated and tested. We demonstrate that the external piezoelectric cantilevers are capable of monitoring the dynamic behavior of the isolated heart *ex vivo*, by measuring the motions of the heart wall. For typical rat hearts with heart rates in the range of ~150–250bpm (beats per minute), cantilevers with dimensions of $t \times w \times L \approx$ 130μm × (0.3–8)mm × (1–19)mm yield electrical signal of ~50–400mV. Measured data can also help identify signatures of various regimes (*e.g.*, from healthy to fatigued, to expiring) in the dynamical evolution during the perfused heart's lifetime. Preliminary tests on parallel multi-channel monitoring with probes positioned at multiple locations on heart surface prove to be valid and useful in obtaining information of regional displacement of heart wall.**

I. Introduction

Heart is the foundation of advanced lives including human being. Heart health and function monitoring are critical, especially for patients who suffer from heart diseases. Heart diseases are the top one cause of death in the United States and several other countries (*e.g.*, a total death of ~600,000 per year in 2007–2009 in US means one death due to heart disease every ~50 seconds) [1], and are the top reason for disease-based deaths throughout the world. This drives researchers to push the limits in advancing heart healthcare technology. We perceive at least the following major challenges: (i) *early detection* of alarms of heart diseases for apparently healthy people – it is desirable to develop devices that are wearable or implanted (with very low pain), especially for old people, athletes, and those who may have family history of heart diseases; (ii) *post-surgery* chronic heart monitoring for patients who are already receiving surgery and other treatment – what is desired includes low-pain implantable solutions that are small, light, enduring, and compatible with telemetry.

To date it has been well recognized that regional strain and stress on heart wall are related to development of disease [2], and studying the electromechanical properties and monitoring heart wall motion can help for heart diseases diagnosis [3]. From an engineering perspective, the heart is an amazing *electromechanical* device with exquisitely elegant functions and structures (Figure 1) enabled by soft materials and tissues – this causes some of the fundamental issues that are challenging the devices and instruments to be interfaced with heart for diagnosis and treatment. The prevalent electrocardiography (ECG) today is easy to use but only retrieves crude signals of overall heart function. Magnetic resonance imaging (MRI) is a powerful tool for studying heart structural and motional details and disease mechanisms. However today's MRI systems are bulky, highly complicated, and expensive, and often suffer from limitations in speed and resolution. We have been exploring a new, low-cost approach of directly probing heart wall motion by using distributed, surface-mount, and miniaturized (*e.g.*, micro and nanoscale) *electromechanical* devices for regional strain/stress and motion sensing. Here we report our initial effort toward this goal, and describe our first experimental results.

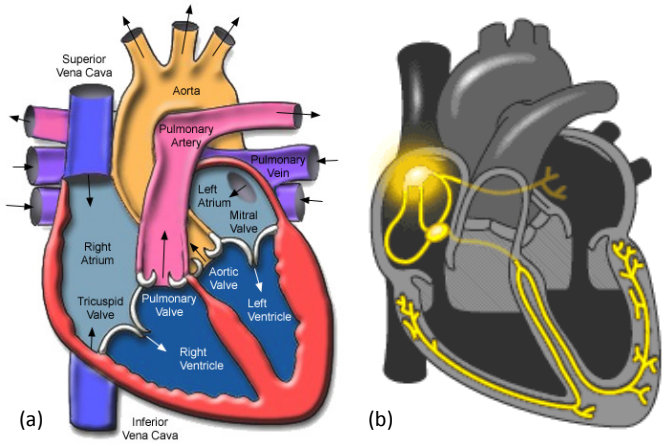


Figure 1. A glance of the heart structures and functions from an engineer's viewpoint. (a) Illustration of the anatomy of the heart, showing the heat chambers, vessels and valves. (b) Illustraton of the *pacemaker* and *distal conduction system*.

II. Heart Wall Motion

The heart is an electromechanical organ with great structural and functional complexities. Its wall motion is depending upon the compromised coronary arterial supply. Myocardial wall motion defects are essential and sensitive markers for coronary artery disease and myocardial ischemia. The capability of directly probing regional or highly localized heart wall motion may also have critical impact on arrhythmia,

restoring synchronization between asynchronous chambers or regions, early alarm of regional muscle fatigue or failure, *etc.*

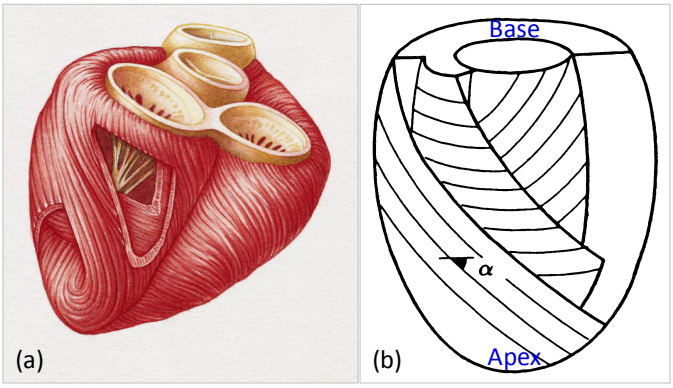


Figure 2. Simplified illustration of the complexity of the heart wall muscle structure and organization – the *basket-weaving* arthitecture of cardiac muscle cells and the myrocardial fibers. (a) A perspective view of the weaving fibers in layers at various depth in the heart wall. (b) A simplified model showing the weaving structure and orientation.

There are several types of muscle cells that participate in and coordinate the complex motions in concert. The most important of these include: (i) the cardiac muscle cells making the weaving myocardial fibers (see Figure 2a), (ii) the vascular smooth muscle cells (abundant in the coronary arterial tree), (iii) the conduction system muscle cells (*e.g.*, in the pacemaker region shown in Figure 1b). For this 'weaved basket', advanced MRI techniques including tagging, harmonic phase, and diffusion tensor MRI (DTMRI) have been developed for regional wall motion and strain assessment [4-7]. This work describes our initial effort and preliminary results toward a convenient electromechanical monitoring and diagnosis system using piezoelectric (PZE) devices, which, with computer aid, can automatically detect and monitor local heart wall motion *ex vivo*.

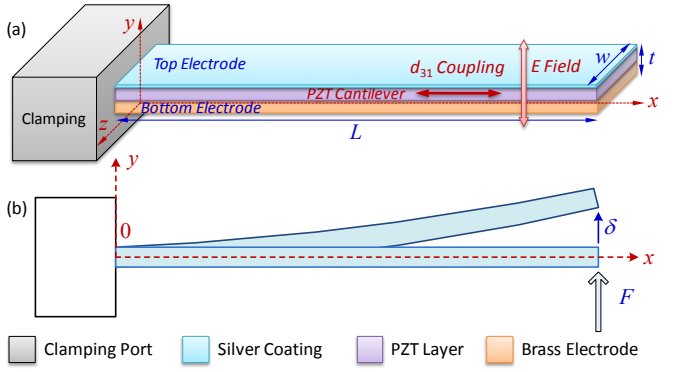


Figure 3. Piezoelectric (PZE) cantilever devices for electromechanical signal transduction. (a) Illustration of a prototypical composite cantilever with an active PZT layer sanwiched between two thin electrodes (metallic coating layers). (b) Simplified illustration of cantilever bending upon application of external force at the cantilever tip.

III. Piezoelectric Cantilever Sensors

A. Piezoelectric Cantilevers

We explore piezoelectric (PZE) device technology because of the direct electromechanical coupling effect in PZE transducers. A PZE device as simple as a singly-clamped cantilever beam (Figure 3) can be conveniently maneuvered to probe static deflections and dynamic motions of other mechanical systems. We have been exploring two scenarios for a PZE cantilever device to interface with a beating heart: (i) making contact between the free end of the cantilever and regions of interest on the heart wall, while keeping the other end of cantilever clamped on a solid (not moving) substrate; (ii) mounting the base (clamped end) of the cantilever on the beating heart's surface and having the cantilever body free to move and vibrate. The former is suited for *ex vivo* studies; the latter is attractive for packaged implanted systems.

Figure 3 illustrates a generic cantilever device based upon a ~65μm-thick PZE lead-zirconate-titanate (PZT) thin film sandwiched between two metal electrodes, a bottom ~65μm-thick brass layer, and a top ~1–5μm silver coating. We exploit the d_{31} coupling in such structures – transverse (out-of-plane) motion of the cantilever tip (free end) induces in-plane strain in the PZT layer and causes surface charge and electrical potential between the two electrodes which is read out for monitoring of motion. Within the scope of this work – cantilever tip in contact with heart wall with quasi-DC movements (beating frequency much lower than the fundamental resonance of the cantilever), the displacement, force, and the voltage signal are in convenient linear relationship, *i.e.*, $V_{PZE} \propto d_{31} \cdot \delta \propto d_{31} \cdot F/k_{eff}$, in a simple lumped parameter model [8].

TABLE I. Piezoelectric Materials of Interest

Material	**Piezoelectric Coefficients** [pm/V or pC/N]	**Young's Modulus** [GPa]	**Density** [g/cm^3]
PZT	d_{33}~100–600, d_{31}~50–300 d_{15}~100–800	~40–150	~7.5–7.8
ZnO	d_{33}~10, d_{31}~4, d_{15}~3	~30–140	~5.6
AlN	d_{33}~5.6, d_{31}~2.6, d_{15}~2.5	~330–410	~3.26
PVDF	d_{33}~20–30, d_{31}~20, d_{15}~10–40	~2–15	~1.76

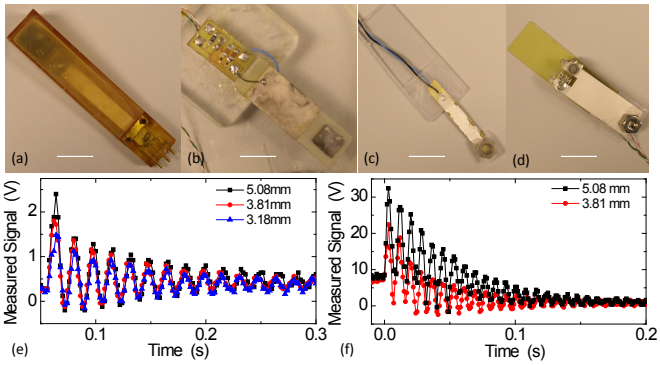


Figure 4. The first generation of our prototyped piezoelectric (PZE) devices based upon flexural-mode cantilevers using PZE thin film materials (*e.g.*, PZT and PVDF). (a) A vibration energy converter in plastic package with complementary dual PZT layers, easily generating ~10V level voltage signals from human body movements. (b) A PZT-based device with on-board energy conversion and harvesting circuit. (c) A much thinner PZT-based device. (d) A PZE cantilever based on flexible PVDF material. Scale bars: 1cm. (e) Measured time-domain voltage waveform (peak voltage ~2V) due to ring-down oscillations of a ~30Hz resonanst mode. (f) Voltage (peak voltage ~32V) measured from ~110Hz oscillations. Legend: initial tip deflection.

B. Materials of Choice

We choose PZT as the active material for its large PZE coefficient and easy availability for this study. Table I displays a short list of materials that we find interesting for employment in our studies. We have also been exploring devices made of polyvinylidene fluoride (PVDF) because of its attractive and promising properties for implants on flexible substrates. Other

materials of potential interest include zinc oxide (ZnO) and aluminum nitride (AlN), particularly for recent developments in engineering them into micro and nanoscale functional devices that could be integrated and packaged in implanted microsystems.

C. Prototyped Devices

We first demonstrate several prototyped PZE cantilevers using PZT and PVDF. The prototyped devices (see examples in Figure 4) are first calibrated in DC/static operation, and are then extensively tested in pulse and resonant modes. Early generations of cantilevers are on the few-mm- to 1cm-scale in size, with fundamental flexural resonance frequencies in the ~0.1kHz to ~1–100kHz ranges, and easily generates voltage signals up to ~10V. Our preliminary data and estimation show that these devices and their performance can be suited for both quasi-DC and resonant-mode applications, operating at atmospheric pressure, either in customized macroscopic plastic packages, or in micro polymeric thin film packages.

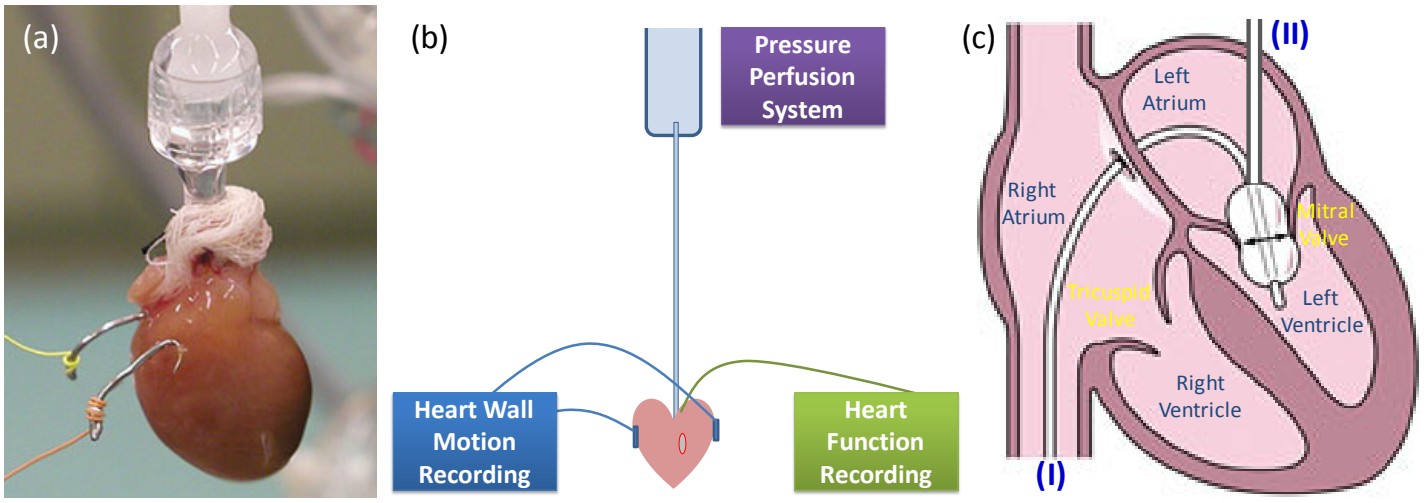


Figure 5. The heart perfusion system for *ex vivo* experimental studies with the isolated live rat heart. (a) Picture of a prototypical perfused rat heart with surface-mount external bulky electrodes for physiological measurements (*e.g.*, see http://vflab.org). (b) Highly simplified schematic of our experimental approach – using both the external piezoelectric (PZE) devices on heart wall, and the conventional approach for heart function recording system in a canonical rat heart perfusion system. (c) Illustration of the scheme of using a balloon (*i.e.*, simiar to *balloon valvuloplasty*) and its associated external perssure sensor for heart function recording. (I) and (II) are two specific options for implanting the balloon; we use option (II) in all the tests presented in this work.

IV. Experimental Techniques

In this early-stage effort of our exploration, as illustrated in Figure 5, we combine the piezoelectric (PZE) cantilever monitoring technique with well-established (commercially available) heart function recording systems. This helps to reliably evaluate the new approach and calibrate the measurements and the devices, both qualitatively and quantitatively, against today's standard protocols.

A. Heart Perfusion System and Heart Function Recording

The conventional *real-time* heart function recording is realized in a heart perfusion system, as illustrated and shown in Figure 5 and Figure 6. The system is based on the classical Langendorff technique for isolated heart perfusion [9]. This allows for convenient and prompt *ex vivo* studies, and many 'brute-force' (*e.g.*, Figure 5a) experiments, on isolated hearts. In the particular case of this study, the perfusion system greatly facilitates the continuous tests with the piezoelectric cantilever probes in a considerably long time (depending on the heart lifetime in the perfusion system, ~1-3 hours typically). It provides not only the live heart, but also a parallel monitoring option as a control experiment.

Male Sprague-Dawley rats of 10–12 weeks old are heparinized (1000 units/kg, i.p.) and anesthetized by sodium pentobarbital (85 mg/kg. i.p.). The heart is excised, cannulated, and perfused with Krebs-Henseleit (KH) buffer containing (in mM) 118.5 NaCl, 4.7 KCl, 1.2 $MgSO_4$, 1.2 KH_2PO_4, 1.5 $CaCl_2$, 11.1 glucose, and 25 $NaHCO_3$. Isolated hearts are perfused at a constant pressure in the Langendorff-type perfusion system. The perfusate was maintained at 37°C and equilibrated with 95% O_2-5% CO_2.

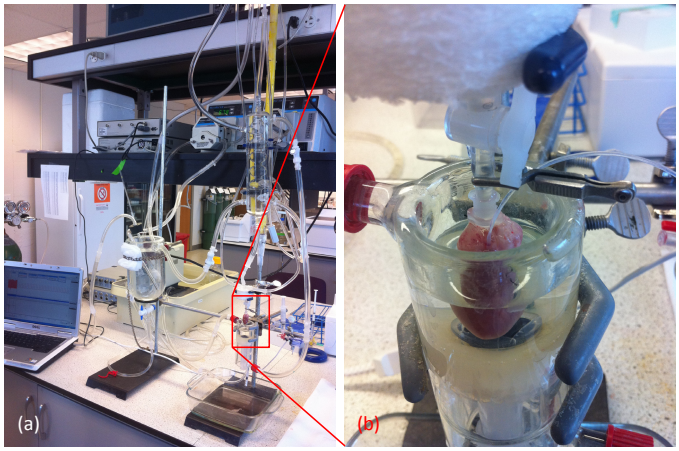


Figure 6. Pictures of the rat heart perfusion system employed in this work. (a) The overall Langendorff perfusion system with rat heart, pump, circulation lines, the implanted balloon sensor and its data acquisition system (vendor: ADInstruments). (b) Close-in view of the rat heart in the perfusion buffer.

Embedded in the heart perfusion system, as shown in Figure 5c, we insert a water-filled latex balloon into the left ventricle (*i.e.*, similar to *balloon valvuloplasty*). The balloon is connected to an external pressure transducer to record the left ventricular developed pressure (LVDP) and heart rate (HR). The heart's rate-pressure-product (RPP), *i.e.*, the product of LVDP and HR, is then calculated as an index of the workload.

The measured rat heart rate can typically be in the range of ~150–250 bpm (beats per minute) at the beginning of the experiment (with a fresh and healthily perfused rat heart). The heart rate decreases gradually as the time elapses. The temperature of the perfusion buffer is a critical factor for keeping the heart alive. When temperature deviates, the heart function could degrade, and the amplitude of heart wall movement may decrease dramatically. Figure 6 demonstrates pictures and details of the rat heart perfusion system we have been implementing in this study.

B. *Piezoelectric Cantilever Monitoring System*

We have built a convenient desktop apparatus for interfacing the piezoelectric cantilever probes with the rat heart in the operating perfusion system. Cantilever probes can be positioned and adjusted by moving the arms (to which the cantilevers are clamped) on the stage with control in all three directions. Figure 7 demonstrates a picture of an early generation of the implementation.

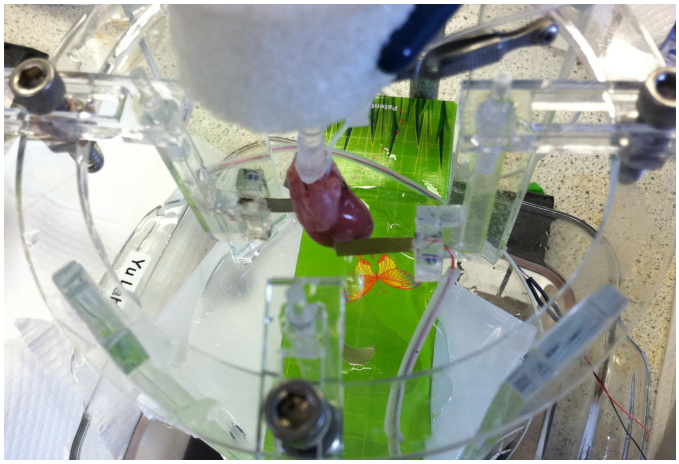


Figure 7. Picture of first-generation exprimental implementation of *ex vivo* monitoring of rat heart wall motion, by using surface-contacting piezoelectric (PZE) cantilever probes. For each cantilever, its one end is clamped, and the other end (tip) is in contact with the heart wall muscle. In paralle, the perfused rat heart's basic function is also being monitored using the balloon valvuloplasty technique. In this particular picture, the free ends of a pair of cantilevers are gently contacting the left and right ventricle of the heart.

TABLE II. PARAMETERS OF SELECTED TESTED DEVICES

Cantilever Device ID	**Length** [mm]	**Width** [mm]	**Resonance** [Hz]
(I)-A	18.0	7.0	222.0
(I)-B	19.0	7.3	214.0
(II)-A1	1.5	0.3	–
(II)-A2	1.3	0.4	–
(II)-A3	1.4	0.3	–
(II)-B1	1.6	0.4	–
(II)-B2	1.5	0.3	–
(II)-B3	1.4	0.4	–

As we expect these cantilever devices to operate while interfacing with perfused hearts in physiological solutions, we need to package the devices so that their electromechanical performance would not be compromised by any corrosion or contaminants. Prior to testing, every device is coated by a thin layer of parylene C (thickness on the order of ~5μm).

We note that in Figure 7 the liquid solution for nurturing the heart is temporarily moved away. In the present generation of setup, both the heart and all the cantilevers and device arrays (all micropackaged with ~5μm parylene C thin layer) are immersed in the fluid. Basic parameters of two generations of devices tested in this work are summarized in Table II.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

Extensive experimental observations and measurements have been performed. First, the heart function is recorded using the balloon implanted in the left ventricle, which provides a reference and calibration for the performance of the perfused rat heart. Then the piezoelectric cantilevers are applied to make contact to the heart wall for direct electromechanical probing.

A. *Heart Function Recorded by Balloon in Left Ventricle*

Without engaging any piezoelectric devices, real-time heat function is recorded by the left ventricular balloon and its associated pressure sensor. Typical data traces of LVDP and left ventricular pressure changing rate *dp*/*dt* are recorded for ~1–2 hours or even longer, throughout the whole lifetime of the perfused heart. Figure 8 shows the measured data in a very short 12s time interval. Table III summarizes the results from a few repeated measurement runs by only using the implanted balloon for monitoring and recording.

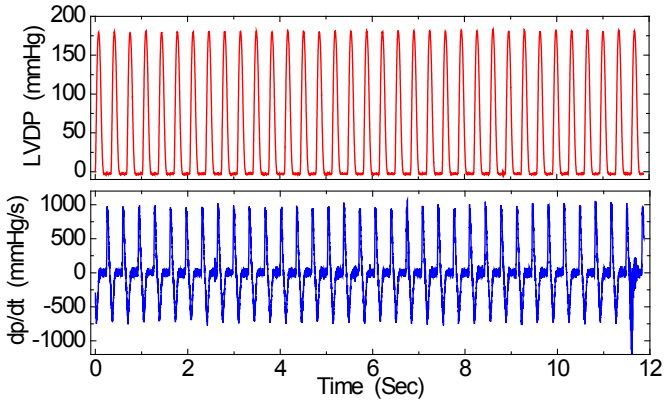


Figure 8. Representative data of the perfused rat heart under healthy condition, measured in real time by only the balloon implanted in left ventricle (no any external cantilever probes touching the heart wall). (a) Left ventricular developed pressure (LVDP, in mmHg) as a fundtion of time. (b) Measured pressure changing rate *dp*/*dt* as a function of time.

TABLE III. MEASURED HEART PARAMETERS WITHOUT CANTILEVERS

Test Run ID	**Heart Rate (HR)** [beats per minute, bpm]	**LVDP** [mmHg]	**RPP** [mmHg·bpm]
1	153 ± 3	135 ± 2	20655 ± 710
2	150 ± 3	120 ± 2	18000 ± 666
3	184 ± 5	102 ± 2	18765 ± 890

B. *Measurements with Piezoelectric (PZE) Cantilevers*

Prior to using the cantilevers for perfused heart wall motion probing, we first perform a 'dry-run' test by using air-filled balloons to mimic simplified heart motions. Because the fundamental flexural-mode resonance frequencies of the PZT cantilevers in this work are usually in the range of ~200Hz to ~100kHz range, the rat heart motions are well in the close-to-DC or quasi-DC range. Because in all our tests, the cantilevers are in contact with the air balloon (*i.e.*, heart model) or perfused heart at the cantilevers' tips, the tip motions closely follow the contractions and heart beat cycles. The heart motion does not drive the cantilever into its flexural-mode resonance.

Such quasi-DC operations of the cantilever probes lead to a simple linear relation between the piezoelectric voltage output and the detected displacement of heart wall motion. The 'dry-run' test (using air balloon to simulate a beating heart) indicates that for the early generations of devices (relatively large), *e.g.*, both (I)-A and (I)-B (listed in Table II), have high conversion *responsivity* (*i.e.*, gain), ~3V/mm (voltage per unit cantilever deflection at its tip). In both the 'dry-run' with heart models and the perfused heart tests, the output voltage signals are directly recorded and monitored by the oscilloscope using Labview/DAQ software.

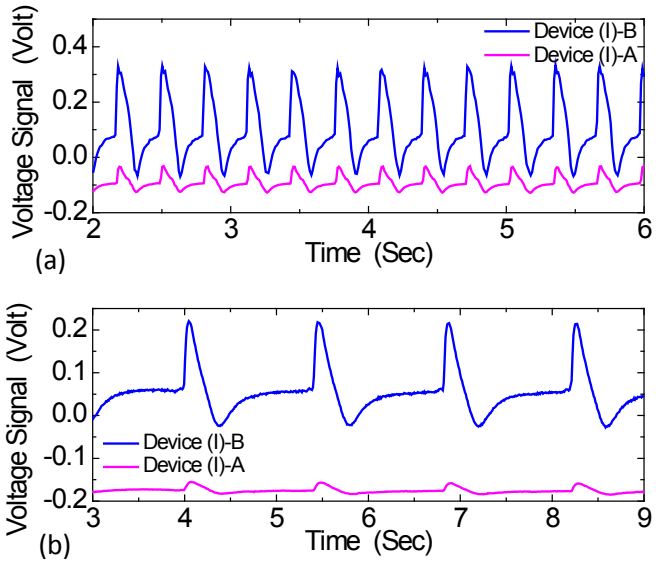


Figure 9. Voltage signal measured from a pair of PZT cantilevers touching the heart wall. (a) The free ends (tips) of the two cantilevers touching the heart wall of the left and right ventricle, respectively. (b) The free ends of the cantilevers are touching the base and apex of the heart wall, respectively.

Figure 9 demonstrates the voltage signals representing the local heart wall motions, probed by a pair of PZT cantilevers. In all these cases, the free ends of the cantilevers just barely touch the heart wall. When the cantilevers tips make contact to the left and right ventricles (Figure 9a) respectively, the extracted heart rate is HR~229bpm. Measured voltage signal is V_{pp}~150mV, and ~350mV, for device (I)-A and (I)-B respectively. The asymmetry in the data amplitude from this pair of devices is due mainly to the fact that one cantilever device has been pre-bent (with a transverse crack developed but not yet broken, at ~1/3 length near the clamped end), and thus has much less strain developed given the same displacement at the tip. When the devices' tips are placed against the base and the apex at a later time, the data show a lower HR~43bpm, and lower V_{pp}~50mV and 150mV, respectively, mostly due to the heart degradation during the time of transition.

The effects of probing the heart wall motion by making contacts with different depths and strengths are also explored. A second generation (group (II) listed in Table II), smaller cantilevers, in small parallel fingers-alike arrays, are used. With control of the positioning arms, the devices are first carefully placed to just barely touch the heart wall, and then gradually move toward the heart center to start gently pressing the heart wall. Figure 10 displays the measured data from a pair of small cantilevers, (II)-B1 and (II)-B2, in contact with the heart wall under three different depth/strength conditions.

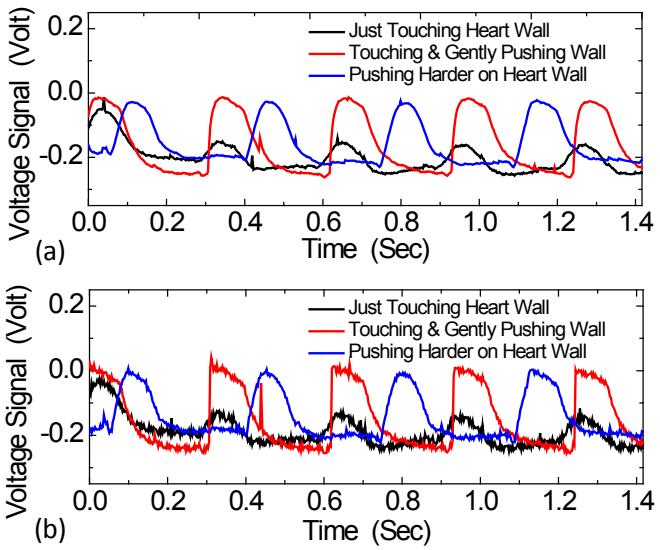


Figure 10. Voltage signal measured from a pair of smaller cantilever devices with varying the depth of the cantilever tip touching the heart wall. (a) Data from device (II)-B1. (b) Data from device (II)-B2. The data shown are truncated from much longer traces. For each device, data for three contact conditions are taken at three time intervals in series. Throughout these time intervals, the implanted balloon recording ensures that the heart functions are normal and stable. The offsets on the time axes are not adjusted among the different traces.

In case of just touching but not pressing the heart wall, larger devices (group (I) in Table II) yield larger V_{pp}, which is evident from Figure 9 and Figure 10. As cantilever tips get to press against the heart wall gradually, the signals first increase as expected, and then there is no more appreciable signal increase observed with further pressing of the devices against the heart surface. Such information can help us better understand the strength of the heart wall motion, especially with more advanced future generations of devices. We also note that the data traces from larger devices (gently touching heart wall, Figure 9) appear to be proportional to the *dp*/*dt* data in Figure 8, while the voltage signals from the smaller devices (under all contact conditions, see Figure 10) seem to have the shape similar to that of the LVDP curve in Figure 8. We are currently making more effort to investigate this intriguing phenomenon.

As we aggressively miniaturize these devices by using micromachining techniques, their apparent signal levels decline. Nonetheless, with significant volume reduction they become better suited with flexible substrates [10,11] and packages that are more amenable to 'harsh' environments for implants in living bodies. We envision that it is also possible for us to take advantage of the resonant operations of the PZE devices, combined with air-cavity packages possible in flexible substrate, cantilever- and membrane-structured micro/nano resonators in various frequency ranges can be exploited for local heart wall motion monitoring. Further, the same types of PZE devices in micropackages can also be employed for energy conversion from heart beats [12], which could be exploited for self powering low-power implants. Moreover, for miniaturized devices and chip-scale implants, the signal transmission could be not only wired but also wireless. As all these technical components are getting ready, the approach of using miniaturized PZE devices explored in this work can lead to both external heart function monitoring systems in research and clinic labs (*e.g.*, supplementing the

conventional implanted balloon recording), and implanted systems for surgery and patients' chronic heart monitoring – and in all these applications, capable of offering high sensitivity and high spatial and temporal resolutions for parallel readout of heart wall motion at various locations.

VI. Concluding Remarks

In summary, we have shown that using external surface-contact or surface-attached piezoelectric (PZE) cantilevers can probe the rat heart wall motion and the heart function *ex vivo*. The cantilevers in the present work are made of PZT thin films with strong PZE effect. The d_{31} coupling effect in the cantilever is exploited to transduce the flexural mechanical motion of the cantilever into electrical signal for readout. The cantilevers have been limited to quasi-DC operation (heart beating rate much lower than cantilever's resonance frequency) with their tips closely following the movements of the regional heart wall. Devices with sizes in the mm-scale and sub-mm-scale are tested. The preliminary tests verify the feasibility of monitoring heart wall movements with good resolutions in the time domain, and at different locations on heart surface. The heart wall displacement extracted from the measurement is typically ~0.1–0.3mm, which is consistent with MRI measurement. The output power level of typical PZE cantilevers we have tested is in the range of ~1–10μW. Combined with advances in PZE materials at micro and nanoscale, implantable and flexible materials, and micropackaging techniques, this approach is expected to have the potential of being implanted, as well as offering very high spatial and temporal resolutions by employing further miniaturized devices.

Acknowledgment

We thank the Louis Strokes Cleveland Medical Center of the Department of Veterans Affairs and the Case School of Engineering for financial support. We are indebted to C. A. Zorman, M. A. Rogonjic, and K. N. Kortepeter for their administrative support. We are grateful to the IEEE UFFC – IFCS/EFTF 2011 for the Student Travel Support Award (for R.Z.). We thank R. C. Roberts and S. B. Lachhman for help on materials and instruments.

References

[1] K. D. Kochanek, J. Q. Xu, S. L. Murphy, A. M. Minino, H.-C. Kung, "Deaths: preliminary data for 2009", *National Vital Statistics Reports*, vol. 59, no. 4, pp. 1-68 (2011).

[2] D. Jeyaraj, L. D. Wilson, J. Zhong, C. Flask, J. E. Saffitz, I. Deschenes, X. Yu, D. S. Rosenbaum, "Mechanoelectrical feedback as novel mechanisms of cardiac electrical remodeling", *Circulation*, vol. 115, pp. 3145-3155 (2007).

[3] E. Nagel, H. B. Lehmkuhl, W. Bocksch, C. Klein, U. Vogel, E. Frantz, A. Ellmer, S. Dreysse, E. Fleck, "Noninvasive diagnosis of ischemia-induced wall motion abnormalities with the use of high-dose dobutamine stress MRI: comparison with dobutamine stress echocardiography", *Circulation*, vol. 99, pp. 763-770 (1999).

[4] W. Liu, M. W. Ashford, J. J. Chen, M. P. Watkins, T. A. Williams, S. A. Wickline, X. Yu, "MR tagging demonstrates quantitative differences in regional ventricular wall motion in mice, rats, and men", *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 291, pp. H2515-H2521 (2006).

[5] J. Zhong, W. Liu, X. Yu, "Characterization of three-dimensional myocardial deformation in the mouse heart: an MR tagging study", *Journal of Magnetic Resonance Imaging*, vol. 27, pp. 1263-1270 (2008).

[6] J. Garot, D. A. Bluemke, N. F. Osman, C. E. Rochitte, E. R. McVeigh, E. A. Zerhouni, J. L. Prince, J. A. C. Lima, "Fast determination of regional myocardial strain fields from tagged cardiac images using harmonic phase MRI", *Circulation*, vol. 101, pp. 981-988 (2000).

[7] J. J. Chen, S.-K. Song, W. Liu, M. McLean, J. S. Allen, J. Tan, S. A. Wickline, X. Yu, "Remodeling of cardiac fiber structure after infarction in rats quantified with diffusion tensor MRI", *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 285, pp. H946-H954 (2003).

[8] S. P. Timoshenko, J. M. Gere, *Mechanics of Materials*, New York: Van Nostrand Reinhold Co. (1972).

[9] M. Skrzypiec-Spring, B. Grotthus, A. Szelag, R. Schulz, "Isolated heart perfusion according to Langendorff – still viable in the new millennium", *Journal of Pharmacological and Toxicological Methods*, vol. 55, pp. 113-126 (2007).

[10] Y. Qi, N. T. Jafferis, K. Lyons, Jr., C. M. Lee, H. Ahmad, M. C. McAlpine, "Piezoelectric ribbons printed onto rubber for flexible energy conversion", *Nano Letters*, vol. 10, pp. 524-528 (2010).

[11] X. Chen, S. Y. Xu, N. Yao, Y. Shi, "1.6V nanogenerator for mechanical energy harvesting using PZT nanofibers", *Nano Letters*, vol. 10, pp. 2133-2137 (2010).

[12] W. H. Ko, "Piezoelectric energy converter for electronic implants", *U.S. Patent*, No. 3,456,134 (1969).

Influence of non-ideal clamping in microcantilever resonant frequency estimation

L. Fadel-Taris, C. Ayela, I. Dufour
Université de Bordeaux, CNRS, IMS Laboratory
Talence, France
cedric.ayela@ims-bordeaux.fr

F. Josse, S.M. Heinrich
Marquette University
Milwaukee, USA
stephen.heinrich@marquette.edu

D.Saya
LAAS-CNRS
Toulouse, France

O. Brand
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, USA

Among the most promising sensing platforms are resonating microcantilevers due to their high sensitivity and wide application range. A key parameter of the device implementation is the predicted value of the resonant frequency that depends on the modeling and considerations of relevant physical phenomena. In fact, the estimation based on the conventional, perfectly clamped, Bernoulli-Euler cantilever beam does not lead to satisfactory accuracy in certain cases. Hence this work investigates two system characteristics that may affect the resonant frequency (a support effect and a so-called rim effect) and provides solutions for a straightforward estimation of rim dimension using resonance behavior of the cantilevers.

I. INTRODUCTION

Among the most promising sensing platforms are resonating microcantilevers due to their high sensitivity and wide application range for chemical sensing. A key parameter of the device implementation is the predicted value of the resonant frequency, which depends on the modeling and considerations of relevant physical phenomena. The classical expression used for the first transverse resonant frequency of a clamped-free microcantilever is obtained from the solution of the Euler-Bernoulli equation for a cantilever in vacuum [1]:

$$f_0 = \frac{\lambda_0^2 h}{2\pi L^2}\sqrt{\frac{E}{12\rho}} \qquad (1)$$

With $\lambda_0 = 1.875$, h the cantilever thickness, L the cantilever length, E and ρ the Young's modulus and mass density of the cantilever material, respectively. In fact, when compared to measurement, this equation does not lead to a satisfactorily accurate result in certain cases. For example, in the case of a silicon cantilever 504µm x 100µm x 20µm, with the cantilever length parallel to the <110> direction of the silicon <100> wafer, the measured resonant frequency is 94.4kHz, whereas the theoretical resonant frequency using (1) with the mechanical properties of silicon in this configuration (E = 169GPa, ρ = 2330kg/m^3) leads to 108.3kHz (15% higher than the measured value). Assuming that the geometry of the cantilever and the boundary conditions are not the error sources, two possible reasons for the difference between the theoretical and measured resonant frequency may be envisioned: the presence of the air as the surrounding medium instead of vacuum and the shear strain which is not considered in the Euler-Bernoulli equation.

The presence of the surrounding medium can be taken into account by considering the hydrodynamic force per unit length exerted by the air on the cantilever. This force per unit length is composed of both a viscous part proportional to the cantilever velocity (term noted g_1) and an inertial part proportional to the cantilever acceleration (term noted g_2) [2]. Equation 1 has to be modified by taking into account this force, resulting in [3]:

$$f_r = f_0 \frac{\sqrt{1-\frac{1}{2Q^2}}}{\sqrt{1+\frac{g_2}{m_L}}}; \qquad (2)$$

with m_L the cantilever mass per unit length and Q the quality factor defined by [3]:

$$Q = \frac{2\pi\sqrt{1+g_2/m_L}}{g_1/m_L} f_0 \qquad (3)$$

Using these equations and appropriate values for g_1 and g_2 [3], the resonant frequency in air (107.1kHz) is smaller than in vacuum but the difference is not sufficient to predict the measured resonant frequency.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

The Timoshenko beam theory [4] allows taking into account the shear strain and rotational inertia effects which are not considered in Eq.1. Again, these effects are not significant enough to account for the decrease in the resonant frequency that is seen in the measurements.

This leads one to consider that frequency discrepancy may be due to variations in the geometry and/or the boundary conditions. As detailed in the following section, the fabrication process may result in cantilever and support geometry that is not exactly the same as the one considered in the solution of the Euler-Bernoulli equation for a perfectly clamped cantilever. In this context, the work presented in this paper investigates two system characteristics that may affect the resonance mode and associated resonant frequency: the elasticity of the support which doesn't clamp ideally the supported end of the cantilever and the effect of a fabrication-induced silicon undercut that can occur during the release process of the suspended structures by Deep Reactive Ion Etching (DRIE).

II. MOTIVATION

Classically, silicon microcantilevers are fabricated by means of silicon standard micromachining techniques using a process in which release of the structures is based on a backside etching of the wafer. For this, Silicon-On-Insulator (SOI) substrates are commonly used in which a thin silicon dioxide layer insulates the backside silicon etching, either using a wet or dry etching process. This results in free-standing microcantilevers clamped on a thick silicon support. An example of silicon microcantilevers with silicon support is shown on figure 1. From this figure, a free-standing microcantilever is clearly visible, but a silicon undercut at the clamped end of the cantilever is also visible, due to silicon over-etching during backside etching. (In the present case, silicon backside etching is provided by means of DRIE.) A consequence is a loss of rigidity at the supported end of the cantilever, resulting in possible mechanical deformation of the undercut part, which in this paper will be referred to as the "rim" (Fig. 1).

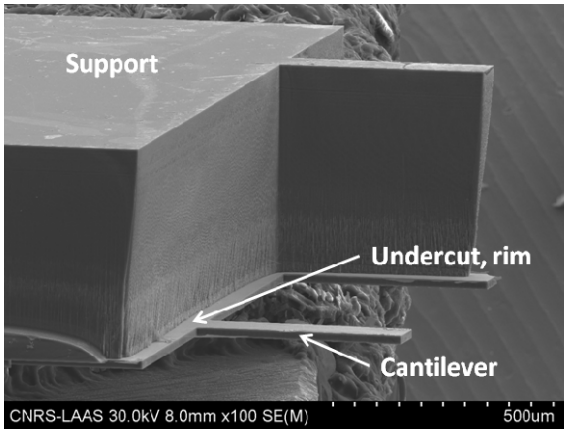


Figure 1: SEM picture of a microcantilever released by DRIE on a SOI wafer. Visualization of the silicon support and over-etching (rim).

In the present analysis, two origins in the discrepancy between theoretical and experimental values of resonant frequency of microcantilevers have been identified: a support compliance effect and an effect due to silicon over-etching, i.e., the so-called rim effect. First, the impact of both effects on the resonant frequency has been simulated using COMSOL. For this, a cantilever with ideal clamping (no rim, rigid support) has been simulated and the results compared to resonant frequency predictions for the case in which the structure is supported by an elastic rim attached to a rigid support block (rim length = 35µm in the simulation). Then, the impact of both effects (combined) has been simulated. The comparison of results is indicated in the schematic diagram of figure 2. This figure shows that the combined effect of rim and support is significant: the simulated resonant frequency is 11.3% lower than for the ideally-clamped cantilever. The results also show that the rim effect dominates the support compliance effect since a decrease of resonant frequency of 10% is due to the rim and thus, 1.3% due to the support elasticity. In this context, the systematic and direct (i.e. without the need of a backside observation of the chip) estimation of the rim length is of particular interest. This will lead to an appropriate use of the structures in the broad range of applications covered by resonant microcantilevers. The objective for the present work is thus to develop a method for using the resonance behavior of microcantilevers to determine the rim length at the supported end of the structures. Future work will include the development of an analytical expression for the cantilever resonant frequency which takes into account the rim effect.

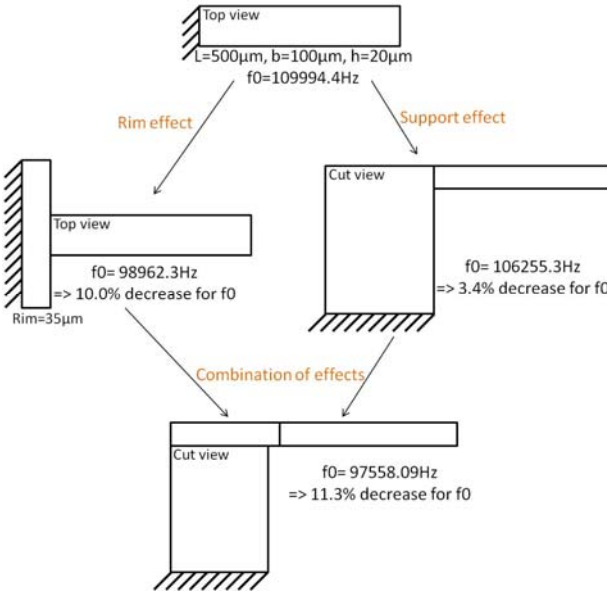


Figure 2: Microcantilever geometries and resonant frequencies obtained by COMSOL simulation in the case of a) ideal clamping, b) rim effect, c) support effect, d) combination of effects.

III. Experimental Section

A. Cantilever fabrication

Free-standing microcantilevers have been fabricated by standard micromachining techniques. Cantilevers are rectangular shaped with a length of 500µm and a width of 100µm. Each cantilever can be actuated individually by means of a Laplace force by incorporating both a patterned gold layer allowing a local current flow and an external magnet.

The main steps of the fabrication process are as follows. The starting substrate is a 100 mm-diameter, <100>, N-type Silicon-On-Insulator (SOI) wafer, with a 1 µm-thick buried oxide and a 20 µm-thick top silicon layer (resistivity of 4-6 Ω.cm). A first step consisted in the deposition of 300 nm of Plasma Enhanced Chemical Vapor Deposition (PECVD) silicon dioxide on the entire SOI wafer before the sputtering of Ti/Au (100 nm / 700 nm) for the electrode used for electromagnetic actuation. The film was lifted off with an AZ nLOF negative photoresist to define the electrode characterized by a width of 10 µm. A passivation silicon oxide film (300 nm thick) was then deposited by PECVD. Contact pads were opened by wet etching of oxide using HF buffer. To finish, the microcantilever shapes were defined by a front Reactive Ion Etching of silicon, followed by vertical sidewalls etching on the backside of the SOI wafer using the Deep Reactive Ion Etching technique to release the structures. The 1 µm-thick SiO_2 acts as an etch stop layer for the dry silicon etching. This layer was then removed by Reactive Ion Etching. The cantilever chips were then mounted on a PCB by gluing the silicon support (figure 1) with epoxy glue. Wire bonding ensures communication between the chips and the PCB.

Figure 3 shows a close-up view of one cantilever obtained by dual-beam optical interferometry (Veeco NT9080). As shown in this figure, the rectangular geometry of the cantilever is clearly defined with respect to the design, as is the gold electrode used for integrated actuation. Also, as shown in figure 1, the release of the cantilever is clearly visible.

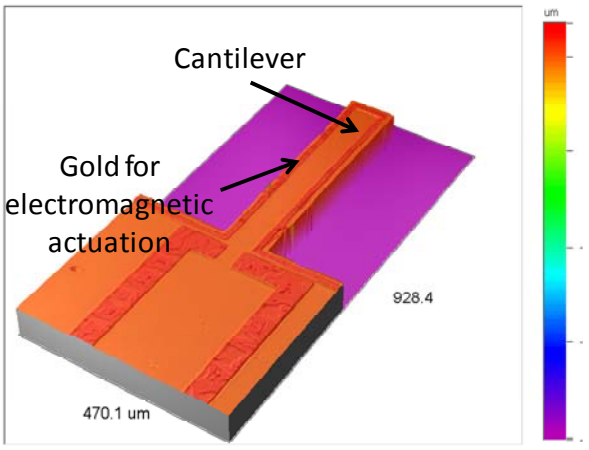


Figure 3: Geometry of the chip obtained by optical profilometry.

B.Rim pre-estimation by means of optical profiler

Prior to the development of methods for estimating the rim length using the resonant behavior of microcantilevers, the value of rim dimension is estimated by means of a dual-beam interferometry profiler. This optical tool was found to be the best compromise to achieve a balance of simplicity, accuracy and minimal time consumption. Concerning the measurement set-up, the cantilever chips were placed vertically under the objective of the profiler, so that the sidewall etching profile of the backside of the SOI wafer could be studied. An example of a profile acquired is shown in figure 4. From this figure, one sees that the verticality of the sidewall is not perfect, and the silicon over-etching is clearly visible. An undercut below the top silicon layer is evident, from which the rim length may be determined. For the different chips tested, the values of rim length that were measured are summarized in table 1. From this table, it can be seen that the rim lengths range from 15 to 35 µm. The values obtained are consistent with the position of the chips on the wafer, since silicon over-etching decreases as one moves from the center to the periphery of the SOI substrate. (Chip A7 is at the periphery of the wafer while chips A100 and A102 are close to the center of the wafer.)

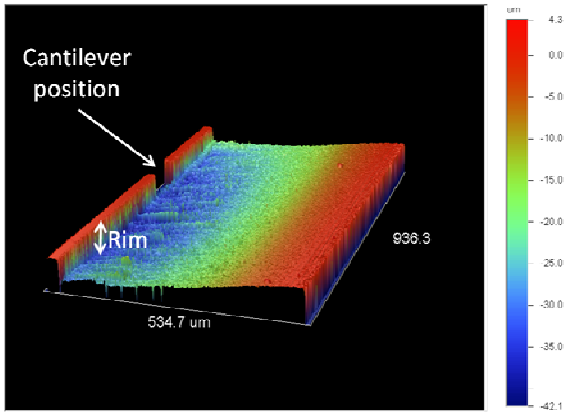


Figure 4: Sidewall etching profile measured by optical profilometry.

Table 1: Measured rim length (optical profiler).

Chip #	Measured rim length (µm)
A100	34.7 ± 0.8
A102	32.2 ± 1.1
A63	35.5 ± 1.4
A7	15.6 ± 0.5
A28	26.7 ± 3.2

C. Rim length evaluation by means of cantilever resonance

In the present work, two methods based on the resonance behavior of microcantilevers are proposed to estimate the dimension of the rim, whose existence is primarily responsible for the inaccuracy of the conventional method for

estimating resonant frequency (1). The first method uses the value of the microcantilever resonant frequency, while the second one uses the deflection profile at resonance (i.e., the deformed beam shape) of the actual cantilever. For both methods, the first out-of-plane flexural resonant mode has been studied using a Polytec MSA-500 optical vibrometer. Actuation of the structures is performed by electromagnetic forces, while the deflection spectrum is acquired via the laser vibrometer, allowing a precise determination of resonant frequency (<1 Hz resolution). Also, the vibrometer system allows a specific meshing of the vibrating structure, so that a specific resonant mode shape can be studied quantitatively. Making use of this capability, the cantilever deflection profile at resonance has been acquired. A resulting ratio (R) between deflection at the tip of the cantilever and the deflection at the clamped-end (i.e. at the beam/rim interface) has been calculated and compared to a calibration curve obtained via Finite Element Modeling (FEM) using COMSOL. For the resonant frequency method, a similar approach is proposed, but the values that are compared are those of resonant frequency. Using both methods, a rim value is determined by comparison between experimental data and the calibration curve.

IV. Results and Discussion

By utilizing the resonance behavior of the cantilevers, the two methods outlined above have been employed to obtain estimates of the rim length, i.e., without the need of a backside (or side) observation of the chip. In the following, the details of the methods are given in addition to the results.

A. Resonant frequency method

The resonant frequency method is based on the fact that the resonant frequency clearly depends on the value of rim length because the presence of the rim induces a different mechanical rigidity at the supported end of the cantilever. Indeed, in figure 5 a calibration curve has been established via FEM where the simulated structure corresponds to the one designed, composed of bare silicon, PECVD SiO_2 and a gold electrode. As expected, the figure shows a decrease of resonant frequency when the rim length increases. This is mainly due to the loss of rigidity at the clamped-end of the cantilever when rim length increases. Also, in this figure has been plotted the value of resonant frequency measured on chips A7, A28, A100, A102., These chips, composed of cantilevers characterized by the same geometry, provide different resonant frequencies ranging from 92.5 kHz to 97.9 kHz. This indicates that different values of rim length are present for the tested chips. Indeed, by comparison between the calibration curve and the experimental values, rim length values ranging from 20.4 µm and 49 µm, have been determined. However, these values do not perfectly fit with values determined with the optical profiler, as shown in table 2. Note that in all cases, the proposed method overestimates the rim length, indicating that there exists other "softening effects" (e.g., support compliance) that the proposed method does not include. A discrepancy between 14.4% and 30.8% in rim length estimation is observed for three of the four chips. The remaining chip (A28) has an 83.5% difference because the measured values of the frequency and rim length do not follow the trend of the other data (i.e., larger rim length should result in lower frequency). Thus, this data point may be suspect. Note that a slight loss of accuracy may result from other parameters not considered in the simulation, such as the exact thickness of the top silicon substrate or the thickness nonuniformity of the gold electrode. However, with the resonant frequency method, a simple and rapid estimation of rim has been demonstrated.

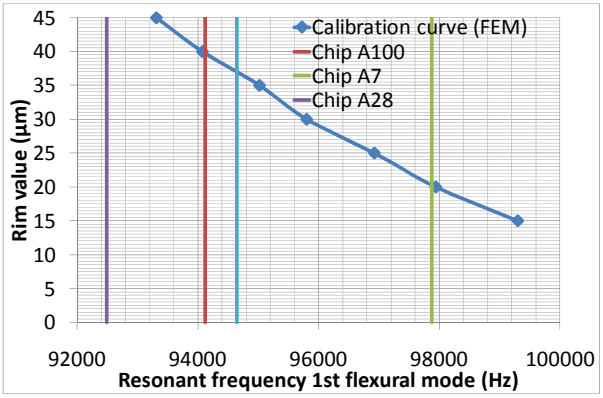


Figure 5: Determination of the rim length value using the resonant frequency measurement and the one obtained by COMSOL simulation.

Table 2: Comparison of the rim length estimation using resonant frequency measurement and the one obtained by optical profilometry.

	A100	A7	A28	A102
estimated rim length (optical profiler) (µm)	34.7	15.6	26.7	32.2
measured resonant frequency (Hz)	94116.2	97877.9	92492.1	94648.4
estimated rim length with resonant frequency (µm)	39.7	20.4	49	37

B. Resonant profile method

Since the rim influences the rigidity of the supported end of the cantilever, the large deflection obtained at resonance can induce a deformation in the rim structure and this deformation will, of course, depend on the rim length. With this in mind, the influence of the rim length on the deflection profile of the cantilever at resonance has been studied. In figure 6, the normalized resonant profiles of cantilevers for different values of rim length have been simulated by FEM. From this figure, the dependence of the resonant profile on rim length is evident, so that the ratio between tip deflection and deflection at the clamped-end (rim/cantilever interface) is proposed as a metric for rim length estimation. By comparison with the calibration curve obtained by FEM, values of rim length have been estimated, ranging from 18.3 µm to 26.7 µm as shown in table 3. With this method, it is observed that the rim length tends to be underestimated (in contrast to the frequency method) with the discrepancies being larger (between -30.7% and 17.3%) than those obtained

with the resonant frequency method (Again, the largest difference (-30.7%) corresponded to Chip A28).
However, given the simplicity of the method and the possible error sources, mainly the difficulty in determining the exact position of the clamped-end of the cantilever for the determination of the deflection ratio, the values of rim length obtained give an approximate estimation of this important parameter.

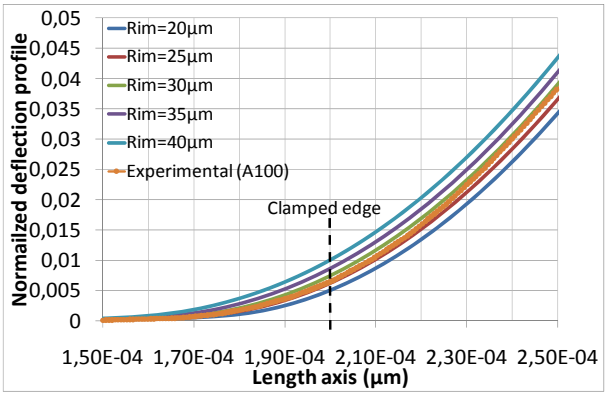


Figure 6: Deflection profile at resonance obtained for different values of rim length (COMSOL simulations).

Table 3: Comparison of the rim length estimation using the resonant profile method and the one measured by the optical profiler

	A100	A7	A28	A102
estimated rim length (optical profiler) (µm)	34.7	15.6	26.7	32.2
measured deflection ratio (w_tip/w_clamp)	151.24	220.87	213.2	176.6
estimated rim length with ratio (µm)	26.7	18.3	18.5	22.7

C.Summary

The values of rim provided by optical profilometry (reference method) and the ones obtained with the alternative methods proposed for the four tested chips are summarized in figure 7. From this figure, it can be seen that the resonant frequency method tends to overestimate the rim length, while the resonant profile method tends to overestimate it (the exception being chip A7). Thus, by combination of both methods, a rim length range for each chip can be determined, while the average value gives an approximate value of rim length with a discrepancy magnitude as low as ~4%, as obtained for the chip A100. These results are encouraging in that they provide some motivation for (a) further development of a combined method based on the resonant frequency and resonant profile methods, and (b) development of analytical models that show how various system parameters, including rim length, influence the resonant behavior of a microcantilever/rim system. We envision the successful development of these types of models as being of paramount importance when using single or multiple (coupled) cantilever devices in sensing applications.

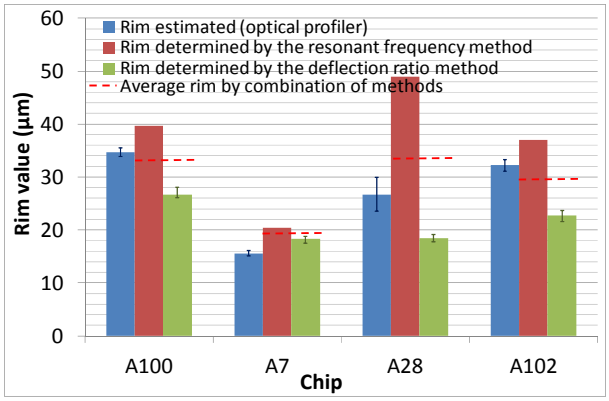


Figure 7: Summary of rim length estimation for the different methods investigated.

V. CONCLUSION

In the present work, two straightforward methods based on resonance behavior of microcantilevers have been proposed and evaluated for the systematic determination of rim length due to silicon over-etching that can occur during DRIE. Using both the cantilever's observed resonant frequency and its deflection profile at resonance, the rim length was determined via comparisons with calibration curves obtained via FEM simulations. By combining both methods, good estimates of the rim length value have been achieved. However, due to the time-consuming nature of performing the requisite FEM simulations, the development of analytical solutions to the problem is desirable. Such analytical models are now under development, so that a better understanding of the complex interplay of system parameters may be obtained for resonating microcantilevers having non-standard support conditions.

[1] R. D. Blevins, Formulas for natural frequency and mode shape, Ed. Van Nostrand Reinhold Company, New-York, ISBN 0-442-20710-7, 1979.

[2] L D Landau, E M Lifshitz, A Course in Theoretical Physics - Fluid Mechanics, Pergamon Press Ltd., Vol. 6, 1987.

[3] I. Dufour, S. Heinrich, F. Josse, "Theoretical analysis of strong-axis bending mode vibrations for resonant microcantilever (bio)chemical sensors in gas or liquid phase", Journal of MicroElectroMechanical Systems (IEEE/ASME Publication), Vol. 16, 2007, pp. 44-49.

[4] Timoshenko et D. H. Young, Advanced Dynamics, McGraw-Hill Book Co., 1948

RF oscillators stabilized by temperature compensated HBARs based on $LiNbO_3$/Quartz combination

T. Baron, G. Martin, E. Lebrasseur, B. François, S. Ballandras
FEMTO-ST, UMR 6174 CNRS-UFC-ENSMM-UTBM
Time and Frequency Department
Besançon, France
sylvain.ballandras@femto-st.fr

D. Gachon
Laboratoire PROMES-CNRS, UPR 8521
Université de Perpignan Via Domitia,
Perpignan, France

P.-P. Lasagne, A. Reinhardt
CEA-LETI, LCRF,
Grenoble, France

Luc Chomeloux
SENSeOR SAS
Mougins, France

Jean-Marc Lesage
DGA, CELAR
Rennes, France

***Abstract*—In this work, we develop a radio-frequency oscillator operating near the 434MHz-centered ISM band to validate the temperature compensated capability of HBAR based on combining $LiNbO_3$ and Quartz single crystal plates for such purposes. Electrical and thermoelectric characterizations have shown QF product in excess of 10^{13} and a third order frequency-temperature behavior. A phase noise better than -160dBc/Hz at 100kHz has been measured as well as a -165dBc/Hz level at 1MHz from the carrier. First results showing that the resonator stability is better than 10^{-9} under room conditions. Further work is expected to lower that level well below 10^{-10}.**

I. Introduction

Radio-Frequency (RF) oscillators can be stabilized by various resonating devices. Their stability is mainly conditioned by the spectral quality of the resonator even if the oscillator loop electronics must be optimized to lower the generated noise as much as possible. For mid-term stability, temperature compensation as a key point and allows for notably improving the corresponding figure of merit. In a previous work [1], we have shown the possibility to build temperature compensated High overtone Bulk Acoustic Resonators (HBARs) combining $LiNbO_3$ and Quartz single crystal plates for the excitation of high order modes exhibiting cubic frequency-temperature behavior. In this work, we develop a radio-frequency oscillator operating near the 434 MHz-centered ISM band to validate the capability of the above-mentioned HBAR for such purposes. The composite substrate have been built using a 3-inch (YX*l*)/163° $LiNbO_3$ cut wafer bounded and thinned down to 15 µm onto a 350µm thick (YX*lt*)/34°/90° quartz base.

Single-port resonators operating near 434 MHz (exploiting the third harmonic of the thinned niobate plate as the HBAR transducer) have been then manufactured. Electrical and thermoelectric characterizations have shown quality factor of the resonance in excess of 25000, yielding a QF product in excess of 10^{13} and a third order frequency-temperature behavior. A SAW filter was used to select the ISM band and to filter the high spectral density HBAR response. The oscillator then has been measured using a phase noise automatic bench (R&S FSUP26 spectrum and phase noise analyzer). A phase noise better than -160dBc/Hz at 100kHz has been measured as well as a -165dBc/Hz level at 1MHz from the carrier. Mid term stability also have been achieved, first results showing resonator stability better than 10^{-9} under room conditions. Further work are expected to lower that level well below 10^{-10}.

The first section of the paper discusses the design of the resonator based on thermoelastic computation of the temperature coefficients of frequency (TCF) of each mode of the structure. The basic theory is first recalled and general relations are provided to select the material stack thicknesses for a given mode order/frequency operating condition. A very brief description of the device fabrication is provided and experimental characterization of the resonators is reported. The oscillator architecture is depicted and the stability measurements detailed. As a conclusion, comparison with SAW-based oscillator is achieved and possible improvements of the oscillator is discussed

II. Device analysis and design

A. HBAR Temperature Coefficient of Frequency (TCF)

The celebrated Campbell&Jones method [2] is used here for predicting the TCF of any mode of a given HBAR. As it has been reported hundred times in previous papers, only the main basic equation is reported below:

$$f=\frac{V}{2e} \rightarrow \frac{df}{f}(T)=\frac{dV}{V}(T)-\frac{de}{e}(T) \qquad (1)$$

which means that the frequency changes due to temperature variations are computed as the difference between the development of the velocity and of the stack thickness vs temperature. As discussed in [1], Lithium Niobate and Quartz have been associated for the fabrication of shear-waver based HBARs. $LiNbO_3$ provides crystal orientations for which very strongly coupled shear waves exist (k_t^2 in excess of 45%) whereas AT cut of Quartz allows for compensating second order frequency-temperature effects [3]. Although this idea was already proposed using other material combinations [4], no real design process was presented until now and therefore the possibility to actually determine structures allowing for high frequency operation with first order TCF smaller than 1ppm.K^{-1} was quite hypothetic. Second order compensation was even not imagined in [4]. Figure 1 shows the evolution of the TCF_1 of almost all the modes of thinned (YX*l*)/163° $LiNbO_3$ layer atop a (YX*lt*)/36°/90° Quartz plate for various Niobate/Quartz thickness ratios, always assuming a Niobate layer much thinner than the Quartz substrate (which allows for neglecting the differential thermoelastic stress within the stack). One can see that depending on the harmonic number, the TCF_1 changes from +1 to -14ppm.K^{-1}. Furthermore, depending on the harmonic of the transducer alone, the TCF_1 may notably change and thus it cannot be considered as a simple periodic function vs harmonic number. Therefore, it is mandatory to accurately consider all the actual feature of the structure for an accurate design of a resonator, i.e. the operation frequency, the harmonic number and the thickness ratio for a given structure. To complete this, one should also account for the actual thickness of the device as this parameter will control the possibility to select one (frequency/harmonic number) couple. Finally, it clearly appears that the analysis of such HBAR TCF requires a numerical analysis and that if an intuitive approach allows for a first order definition of crystal orientations, the complicated distribution of energy within the stack vs all the structure parameters induces more intrication in the design process.

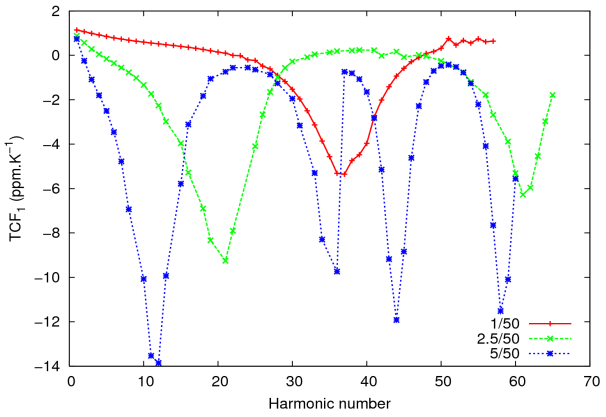


Figure 1. Plot of the TCF of a HBAR built on a (YX*l*)/163° $LiNbO_3$ thinned plate bonded on (YX*lt*)/36/90 Quartz substrate for various Niobate/Quartz thickness ratio (Quartz thickness arbitrary fixed to 50µm)

As an illustration of these considerations, the evolution of the TCF of the most strongly coupled mode of the structure (coinciding with the fundamental mode of the niobate plate alone) has been plotted vs niobate thickness (Fig.2). A TCF_1 vs Niobate/Quartz ratio law has been derived from this curve, yielding the following expression:

$$TCF_1 = 42.5 \times R_t + 3.2 \times \log(200\, R_t) \quad (2)$$

where R_t is the Niobate/Quartz ratio, thicknesses and TCF_1 being expressed in µm and in ppm.K^{-1} implicitly.

Considering now the TCF_1 evolution vs quartz cut angle θ (rotation around X axis), one can set the following complementary law:

$$TCF_1 = (\theta - 35.) \times 5.10^{-6} \quad (3)$$

Using these relations allows one to show that the most coupled mode of the structure is temperature compensated for R_t=1/50 for the (YX*lt*)/34°/90° quartz cut, whereas the (YX*lt*)/33°/90° and (YX*lt*)/32°/90° cuts will be selected for R_t respectively equal to 2.5/50 and 5/50.

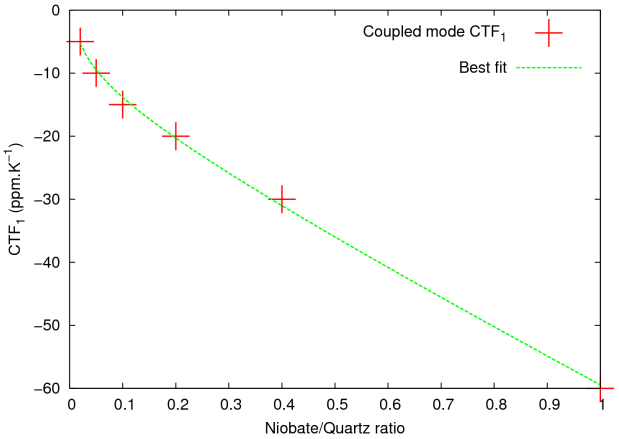


Figure 2. Plot of the TCF of a HBAR built on a (Yxl)/163° $LiNbO_3$ thinned plate bonded on (Yxlt)/36/90 Quartz substrate for various Niobate/Quartz thickness ratio (Quartz thickness arbitrary fixed to 50µm)

B. 434 MHz resonator design

The design of the 434 MHz centered resonator has been done considering our capability to thin down $LiNbO_3$ wafers. A typical thickness of 15µm±2µm then must be considered, the Quartz thickness being deduced along the theoretical analysis presented above. Assuming a phase velocity of the shear mode in the $LiNbO_3$ layer near 4000m.s^{-1}, the third harmonic of the fundamental mode then is easily excited at the expected frequency. To achieve the temperature compensation of the mode, the quartz cut angle and plate thickness then has to be fixed. A R_t value of 2.5/50 is considered as a trade-off between the spectral density of modes and the ease to manufacture the material stack.

Consequently, the quartz thickness is found near 270µm for a cut angle corresponding to (YX*l*)/34°/90°. Once again for manufacture simplicity, standard quartz wafer thicknesses are considered, yielding a value of 350µm.The $LiNbO_3$ thickness then is adjusted to 14µm according to the above considerations (R_t actual value 2/50). Validation of such a configuration has been achieved theoretically by computing the admittance of the corresponding device and then its theoretical TCF. An equivalent quality coefficient Q of 20,000 has been considered for Quartz whereas a Q of 100,000 was affected to $LiNbO_3$. Au electrodes are used at the interface (t = 300nm, Q~1000) and Al ones at the top surface (t = 250nm, Q~2000). An effective Q value slightly larger than 20,000 then is predicted for a k_s^2 of 1,9‰ (Fig.3). Temperature compensation is checked by plotting both TCF_1 and second order TCF_2 for each resonance near 434MHz, as shown in Fig.4. For such a configuration, one can see that TCF_1 smaller than 1ppm. K^{-1} can be obtained together with TCF_2 of less than 5ppb.K^{-2}.

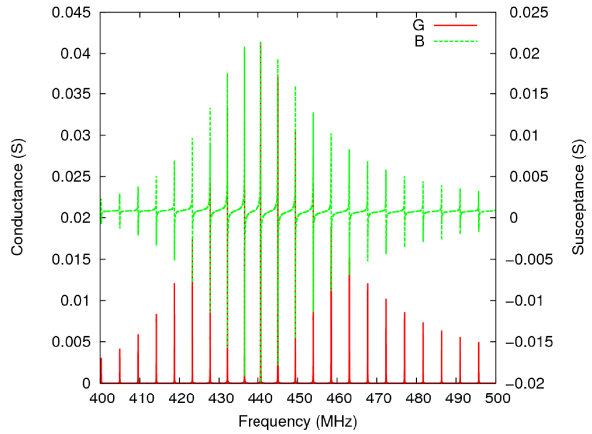


Figure 3. Theoretical admittance of a HBAR built on a 14µm thick (YXl)/163° $LiNbO_3$ plate onto 350µm thick (YXlt)/34°/90° Quartz

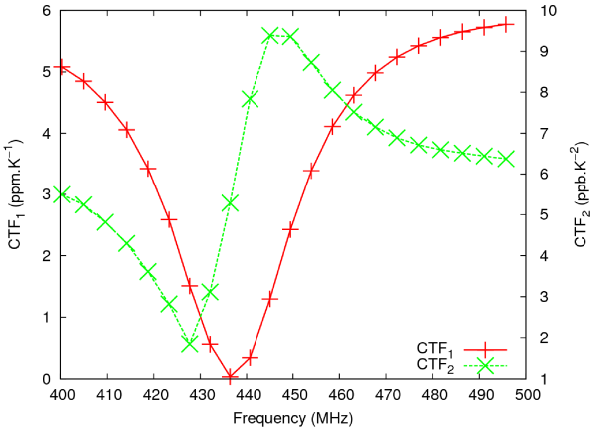


Figure 4. 1st and 2nd order TCF of the above $LiNbO_3$/Quartz HBAR, temperature compensation achieved between 430 and 440MHz.

III. Experimental implementation

A. *HBAR fabrication*

As mentioned above, only some information is reported here as most of the process has been already presented in details [1]. $LiNbO_3$ and Quartz are bonded together using thin Cr/Au films deposited on both wafer sides to be bounded and therefore used both as diffusion layers and interface electrode. A pre-bonding is first achieved using either an EVG bounder or a Suss Microtech CL200 megasonic cleaner. The bounding then is reinforced by applying a quite strong homogeneous pressure during a few tenth minutes, yielding then an almost perfect acoustic adhesion between both material. Top electrodes are then deposited using standard lithography processes.

B. *Resonator characterization*

Conforming to the above design and manufacturing process, several devices have been fabricated and tested to select the most appropriate one for frequency stabilization in the ISM band. Figure 5 shows one example of resonator admittance from which electromechanical coupling k_s^2 and Q factor are deduced, respectively at about 0.2‰ and 25,946, yielding a Q factor of 1.13×10^{13}, much better than Rayleigh wave based resonators on Quartz considering the same frequency. One can note that Q factor in excess of 30 000 also were measured (QF products in excess of 1.3×10^{13}), but for frequencies slightly out of the ISM band. The TCFs have been again tested to demonstrate the predicted compensation effect.

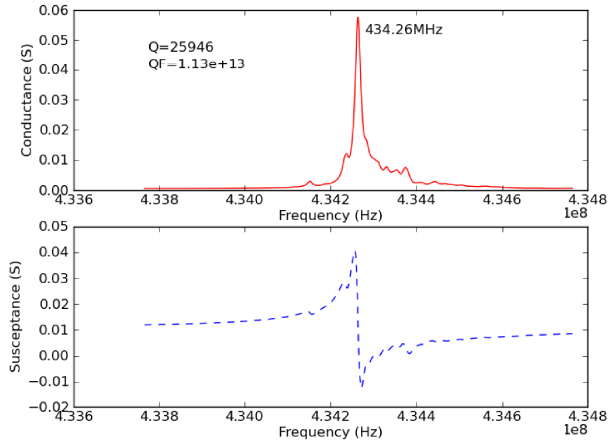


Figure 5. Experimental measurements (admittance) of a HBAR based on (YXl)/163° $LiNbO_3$ plate onto 350µm thick (YXlt)/34°/90° Quartz

C. *Oscillator stabilization*

The implemented oscillator once again is very similar to what has been already presented for SAW devices [5] and therefore will not be detailed here. It is a Colpitts-type negative resistance circuit to which a buffer stage has been

added to isolate the oscillating loop and provide sufficient signal for the oscillator characterization. Figure 6 shows the characterization of the oscillator output using a spectrum analyzer, showing the definition of the resonance and also proving the possibility to actually select the ISM band by filtering.

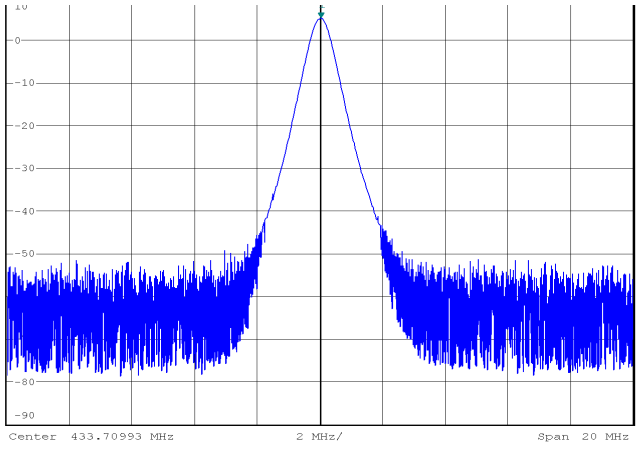


Figure 6. Spectrum of the oscillator output signal, close to the ISM center (433.7MHz)

Figure 7 shows the phase noise of the oscillator. The noise floor is at -165dBc/Hz and the phase noise at 10kHZ is -140dBc/Hz.

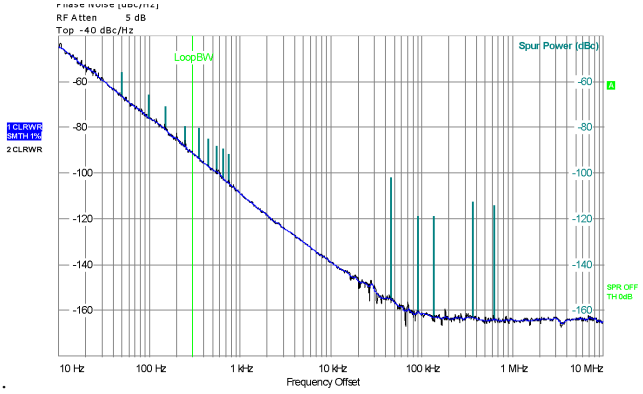


Figure 7. Phase noise of the Colpitts oscillator.

Figure 8 show Allan variance of the oscillator output signal. The oscillator based on HBAR temperature compensated presents a relative stability of near 2×10^{-10} over 2 seconds of integration at room conditions.

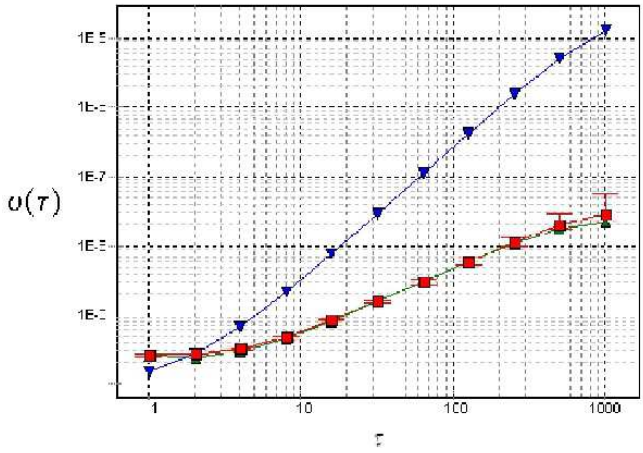


Figure 8. Allan Variance of the oscillator output signal, yielding a relative stability near 2×10^{-10} over 2 seconds of integration at room conditions

IV. Conclusion

In this work, we develop a radio-frequency oscillator operating near the 434MHz-centered ISM band to validate the temperature compensated capability of HBAR based on combining LiNbO3 and Quartz single crystal plates for such purposes. Electrical and thermoelectric characterizations have shown QF product in excess of 1013 and a third order frequency-temperature behavior. A phase noise better than -160dBc/Hz at 100kHz has been measured as well as a -165dBc/Hz level at 1MHz from the carrier. First results showing that the resonator stability is better than 10-9 under room conditions over 2 seconds. Further work is expected to lower that level well below 10-10. To achieve this objective, excitation of resonator will be increase, test will be done under controlled temperature condition, and electromechanical coupling coefficient will be increase. Towards ultimate stability allows us competition with SAW based oscillators.

References

[1] T. Baron et al, Temperature compensated radio-frequency harmonic bulk acoustic resonators, Proc.of the IEEE IFCS, pp. 652 – 655, 2010

[2] J.J. Campbell, W.R. Jones, "A method for estimating crystals cuts and propagation direction for excitation of piezoelectric surface waves", IEEE Trans. On Sonics and Ultrasonics, Vol. 15, pp. 209-217, 1968

[3] S. Ballandras, D. Gachon, RESONATEUR HBAR A STABILITE EN TEMPERATURE ELEVEE, French patent #FR2932333 (B1), PCT extension WO2009156658 (A1)

[4] Curran Daniel R & Al, US Patent #3401275A, 1968-09-10

[5] B. François, Fabrication and Characterization of a SAW-Oscillator-based Sensing System Including an Integrated Reciprocal Counter and a Wireless ZigBee Transmission System, Proc. of the IEEE International Ultrasonics Symposium, 2010

This work is supported by the FCE – DGCIS under grant #092906659/60/61.

Passive Tuning in Lateral-Mode Thin-Film Piezoelectric Oscillators

Mohsen Shahmohammadi, Derya Dikbas, Brandon P. Harrington and Reza Abdolvand
Department of Electrical and Computer Engineering, Oklahoma State University, Tulsa, Oklahoma, USA
mohsen.shahmohammadi@okstate.edu

Abstract — **In this paper, for the first time frequency tuning in lateral-mode thin-film piezoelectric resonators is demonstrated through varying the termination load connected to an isolated tuning port. The dependency of frequency tuning on the design features of the resonator is studied as well. More than 4300ppm tuning range in a thin-film piezoelectric-on-silicon (TPoS) resonator and more than 9000ppm tuning range in a pure piezoelectric resonator is achieved at ~29.5MHz and ~31.5MHz respectively. Also, an oscillator tuning range of more than 2700ppm is reported which shows a great potential for temperature compensated oscillators.**

I. Introduction

Quartz crystal resonators have been widely employed in clock generators in communication systems since they offer high quality factor (Q) and excellent temperature stability. However, the drawback of such resonators is that they are not compatible with CMOS fabrication process, which complicates the integration of the high-performance oscillators with the rest of electronics. This issue will be of much greater importance for realization of future generations of data and communication systems in which oscillators at different frequencies are demanded in a very small footprint area.

MEMS resonator technology has gradually become popular as a solution for the integration problem. Many researchers and companies have presented high-performance MEMS-based oscillators so far [1, 2]. In the past, our group has shown that thin-film piezoelectric-on-substrate (TPoS) MEMS resonators can yield superior power handling and Q (in air) due to the excellent power density and low acoustic loss of their silicon resonant body [3, 4]. Such advantages make TPoS resonators as great candidates to be employed in low-noise oscillator circuits.

Although, lateral-mode piezoelectric resonators such as TPoS devices have been successfully incorporated in low-power and low-noise oscillators, they have limited frequency stability versus temperature (relatively large temperature coefficient of frequency, TCF). Also, methods of frequency tuning have not yet been studied extensively for such devices. Simple and effective tuning schemes can be exploited in temperature compensated oscillators.

In capacitive resonators, frequency tuning is usually pursued through altering the effective stiffness of the resonator by varying the bias voltage across the parallel plate capacitive transducer [5, 6]. In these resonators effective tuning at high frequencies (10's of MHz and above) is only achievable with very small transduction gaps and large tuning voltages, rendering the technique impractical. Tuning based on heating the resonant structure has also been studied in the past [7, 8]. The drawback is that although the frequency stability achieved in such scheme can be considerable, but the large power consumption is prohibitive for all battery-powered applications. In this paper we exploit the well-known concept of altering the mechanical stiffness of a vibrating piezoelectric film through varying the electrical impedance connected to the conducting electrodes. The novelty in this work is in separating the tuning port from the signal port which will be included in the oscillation loop. This way the termination impedance can be altered to the extremes (open and short) without largely affecting the oscillation loop and therefore the tuning range is expanded. We will experimentally investigate the effect of the resonator design features such as tuning electrode position, relative electrode sizes, and relative thicknesses of the layers in the resonant structure on the maximum achievable tuning.

II. Overview of TPoS Technology

As shown in the schematic of Fig. 1, a TPoS resonator consists of a thin piezoelectric film (e.g. aluminum nitride, AlN) sandwiched between two metal electrodes (e.g. molybdenum, Mo) stacked on top of the device layer of a silicon-on-insulator (SOI) substrate [9]. The bottom Mo is patterned to reduce the parasitic capacitor of the input and output traces. The top and bottom Mo are dry etched in SF6 and O2 plasma to form the top electrodes and the traces while the AlN is wet etched to clear access to the bottom electrode. The resonator body is defined in a plasma etching step. The backside handle layer silicon is etched in a deep reactive ion etching chamber, which creates the cavity below the resonator. The final release step is done in buffered oxide etch (BOE) solution to remove the buried silicon dioxide layer and create the free-standing structure. TPoS resonators can be designed in either one-port or two-port configurations. In a two-port design (Fig. 1) usually one port is used to excite the resonance and the second port is used to sense the vibration. The electrode pattern is designed to yield a minimum series impedance (maximum output current) at the mode-shape of interest. For example the pattern of Fig. 1 can be used for both

This work was funded by Integrated Device Technology Inc.

fundamental and the third harmonic bulk width-extensional modes of the structure.

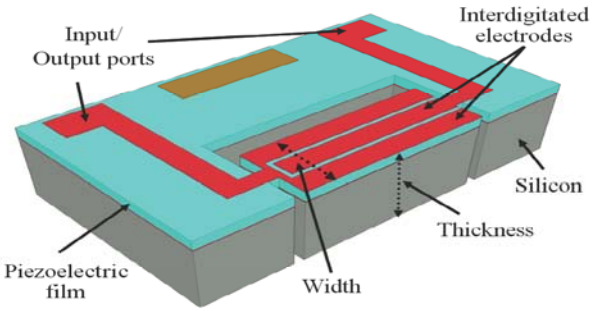


Figure 1. The schematic of a TPoS resonator

Using the TPoS technology we have successfully demonstrated very high quality factor, low motional impedance, and high power handling in air. A frequency response of an optimized fundamental-mode TPoS resonator similar to the schematic picture of Fig. 1 and designed for operation at 27MHz is shown in Fig. 2. The combination of 44500 unloaded quality factor and 200Ω motional impedance in air as demonstrated in this figure is a unique feature not reported by other technologies. Also, there are no significant spurious resonance peaks in the vicinity of the main peak as captured in the 20MHz span of Fig. 2.

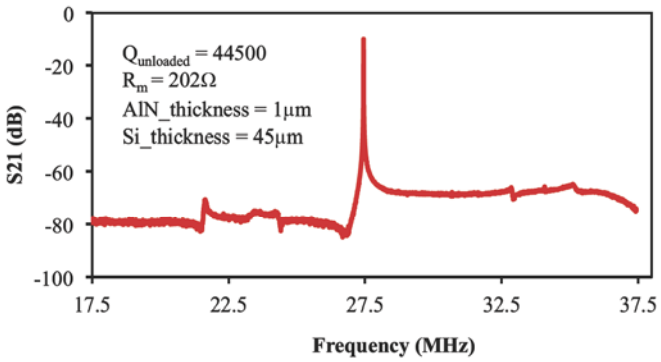


Figure 2. The frequency response of a typical TPoS resonator in air

III. Passive Frequency Tuning in TPoS Resonators

It is known that the effective modulus of a piezoelectric layer (and consequently the resonance frequency) can be altered by varying the shunt impedance connected to a pair of conductive electrodes that cover the surfaces of the film [10]. As shown in Fig. 3 the modulus will change between a minimum at short termination and a maximum at open termination as the parallel impedance is varied. For a short termination the effective modulus (E_{Short} in Fig. 3) depends only on the material compliance (S^E) while for other terminations it is a function of not only the material compliance but also the piezoelectric properties and the shunt impedance (d, ε_p, Z_T and Z_p). In Fig. 3, E is the effective modulus, d is piezoelectric constant, ε_p is the permittivity of piezoelectric material and Z_T and Z_p are shunt impedance and the capacitive impedance of conductive electrodes respectively.

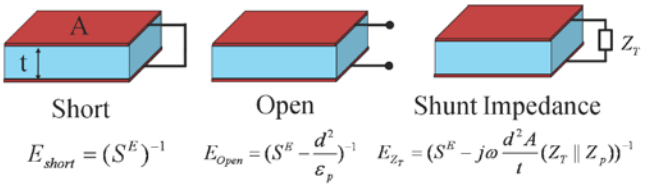


Figure 3. Dependency of the effective modulus of a metalized piezoelectric film on the electrical termination [10]

We have exploited this concept in the lateral-mode TPoS resonators. The proposed configuration for the frequency tuning purpose is shown in Fig. 4. One of the two terminals in the two-port resonator (port 2) is designated to tuning. Varying the shunt impedance (Z_T) connected to this port changes the effective modulus of the resonant body and the minimum and maximum impedance determines the resonance frequency limits. The other port (port 1) will then be connected to the oscillator circuit. To investigate the tuning range the reflection S-parameters (S11) are measured from port one as the second port is either floating (creating open termination) or connected to a 50Ω terminated probe (mimicking close to a short termination). The resonator impedance is then calculated from the measured S11 data.

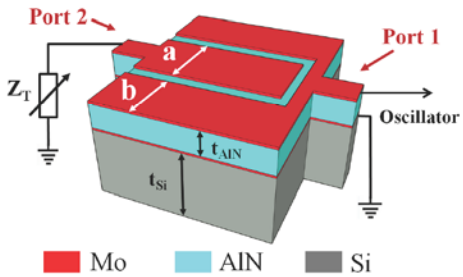


Figure 4. The proposed TPoS resonator electrode configuration for passive tuning purposes

IV. Measurements

A. Tuning in TPoS Resonators

TPoS resonators with variable electrode pattern were fabricated on SOI wafers with different silicon thicknesses for characterization of the tuning range. Two sets of pure AlN devices [11] were also fabricated on bare silicon wafers with 1.5 and 2μm thick AlN film for comparison. The frequency response of a 29MHz fundamental-mode resonator is shown in Fig. 5. This design is fabricated on a SOI wafer with 5μm silicon and 2μm AlN. The impedance vs. frequency plot is calculated from the measured reflection S-parameter (S11). Measurement is repeated for two cases; 50Ω (close to short) and open termination at port two. The results are plotted in Fig. 6. As seen in this figure the resonance frequency is shifting more than 116kHz at ~29MHz (4000ppm) between short and open case. It is observed that the attributes of the measured frequency response other than the resonance frequency are also affected by the termination on the tuning port (e.g. the Q is higher for the open termination). However, the change is not significant to an extent that could disrupt the operation of an oscillator circuit.

As we will discuss in the next few sections the maximum achievable tuning range is a function of electrode position, electrode size and relative Silicon to AlN thickness.

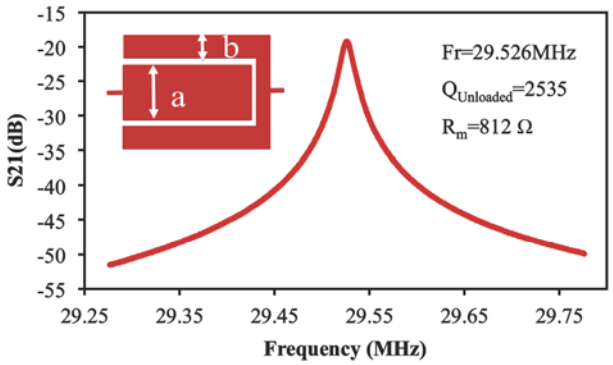


Figure 5. Measured frequency response of a TPoS resonators with a/b=4, Silicon thickness = 5µm, AlN thickness= 2µm

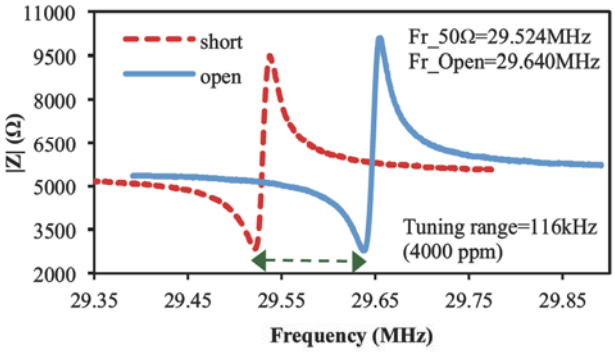


Figure 6. The measured frequency shift for the device above on probe station

1) *Effect of electrode position*

The effect of the electrode position on tuning range is studied by measuring the tuning range for two cases. First; port two is used as tuning and S11 measurement is done at port one and next; the configuration is reversed. It is observed that a larger frequency shift is achieved by utilizing the center electrode (port two) for tuning versus port one. For example the tuning results for the device tested above are shown in Table I. This observation can be explained by the fact that the center electrode covers the area that undergoes the maximum stress in fundamental mode as shown in the simulated mode-shape of Fig. 7. Therefore the contribution of the center area on the overall effective modulus of the structure is larger.

Table I. Measured frequency tuning of a TPoS resonators with a/b=4, Silicon thickness = 5µm, AlN thickness= 2µm

		Frequency(MHz)	Tuning Range(ppm)
Tuning Port: 2	Short	29.524	4000
	Open	29.640	
Tuning Port: 1	Short	29.524	340
	Open	29.534	

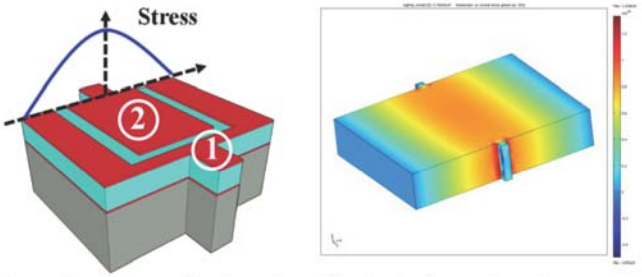


Figure 7. Stress profile along the width of a fundamental-mode resonator

2) *Effect of electrode size*

Several TPoS resonators with different ratios of the center electrode to the side electrode width (a/b) are fabricated and tested. Results show that the frequency tuning range increases as this ratio increases. As it is depicted in Fig. 8 it varies from 1000ppm to 4300ppm for "a/b" ratio of 0.7 to 10.8.

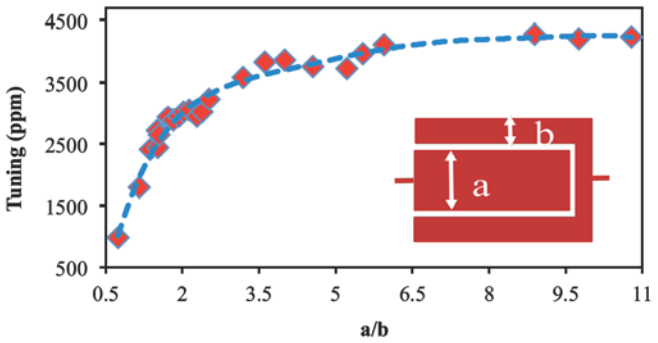


Figure 8. Maximum frequency shift as a function of a/b for TPoS resonators with Silicon thickness= 5µm and AlN thickness= 2µm

3) *Effect of relative Silicon to AlN thickness*

Several devices on different wafers were fabricated and tested in order to study the effect of the relative silicon to AlN thickness on tuning range. It is observed from Fig. 9 that the tuning range increases as the ratio of t_{Si}/t_{AlN} decreases. Table II also shows the tuning range for the same design which is fabricated on two wafers with different AlN thickness. The maximum 4300ppm tuning range at 29MHz is measured for the device with AlN thickness of 2 µm.

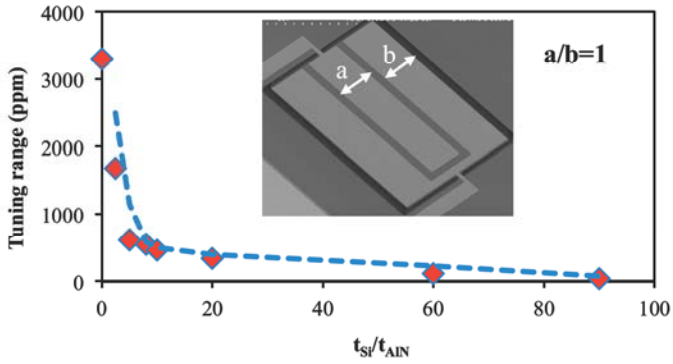


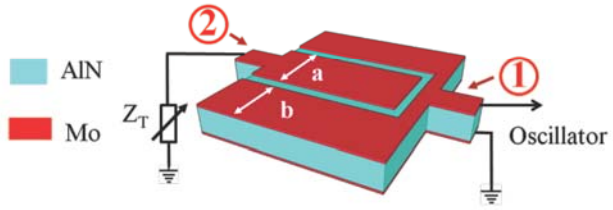
Figure 9. Maximum frequency shift as a function of t_{Si}/t_{AlN} for resonators with a/b=1

Table. II. Frequency tuning range for different AlN thickness, Si thickness=5µm

AlN	a/b	Short (MHz)	Open (MHz)	Tuning Range(ppm)
1.5µm	5.94	28.907	28.989	**2824**
2µm	5.94	29.519	29.647	**4300**

B. *Tuning in Pure AlN Resonators*

Designs are fabricated on a wafer with just AlN similar to resonators in [11]. As it is expected the tuning range increases compared with TPoS resonator since the contribution of the piezoelectric modulus on the overall effective modulus is maximized. The highest overall frequency shift is measured from pure AlN devices which is more than 9000ppm at 31.5MHz. Increasing the thickness of the AlN from 1.5µm to 2µm has also caused the tuning range to increase substantially (Fig. 10).



AlN	a/b	Short (MHz)	Open (MHz)	Tuning Range(ppm)
1.5µm	5.94	30.493	30.745	**8260**
2µm	5.94	31.571	31.860	**9150**

Figure 10. Pure AlN resonator

C. *TPoS Oscillator Characterization*

The schematic of the oscillator circuit used in this work is shown in Fig. 11 where port one on the resonator is integrated in the oscillation loop. The oscillation loop consists of two amplification stages (the second stage is mostly used for phase correction). Output signal waveform is converted to square wave through a comparator in order to enable jitter measurement. On the second port a varactor diode is used to implement the variable impedance.

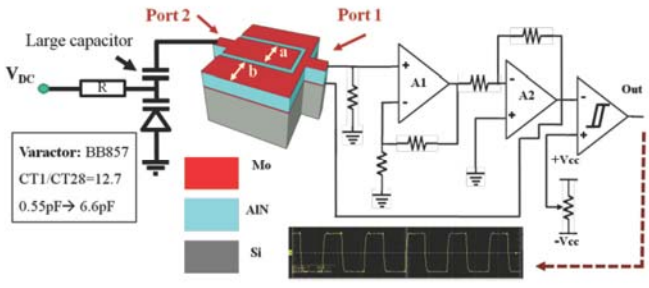


Figure 11. The schematic of the tunable oscillator circuit

Both TPoS and pure AlN resonators were employed in the oscillator. Through applying a tuning DC voltage to the varactor diode the effective shunt impedance at port two is varied. Therefore, the resonance frequency of the resonator and consequently the oscillation frequency of the circuit changes with the DC voltage. Preliminary results show oscillator tuning ranges that are less than the measured frequency shift of the resonator on the probe station. This is because the tuning range is limited by the parasitic capacitor at the tuning port generated by wire bond and PCB tracks which degrade the ideal open condition. Another limitation is the upper and lower capacitance values of the varactor diode which creates a smaller range of shunt impedance compared to what is achievable on the probe station. Maximum oscillator tuning from a TPoS oscillator was measured to be 1200ppm where as a tuning range of 2718ppm was recorded for a pure 2µm thick AlN resonator (Fig. 12) Assuming an approximate TCF of -25ppm/C for the resonator[11], 2718ppm tuning range can compensate for 110°C of temperature change.

The oscillator jitter is measured for the whole tuning range which changes from 22ps to 23ps (cycle-to-cycle standard deviation) between DC voltage of 0v to 25v showing that the jitter is not affected by the tuning impedance.

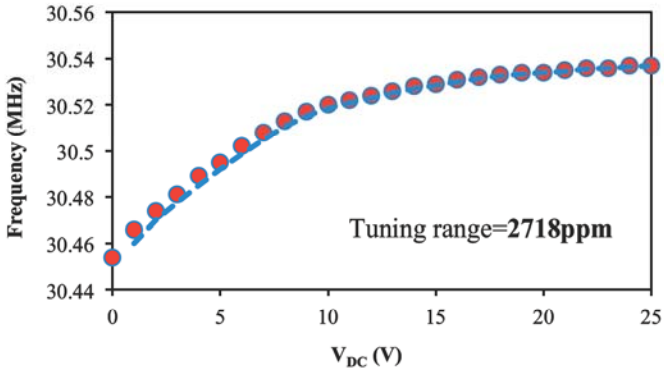


Figure 12. Pure AlN oscillator frequency as a function of DC voltage applied to the varactor

V. Conclusion

In this paper, the novel passive frequency tuning method through varying termination load connected to an isolated tuning port is demonstrated in thin-film piezoelectric lateral-mode resonators. The Effect of resonator characteristics on the frequency tuning range is also studied. Frequency tuning range of more than **4300ppm** in TPoS resonators and more than **9000ppm** in pure AlN devices are reported at ~29.5MHz and ~31.5MHz respectively. An oscillator tuning range of more than 2700ppm is reported which shows a great potential for temperature compensated oscillators.

Acknowledgment

Authors thank Mohammad Jafar Modarres-Zadeh and Hediyeh Fatemi for their contribution in this work.

REFERENCES

[1] W.-T. Hsu and M. Pai, "The New Heart Beat of Electronics - Silicon MEMS Oscillators," in *2007 Proceedings 57th Electronic Components and Technology Conference* (IEEE, 2007), pp. 1895-1899.

[2] H.M. Lavasani, et al., "A 76dBOhm, 1.7 GHz, 0.18um CMOS Tunable Transimpedance Ampflier Using Broadband Current Pre-Amplifier for High Frequency Lateral Micromechanical Oscillators," *IEEE International Solid State Circuits Conference (ISSCC 2010)*, pp. 318-320.

[3] M. Shahmohammadi, B.P. Harrington, and R. Abdolvand, "Concurrent Enhancement of Q and Power Handling in Multi-Tether High-Order Extensional Resonators," in *Proc. IEEE Int. Microwave Symp.,* May 2010, pp. 1452-1455.

[4] M. Shahmohammadi, M.J. Modarres-Zadeh, and R. Abdolvand,"Low jitter thin-film piezoelectric-on-substrate oscillators," in Proc. *IEEE International Frequency Control Symposium (IFCS)*, Newport Beach, CA, 1-4 June 2010, pp.613-617.

[5] F. D. Bannon III, J. R. Clark, and C. T.-C. Nguyen, "High-Q HF microelectromechanical filters," *IEEE J. Solid-State Circuits*, vol. 35, pp. 512–526, Apr. 2000.

[6] G. K. Ho, K. Sundaresan, S. Pourkamali, and F. Ayazi, "Temperature Compensated IBAR Reference Oscillators," Proc. 19th IEEE *International Conference on Micro Electro Mechanical Systems* (MEMS 2006), Istanbul, Turkey, Jan. 2006, pp. 910-913.

[7] A. Hajjam, A. Rahafrooz, and S. Pourkamali, "Sub-100ppb/^{o}C Temperature Stability in Thermally Actuated High Frequency Silicon Resonators via Degenerate Phosphorous Doping and Bias Current Optimization," proceedings, *IEEE International Electron Device Meeting (IEDM)*, 2010.

[8] K. Sundaresan, G.K. Ho, S. Pourkamali, and F. Ayazi, "A Low Phase Noise 100MHz Silicon BAW Reference Oscillator," Proc. of the *IEEE Custom Integrated Circuits Conference* (CICC 2006), Sept. 2006, pp. 841-844.

[9] R. Abdolvand, et al., "Thin-film piezoelectric-on-silicon resonators for high-frequency reference oscillator applications," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 55, pp. 2596-606, 2008.

[10] Clark, William W. 2000. ''Vibration Control with State-switched Piezoelectric Materials,'' *Journal of Intelligent Material Systems and Structures*, 11(4):263–271.

[11] G. Piazza, P.J. Stephanou, A.P. Pisano, "Piezoelectric Aluminum Nitride Vibrating Contour-Mode MEMS Resonators", *Journal of MicroElectroMechanical Systems,* vol. 15, no.6, pp. 1406-1418, December 2006.

High Frequency Dual-Mode Thermal-Piezoresistive Oscillators

Amir Rahafrooz and Siavash Pourkamali
Department of Electrical and Computer Engineering
University of Denver
Denver, CO, USA
E-mail: amir.rahafrooz@du.edu

***Abstract*— This work presents high frequency fully-micromechanical self-sustained oscillators. It has previously been demonstrated that interactions between mechanical strain and thermally generated actuation forces in micromechanical structures through the piezoresistive effect, can result in a positive feedback loop leading to self-sustained oscillation without the need for an electronic amplifier. In our previous work we demonstrated self-sustained oscillation of single-crystalline silicon dual-plate structures with frequencies up to 6.7MHz [1]. This work demonstrates higher frequency thermal-piezoresistive self-sustained oscillators with frequencies up to 36MHz using I^3-shaped resonant structures. Due to the highly nonlinear (mixed flexural-extensional) deformation of actuator/piezoresistors in such structures, they can provide two dominant output frequency harmonics. The dominant harmonic can be selected by changing the resonator bias current.**

I. INTRODUCTION

Micromechanical oscillators have a lot of potential as replacements for conventional mechanical frequency references in electronic systems. They can also be utilized as highly sensitive resonant mass sensors for chemical and biomedical applications. Such oscillators can provide lower cost, smaller size and compatibility with integrated circuit technologies. In order to use a micromechanical resonator as the frequency reference for an electronic oscillator an amplifying positive feedback loop is required. . In case of conventional electrostatic [2] and piezoelectric [3] micromechanical resonators which have been widely studied and developed over the past decade, an external amplification circuitry is required to provide such a feedback.

The other category of MEMS resonators, thermally actuated resonators with piezoresistive readout, have not been studied as intensively and thoroughly as the other conventional MEMS resonators. This is mainly due to the generally presumed disadvantages of thermal actuation, i.e. being a power hungry transduction mechanism and also its established reputation of being suitable for DC or very low frequency applications only. However it has been lately demonstrated that not only such devices can operate at very high frequencies, but also their performance can get better at higher frequencies providing that proper design considerations are met [4]. It is also shown that the power consumption for such structures can reach well below 1mW as their dimensions are scaled down to reach higher frequencies [1]. Furthermore, what is usually ignored is that the power consumption in thermal-piezoresistive resonators makes them active devices. Such devices are capable of absorbing DC power from a source and acting as an AC energy pump initiating and sustaining mechanical oscillations (i.e. playing the role of transistors in electronic circuitry). This concept has been recently demonstrated in [5] showing self-sustained oscillation capability for a thermally actuated flexural mode silicon microstructure oscillating at 1.26MHz. This is enabled by the active nature of such devices and an internal positive feedback loop resulting from the sequential interactions of electrical, thermal, and mechanical events upon passing a constant DC bias current through their structure.

In our previous work we demonstrated self-sustained oscillation of single-crystalline silicon dual-plate structures [1] with frequencies up to 6.7MHz. Furthermore, in order to reduce the power consumption of such resonators, thermal actuators were aggressively narrowed down. Although this method improves the power consumption of the oscillators but it also raises new concerns regarding their performance. Considering the fact that the in-plane extensional stiffness of the dual plate or I^2-BAR structures is determined by their thermal actuators, narrowing them down causes a decrease in the resonance frequency of the structure. Furthermore, defining such narrow features with very high precision is quite challenging leading to significant frequency mismatch between devices in the same batch.
This work demonstrates self-sustained oscillation for higher frequency extensional mode resonators.

II. OSCILLATION MECHANISM AND DEVICE STRUCTURE

The utilized structure in our previous work [1] that is shown in Fig. 1 resonates in its in-plane extensional mode. Since in this resonant mode, the two plates vibrate against each other and not against the anchors, most of the acoustic energy remains in the resonator. In addition, due to the in-plane vibration of the structure, the air damping is minimized. Moreover, the

This work was supported by National Science Foundation under grants MRI-RAPID #1061489 and MRI #0923518.

extensional resonance mode provides relatively high stiffness that keeps a lot of energy in the device and makes the energy loss through air negligible. Therefore, high quality factors and high motional conductances can be measured for such structures. This makes these devices capable of self-sustained oscillation under both vacuum and atmospheric pressures.

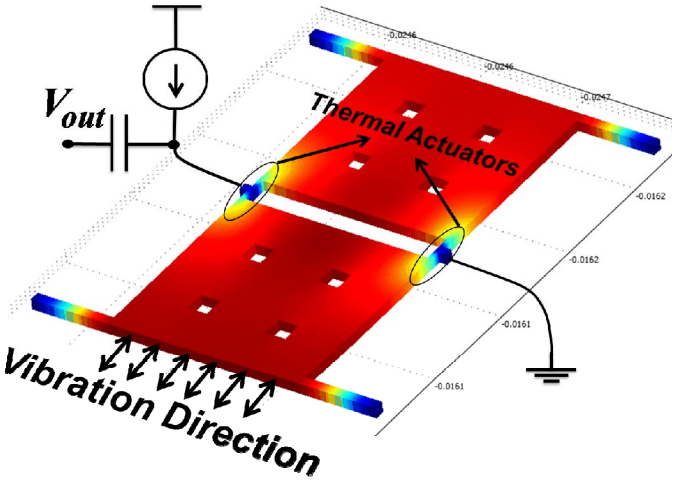


Fig. 1. COMSOL modal analysis, showing the fundamental in-plane resonance mode shape for the dual-plate 4.5MHz resonator which is bias with a constant DC current. Red and blue colors show locations with the largest and smallest vibration amplitudes respectively. The two narrow beams in the middle of this structure are also act as thermal actuators. The used electronic circuitry is to test the oscillation behavior of such resonators

The series of events that leads to self-sustained oscillation can be explained as follows: upon application of a DC bias current, the thermal actuators heat up and expand. Due to the mass (inertia) of the plates, the beams experience an over-expansion after pushing the plates and undergo tensile stress. Since the piezoresistive coefficient has a negative value, the tensile stress decreases the electrical resistance of the actuator beams. When biased with a constant DC bias current, this lowers the ohmic loss in the thermal actuators forcing them to contract. After structural over-contraction (due to the mass of the plates), the actuators will be under compressive stress and will have increased electrical resistance. This translates into a higher thermal power forcing the structure to expand again. If the resulting driving force (due to heating and cooling) in each cycle is large enough to compensate for mechanical losses of the structure, the same sequence is repeated over and over in a self-sustained manner and the vibration amplitude keeps increasing until it is limited by nonlinearities. The 90º phase lags (Fig. 2) resulting from the mechanical and thermal response delays lead to a perfect timing for the abovementioned sequence of physical phenomena to continue.

In order to improve the performance of such oscillators, in this work a wider third bar has been added to the I^2-shaped structures (Fig. 3). In the resulting I^3-shaped structures the stiffness of the resonator is mainly determined by the wider middle bar. Therefore higher resonance frequencies can be achieved while maintaining very narrow actuator beams. For the same reason fabrication induced changes in the dimensions of the narrow thermal actuators will not have a significant effect on the frequency of the resonator. Therefore, it is also expected that fabrication induced frequency variations between similar devices of this types will be highly suppressed.

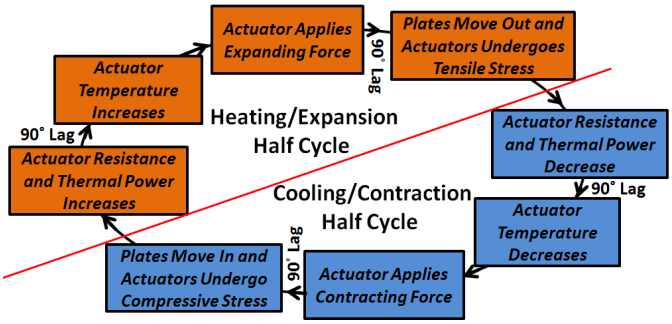


Fig. 2. Sequence of phenomena causing an internal positive feedback loop in thermal-piezoresistive resonators biased with a constant current that lead to self-sustained oscillation.

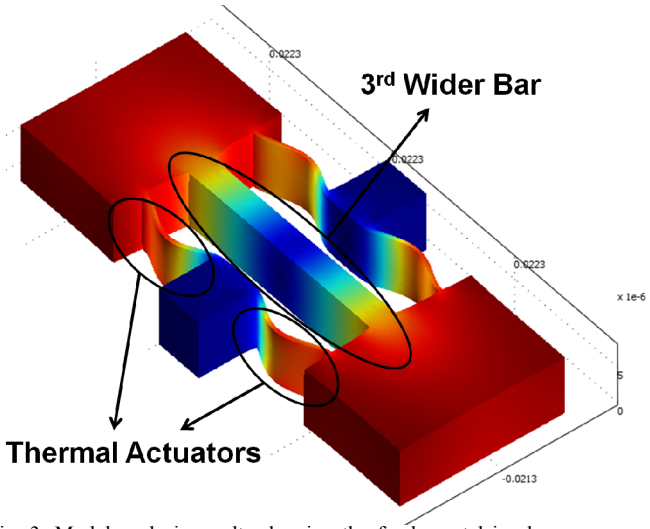


Fig. 3. Modal analysis results showing the fundamental in-plane resonance mode shape for an I^3 resonant structure. In this structure a third wider bar has been added in the middle of the structure to further increase the resonance frequency of the resonator. Red and blue colors show locations with the largest and smallest vibration amplitudes respectively.

III. Fabrication and Measurement Results

Single crystalline silicon I^3-shaped structures were fabricated using the standard single mask SOI-MEMS process [3] on a low resistivity N-type SOI wafer with device layer thickness of 10µm. Since longitudinal piezoresistive coefficient of silicon reaches its maximum value along the <100> direction [6], the resonator actuator beams were fabricated along that direction for optimized transduction. Figure 4a shows the SEM view of a fabricated resonator as well as the schematic view of the required bias current and electrical connections for oscillator operation and measurement.

For this resonator a series of thermal oxidation steps followed by oxide removal in hydrofluoric acid were carried out to narrow down the thermal actuators. Instead of current sources relatively large resistors with values a few times (up to 10X) larger than the electrical resistance of the resonator, were used to provide the resonator bias currents. The AC output

was connected to a high frequency digital oscilloscope. By gradually increasing the bias current, after passing a threshold, the waveform in Fig. 4b with oscillation frequency of 36.54MHz emerged on the oscilloscope. As it can be seen in Fig. 4b, the output signal shape is far from a sinusoidal due to existence of different frequency harmonics. Since the tested resonator has its first in-plane resonance mode at

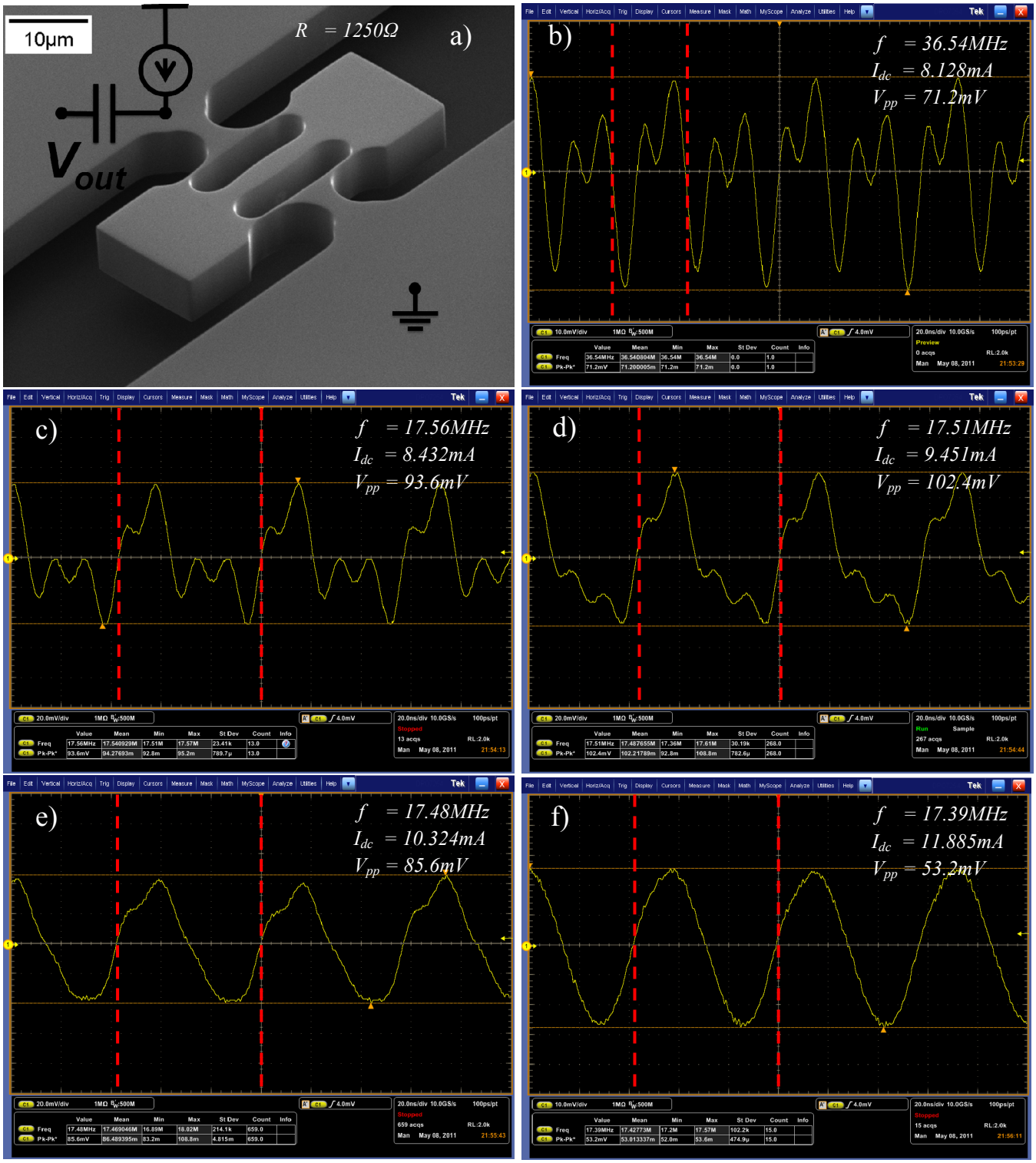


Fig. 4. a) SEM view of a fabricated dual-mode silicon I^3 resonator with oscillation frequencies of b) 36.54MHz at I_{bias}=8.128mA, c) 17.56MHz at I_{bias}=8.432mA, d) 17.51MHz at I_{bias}=9.451mA, e) 17.48MHz at I_{bias}=10.324mA and f) 17.39MHz at I_{bias}=11.885mA both under atmospheric pressure. All the waveforms have the same time interval

17.4MHz, it can be concluded that the second harmonic is the dominant component. The first harmonic shows itself as uneven level of the consecutive peaks in the output waveform. In addition, the small ups and downs in the waveform can be blamed on higher frequency harmonics. It was noticed that by further increasing the bias current, the output voltage waveform constantly changes and at some point the first frequency harmonic with frequency of ~17.5MHz becomes dominant. Figures 4c-4f show how the output waveform changes by increasing the bias current. Figure 4f shows a close to sinusoidal waveform at the same frequency as the mechanical resonance frequency of the structure.

A closer look at the mode shape for the fundamental in-plane resonant mode of the structure reveals that the deformation of the narrow actuator beams is a combination of both extensional and flexural deformations. Figure 5 shows that as the resonator moves from its fully extended state to its fully compressed state in half of the resonance cycle, the narrow support beams go from a fully bent state to their normal straight shape and back into another bent state. Each time the beams flex there will be an increase in their length leading to a frequency component at twice the vibration frequency of the resonator.

As shown in Fig. 4, by changing the applied DC bias current each of these frequency components can become the dominant component determining the output frequency of the oscillator.

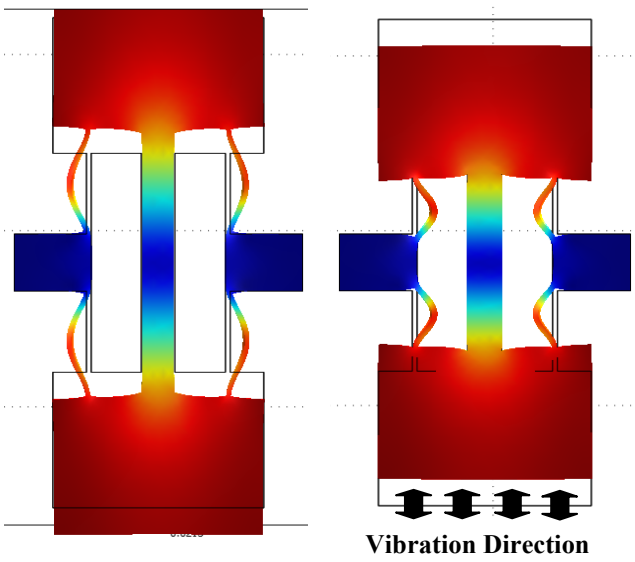


Fig. 5. COMSOL modal analysis results showing the fundamental in-plane resonance mode-shape for an I^3 resonator at its maximum expansion and compression points. It demonstrates that the deformation of narrow thermal actuators is a combination of flexural and extensional deformations and the frequency of the longitudinal deformations of the narrow beams due to flexural deformation is twice the resonance frequency of the resonator.

IV. Conclusions and Future Works

High frequency thermal-piezoresistive I^3-shaped resonators capable of self-sustained oscillation were demonstrated. It was shown that such structures can self-oscillate at higher frequencies compared to the previously demonstrated structures. This design also lowers the fabrication induced frequency variations. Such devices have monolithic crystalline silicon structures and can be fabricated using a very simple CMOS compatible process.

Future work includes further characterization of other versions of such devices and understanding of the factors limiting their oscillation amplitude and long-term as well as short-term frequency stability. Achieving lower power consumptions and higher frequencies are among other future goals.

Although relatively large bias currents were needed leading to relatively high power consumptions for the oscillators, however the power consumption of such structures could be easily lowered to values below 1mW. To do so the resonators should be fabricated on SOI wafers with thin device layer thicknesses rather than using a 10µm thick structure, which is the case in this paper. In addition, the thermal actuators can be further narrowed down to lower their power consumption.

Regarding the stability of such oscillators, phase noise and jitter characteristics could be measured. It is believed that the main source of instability would be the short-term fluctuations in the resonator temperature due to the electronic noise in the bias current and the used bias resistor. Therefore, even higher oscillator stabilities are expected to be achievable via passive temperature compensation of the resonators (e.g. using degenerate doping) [7, 8].

Acknowledgment

Authors would like to thank staffs at Georgia Tech nanotechnology research center and also Dr. Ashwin Samarao for their help with silicon deep reactive ion etching.

References

[1] A. Rahafrooz, and S. Pourkamali, “Fully micromechanical piezo-thermal oscillators,” Proceedings, IEEE International Electron Device Meeting, 2010, pp. 7.2.1 - 7.2.4.

[2] G. K. Ho, K. Sundaresan, S. Pourkamali, and F. Ayazi, “Micromechanical IBARS: tunable high-Q resonators for temperature-compensated reference oscillators,” Journal of Micro-Electro-Mechanical-Systems, Vol. 19, Issue 3, June 2010, pp. 503-515.

[3] C. Zuo, J. V. D. Spiegeland and G. Piazza, "1.05-GHz CMOS Oscillator Based on Lateral-Field-Excited Piezoelectric AlN Contour-Mode MEMS Resonators," IEEE Trans. on Ultrason., Ferroelectr., Freq. Control, vol. 57, no. 1, Jan 2010, pp. 82-87.

[4] A. Rahafrooz, and S. Pourkamali, “High-frequency thermally actuated electromechanical resonators with piezoresistive readout,” IEEE Trans. On Electron Device, 2011, Vol. 58, issue 4, pp. 1205-1214.

[5] P. G. Steeneken, K. Le Phan, M. J. Goossens, G. E. J. Koops, G. J. A. M. Brom, C. van der Avoort and J. T. M. van Beek, "Piezoresistive heat engine and refrigerator," Natrure physics, Vol. 7, No. 4, pp. 354-350, Apr. 2011.

[6] A. A. Barlian, W.T. Park, J. R. Mallon, J., A. J. Rastegar, and B. L. Pruitt,“Fabrication Review: Semiconductor piezoresistance for microsystems,” *Proceedings of the IEEE*, vol. 97, no. 3, Mar. 2009, pp. 513-552.

[7] A. Samarao and F. Ayazi, “Temperature compensation of silicon micromechanical resonators via degenerate doping,” Proceedings, IEEE International Electron Device Meeting, 2009, pp. 789-792.

[8] A. Hajjam, A. Rahafrooz, and S. Pourkamali, “Sub-100ppb/°C temperature stability in thermally actuated high frequency silicon resonators via degenerate phosphorous doping and bias current optimization,” Proceedings, IEEE International Electron Device Meeting, 2010, pp. 7.5.1-7.5.4.

Wafer-scale Packaging for FBAR-based Oscillators

Martha Small, Rich Ruby (IEEE Fellow),
Steven Ortiz, Reed Parker
Wireless Semiconductor Division, Avago Technologies
San Jose, California, USA
martha.small@avagotech.com

Fan Zhang, Jianlei Shi, Brian Otis
Department of Electrical Engineering
University of Washington
Seattle, Washington, USA

***Abstract*— Recent advances in temperature-compensation for FBAR (Film Bulk Acoustic Resonators) have brought this technology forward as a serious contender in the oscillator marketplace. As with any mechanical resonator oscillator, a cost-effective hermetic package combined with circuit technology are critical for commercial application. Billions of FBAR duplexers have been fabricated using Avago Technologies' wafer-scale packaging process, whereby a silicon lid wafer is Au-diffusion-bonded to a base FBAR wafer to make a robust, hermetic package. This paper presents a method for integrating circuitry into the lid wafer to form a sub-0.1 mm^3, sub mW, 1.5 GHz temperature-compensated chip-scale oscillator. Circuit integration, testing and performance will be discussed.**

I. INTRODUCTION

For over 20 years, FBAR or SMR-BAW (Solidly-Mounted Resonator – Bulk Acoustic Wave) with its high f•Q product has been proposed as an attractive alternative to conventional approaches for timing applications [1,2]. Like other mm-scale mechanical resonators, FBAR requires a hermetic package in order to maintain frequency stability. The FBAR is also not readily integrated with circuit technologies for the simple reason that the processing that produces the FBAR – high resistivity, high impedance electrodes; high dielectric constant for the piezoelectric material, and highly variable thicknesses depending on the frequency target – does not produce a viable electronic device. Although integrating a Bulk Acoustic resonator or filter on the same substrate as an IC has been demonstrated [3,4,5], there are three problems with this approach: 1) compromises between technologies must be made to eliminate interference between processing steps (e.g. temperature limitations), 2) the economics for integrating the two technologies on the same wafer is unfavorable [6], and 3) none of the previous work addressed the issue of how to create an open-air cavity around the BAW device while simultaneously providing a hermetic and robust package.

Our previous work demonstrated the potential value of an FBAR [7], a wafer scale hermetic package [8], and a modified FBAR demonstrating temperature stability comparable to quartz (referred to as Zero Drift Resonators, or ZDR [9]). We have also demonstrated low noise sub-mW oscillators using FBAR [10]. The work presented here is the culmination of those advancements. We now present a technique for integrating active circuitry into the lid of the wafer-scale hermetic FBAR package while the FBAR resonator, ZDR, or filter resides on the base wafer. Both sides of an opened and "unfolded" module are shown in Figure 1. The 285µm x 330µm active circuit area is clearly visible, as is the location for the FBAR resonator. This strategy is IC technology "agnostic". Besides an area limitation, there are no special circuit design rules required for our technique, and process modification for the lid/IC wafer is minimal compared to a full SOC (system-on-a chip) implementation.

II. PROCESS DESCRIPTION

The wafer fabrication uses existing processes where possible, which has the advantages both of simplicity and of maintaining native device performance. In the Avago microcap process, the FBAR is fabricated on one wafer while a second "lid" wafer contains through-wafer vias, Au pads, sealing structures, and a recessed air cavity above the FBAR.

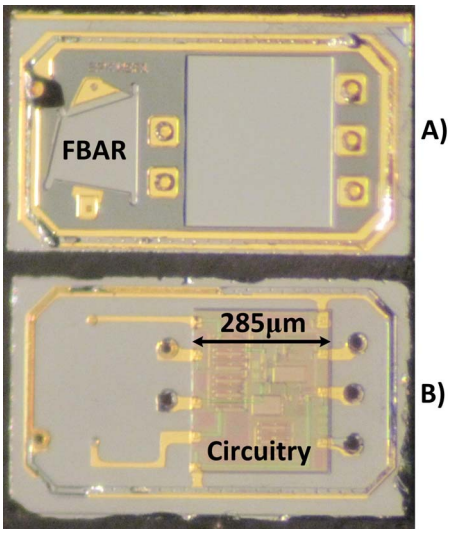


Figure 1. Photograph of disassembled hermetic oscillator package containing A) FBAR die B) lid with integrated active circuitry.

Figure 2 contains a simplified process flow for a standard FBAR/micro-cap part. On the FBAR wafer side, the cavity under the FBAR ("swimming pool") is defined and filled with a sacrificial oxide, the resonators are processed, followed by the patterning of the Au interconnect and sealing material and, finally, the removal of the sacrificial oxide. The Si micro-cap lid is then Au-diffusion bonded to the FBAR wafer and pads patterned on top. The lid wafer is manufactured with standard micromachining processes and through-wafer vias are etched to make external pad connections. Both the FBAR and lid wafers are high-impedance Si to minimize cross-talk and capacitive losses.

Figure 3A contains a schematic flow of the circuit-containing lid wafer. The process flow proceeds as normal with minor modifications for pattern density. Inter-layer dielectric is left in the field until the devices are complete, after which they are removed down to Si. At that point the lid micro-machining can be completed. The cavity etch that leaves a gap above the FBAR removes any remaining epi-Si between contact pads. The final oscillator die with integrated circuitry is shown in Figure 3B. We retained the Au-Au wafer bond that has proven to provide a robust hermetic seal [6]. The circuitry on the lid aligns with a corresponding depression in the FBAR wafer to provide clearance for wafer bonding. Since the depression is patterned along with the depression under the FBAR, no additional masking steps are required. Once the lid is bonded to the FBAR wafer, the lid Au metallization forms the interconnect between the circuitry and the FBAR. Figure 3C contains a SEM cross-section of a bonded die showing the relative position of the FBAR and the lid circuitry (reversed from the schematic). For this demonstration, our circuits used Avago's HP25 silicon bipolar process.

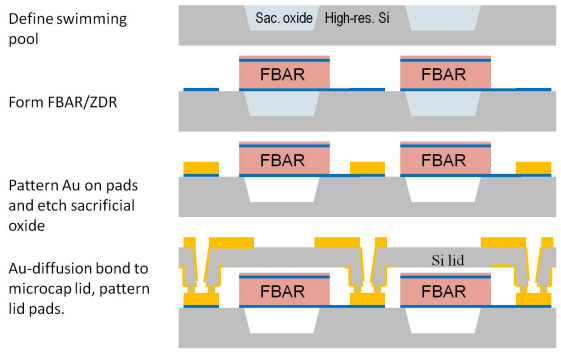


Figure 2. Schematic description of normal FBAR/micro-cap process flow.

III. CIRCUIT IMPLEMENTATION

The concept of integrating active circuitry into the lid has the economic advantage of reusing the lid wafer area for the electronics' pads and from eliminating external connections between the FBAR and the circuit. The size can be further reduced by placing the FBAR directly above the circuitry, which we hope to demonstrate in future iterations. Further economic advantage is gleaned by forcing the circuit size and the resonator (or filter) size to be comparable. The number of external connections is limited by via size; a reasonable via count would be ten or less. In this work, we have six external pads and two internal connections from the lid electronics to the FBAR resonator.

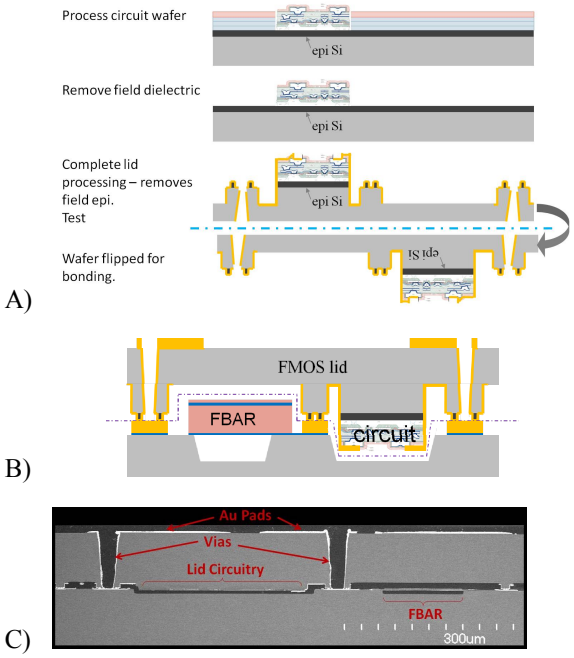


Figure 3. A) Schematic circuit lid flow, B) micro-capped oscillator die, C) cross-section SEM of FBAR/bipolar process showing completed die.

Figure 4A shows the Pierce oscillator and divider that are integrated into the lid. All bias circuitry is integrated and derived internally. The oscillator and bias consume 150µA from a 2V supply. Emitter/Base diode varactors are used for tuning, and a temperature-sensing diode is included for on-die temperature monitoring. The oscillator is AC-coupled to an on-chip divide-by-64 circuit comprising six cascaded stages of divide-by-two bipolar current mode logic (CML). The oscillators operate at 1.5 GHz, yielding a divided clock source of 23.4 MHz. Each divide-by-two circuit includes two cross-coupled D-latches, with the clock input of each latch driven by a level shifter to set the proper DC bias of the input transistors. The maximum operating frequency of the divider-by-two circuit is a function of the D-latch delay and level-shifter delay. The bias currents of the latch and of the level shifter are set accordingly. Because each stage works at half the frequency of the previous one, the bias current of subsequent stages are scaled to reduce power consumption without degrading maximum operating frequency.

A fully-differential Colpitts oscillator was also designed and fabricated in this process. Figure 4B shows the simplified schematic. The Colpitts oscillator has superior cyclostationary noise properties and is thus a good candidate for low jitter reference clocks. However, the Colpitts oscillator requires a higher initial loop gain for oscillator startup. We utilize a gm-boosting technique to reduce power consumption by improving the oscillator startup characteristic. As shown in

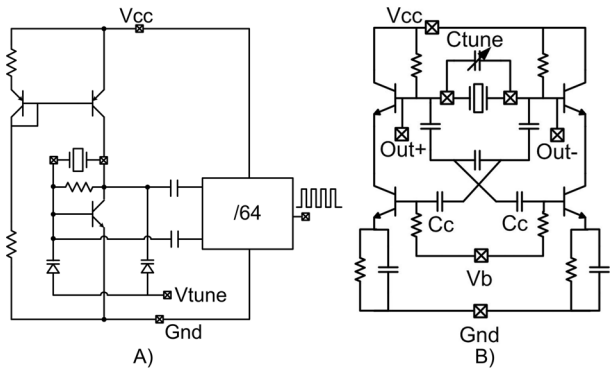


Figure 4. A) Schematic of Pierce oscillator with divider. B) Differential Colpitts oscillator with gm-boosting

Figure 4B, the FBAR resonator is connected between the bases of the top transistor pair. The center point of the FBAR resonator can be viewed as a virtual ground. Differential operation is sustained when the two sides oscillate at opposite phases. By cross-coupling the oscillator output and AC-coupling the signals to the base of the foot transistor pair, the effective loop gain of the oscillator is increased to reduce the bias current requirement. Frequency-dependent emitter degeneration is used to suppress a parasitic mode of oscillation that is not present in typical LC versions of this topology. The differential architecture minimizes common-mode noise introduced power supply and substrate coupling, and thus can further reduce the overall phase noise of the oscillator.

IV. Experimental Results

A. Oscillator Performance

Figure 5 shows the measured transient waveforms from the oscillator at the RF carrier and the divided output. The measured oscillator phase noise at the 1.5 GHz carrier is -118 dBc/Hz @ 100 kHz offset. Although we typically achieve quality factors (Q) of over 2000 for temperature compensated resonators [8], our first devices had a fixable design error limiting the Q of these devices to about 500 to 800. The differential Colpitts oscillator consumes 300µA from a 1.8V supply and provides a phase noise of -124 dBc/Hz @ 100 kHz offset.

These chips integrate temperature compensated FBARs (zero drift resonators, or ZDR) which remove the linear temperature drift around a fixed turn-over temperature (TOT), yielding a residual 2nd order temperature dependence on the order of -15 ppb/˚C^2. The ZDR in these devices have a TOT of approximately -120˚C, which will be moved to 50˚C in subsequent runs. To improve oscillator temperature stability to below 100ppm (over a 100˚C temperature range), a feedback loop measuring temperature is necessary. The on-die temp sensing diode provides a nearly instantaneous read-out of the temperature that closely matches the FBAR temperature. Fig. 6 shows the temperature and frequency as measured by the integrated diode thermometer and a reference sensor placed 10 mm away from the oscillator. A 2W resistive heater on the board was activated for roughly 1 minute to simulate a PA or other heat source. Clearly, the integrated diode will greatly shorten the feedback cycle time and improve temperature compensation accuracy. To test the aging of the wafer-scale packaged resonator, the oscillator frequency shift was measured after 168 hours at 125˚C and again after 336 hours at 125˚C. Less than 1ppm of frequency drift was observed.

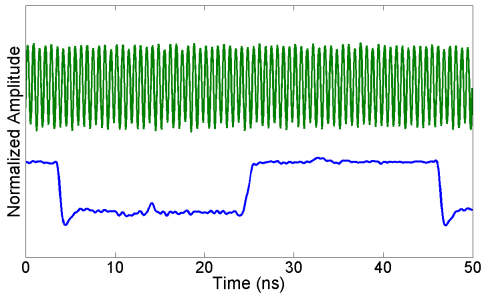


Figure 5. Measured signal at 1.5 GHz RF output (top) and 23.4 MHz divide-by-64 output (bottom)

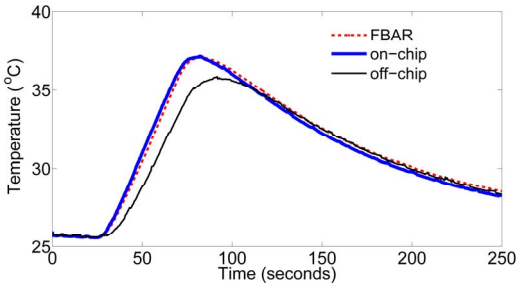


Figure 6. Measured temperature response as measured by the ZDR frequency response, an integrated diode thermometer, and a reference diode placed 10 mm away from the oscillator.

The on-die varactor is used to pull the resonant frequency of the oscillator to compensate for temperature drift as well as any frequency variation due to processing uncertainty. Figure 7 shows the measured pulling range of the oscillator using the integrated varactor (+/- 500 ppm).

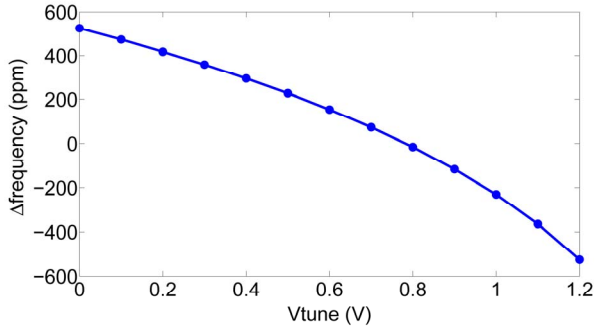


Figure 7. Measured tuning curve of FBAR oscillator.

We used bipolar devices in our first demonstration of this process for three reasons. First, this process has a higher f_T than comparable legacy inexpensive CMOS processes. Secondly, the lower 1/f noise of BJTs provides a significant improvement (6-8dB) in close-in phase noise. Finally, the gm/I_d ratio is 2-3x higher than a MOSFET biased for high f_T.

The properties of this bipolar process thus complement the CMOS circuitry of any accompanying chip.

Different applications of this technology will lead to different optimal partitioning. For example, a high-performance PLL could use the FBAR/bipolar chip as a low noise VCO while the digital functionality were integrated onto a separate CMOS chip. Table 1 is a performance summary of both oscillators implemented in this process. Figure 8 shows the measured phase noise profile of both oscillators.

TABLE I. TABLE 1. DESIGN AND PERFORMANCE SUMMARY

	Pierce	Differential	ISSCC '06 [3]
Integrated FBAR	Yes	Yes	Yes
Hermetic	Yes	Yes	No
Freq (GHz)	1.5	1.5	5.4
Icc (µA)	150	300	1700
Vcc (V)	2.0	1.8	2.7
Tempco (ppm) (0-80˚C)	+/- 200	+/- 200	> +/- 1000*
L(100 Hz) dBc/Hz	-49	-51	-33
L(1 kHz) dBc/Hz	-76	-80	-63
L(10 kHz) dBc/Hz	-98	-104	-93
L(100 kHz) dBc/Hz	-118	-124	-118
*Based on theoretical uncompensated BAW tempco of ~-27ppm/˚C			

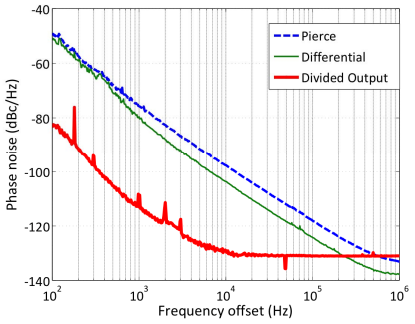


Figure 8. Phase noise of both oscillators measured with an Agilent 5052 signal source analyzer. The 1/f corner for the Pierce is 800 Hz, while the Differential Colpitts is 3 kHz. The phase noise of the ÷64 output is also shown, indicating a 20·log(64) improvement in phase noise. The phase noise floor at 10kHz is the instrument noise floor of the divided output due to a low output signal level. This is not a fundamental limitation.

B. Wafer-level Test

One key benefit of the wafer-scale packaging, as opposed to 2-chip, approach to making FBAR oscillators is the ability to perform wafer-level testing on completed oscillators. Resonator frequency testing with a network analyzer is limited to 5-10 ppm accuracy due to limitations of the measurement and even to parasitic inductance introduced by small variations in probe placement. Oscillator testing reproducibility is <0.2 ppm in initial testing, making it suitable for evaluation of aging effects as well as die screening. Figure 9 shows a map of one of the first wafers made, with the functional ÷64 (23.4 MHz) die shown in green.

V. CONCLUSIONS

We have presented the first circuits implemented in a new single-chip FBAR/Bipolar process. An ultra-low power

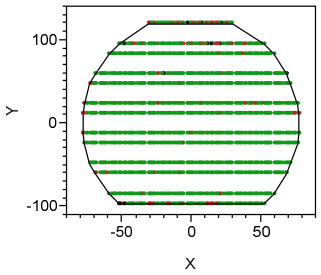


Figure 9. Wafer map showing functional oscillators (in green) from one of the first chip-scale packaged oscillator wafers.

150 µA Pierce oscillator was realized, which achieves a phase noise of -118 dBc/Hz at 100kHz offset. A low noise differential Colpitts oscillator was also implemented, achieving a phase noise of -124 dBc/Hz at 100kHz. The single-chip hermetic package contains all necessary components for a miniaturized frequency reference, including oscillator, high Q resonator, dividers, temperature sensing diode, and frequency tuning varactor.

ACKNOWLEDGMENT

We would like to thank Duane Fasen, Andy Weilert for integrating all of our design rules into reticle services, Mark Dumke and Mark Urfer for their support on the bipolar process, the management team of the Wireless Semiconductor Division of Avago, and Julie Hu for layout/testing assistance.

REFERENCES

[1] Driscoll, M.; Moore, R.; Rosenbaum, J.; Krishnaswamy, S.; Szedon, J.; "Design and Evaluation of UHF Monolithic Film Resonator - Stabilized Oscillators and Bandpass Filters," *Microwave Symposium Digest, 1987 IEEE MTT-S International* , vol.2, no., pp.801-804, May 9 1975-June 11 1987

[2] Khanna, A.P.S.; et al., "A Film Bulk Acoustic Resonator (FBAR) L Band Low Noise Oscillator for Digital Communications," *Microwave Conference, 2002. 32nd European* , vol., no., pp.1-3, 23-26 Sept. 2002

[3] L. Elbrecht et al., "Integration of Bulk Acoustic Wave Filters: Concepts and Trends", International Microwave Symposium, TU-5D, 2004.

[4] M. Dubois et al., "Monolithic Above-IC Resonator Technology for Integrated Architectures in Mobile and Wireless communication", IEEE Journ. Solid-State Circ., v. 41, no. 1, 2006, pp. 7-16.

[5] M. Aissi et al., "A 5.4 GHz 0.35um BiCMOS FBAR Resonator Oscillator in Above-IC Technology", IEEE International Solid-State Circuits Conference (ISSCC), 2006.

[6] R. Aigner, "High performance RF-Filters suitable for above IC integration: Film Bulk-Acoustic-Resonators (FBAR) on Silicon", IEEE Custom Integrated Circuits Conference, CICC, 2003

[7] R. Ruby, et al., "High-Q FBAR filters in a wafer-level chip-scale package," *Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC. 2002 IEEE International* , vol.1, no., pp.184-458 vol.1, 2002

[8] R. Ruby et al, "Ultra-Miniature High-Q Filters and Duplexers using FBAR Technology", IEEE International Solid-State Circuits Conference, 2001, pp. 120-121

[9] W. Pang et al., "A Temperature-Stable Film Bulk Acoustic Wave Oscillator", IEEE Electron Device Letters, April 2008, v. 29, #4,pp. 315-318

[10] S. Rai et al., "A 600 µW BAW-Tuned Quadrature VCO Using Source Degenerated Coupling", IEEE Journal of Solid State Circuits, Vol 43, No. 1, pp.300-305, January 2008.

Low Phase Noise Quartz BAW Oscillator for Space Applications

T. McClelland, J. Zacharski, C. Szekely, E. Mauskop, D. Bogomolov
Frequency Electronics, Inc.
Uniondale, NY/USA
tomm@freqelec.com

Abstract— **A low phase noise, 5 MHz, quartz BAW oscillator has been developed for space applications. Considerable effort has been made to optimize the quartz resonator design used in this oscillator. Both SC and LD resonator cuts have been studied. Although promising results have been achieved with LD cut resonators, the best results have been achieved with 5 MHz, fifth overtone, SC cut resonators. Single sideband phase noise of -133/f ^{3} dBC/Hz for offsets from 0.1 to 2 Hz (equivalent to $\sigma_y(\tau)$ = 7E-14, for τ = 1 sec) has been achieved. This performance is within a factor of two of what is generally accepted as the theoretical phase noise limit for BAW quartz oscillators at this frequency. Details of resonator design, and circuitry are discussed, and phase noise data are presented. Although the results presented have been achieved using commercial electronic parts in the oscillator, the fundamental oscillator architecture is such that it can be implemented using readily available space qualified parts.**

I. INTRODUCTION

The work described here is inspired by a desire to push the envelope of short term frequency stability performance of precision quartz oscillators; with an ultimate goal of implementing this improved performance in quartz oscillators designed for space applications. In order for any results of this effort to be applicable to space applications, it is in general necessary that:

1. All circuit components must be capable of meeting space quality and reliability guidelines (e.g. hermetically sealed packages, space radiation tolerant, etc.)
2. Oscillator and all components must be capable of surviving launch vibration and shock environments

Specific space applications typically impose additional design constraints, however, it is not the purpose of the work described here to meet specific space program requirements, but rather to develop oscillator capability which does not involve any circuit components or mechanical structures which would preclude use in space applications.

With these constraints in mind, the focus of this effort is improvement of short term frequency stability, as defined by the Allan deviation, for averaging times of 0.1 to 100 seconds. Based on the material properties of quartz, the best performance in this domain is expected from oscillators utilizing quartz resonators with the lowest feasible resonant frequency, f_0, since $S_\varphi(f) \propto f_0^4$ ($\sigma_y(\tau) \propto f_0$) [1], [2], [3]. In practice, 5 MHz is the lowest feasible frequency, because lower frequency resonators which can meet the above criteria for space applications are difficult to manufacture. For this reason, the work described here focuses on 5 MHz quartz oscillators and resonators.

More specifically, this work has focused primarily on oscillators using SC cut quartz resonators, given the proven track record and many advantages of such resonators [4]. Accordingly, most of the results shown below are for oscillators employing SC-cut resonators. Recent work, however, has suggested some potential advantage to LD-cut, or Low isochronism Defect, quartz resonators. These resonators are in the same doubly rotated family as the SC cut, but with a somewhat larger φ angle, shown to result in a minimum sensitivity to resonator drive current (amplitude-frequency effect) [5], [6], [7]. Therefore some preliminary effort has been devoted to investigation of oscillators using LD-cut quartz resonators, the results of which are reported below.

For well designed oscillators at this frequency, the quartz resonator in general determines the output noise, and hence the frequency stability. Therefore, this effort has focused mainly on the design of the quartz resonator. Some effort is required to insure that the noise of the sustaining amplifier can be maintained below the level at which it contributes to the oscillator noise within the resonator bandwidth, but the primary effort is to improve the resonator noise performance.

One of the impediments to progress in this arena is the relatively long time required to make meaningful measurements. The introduction of the SC-cut actually improved this situation, since it requires considerably less warm-up/stabilization than comparable performing AT-cut resonators. Nonetheless, given the necessity of tight oven

978-1-61284-110-6/11/$26.00 ©2011 IEEE

control, stabilization times of several hours to several days are required to get meaningful results regarding the underlying performance of these devices out to 100 seconds averaging time. The authors have found that it is generally most efficient to gauge performance of such devices by measuring the phase noise at 1 Hz offset, with attention to the slope of the phase noise vs offset frequency plot.

The frequency stability of the best 5 MHz quartz oscillators, for averaging times of 0.1 to 100 seconds, is determined almost exclusively by 1/f *frequency* noise [3]. This being the case, the *phase* noise is proportional to $1/f^3$, and is related to the Allan deviation, $\sigma_y(\tau)$ (which is independent of averaging time for 1/f *frequency* noise), by [8]:

$$S_\varphi(\mathrm{f}) = \frac{(f_0)^2}{2\,\mathrm{Ln}(2) f^3}\sigma_y^2(\tau\ = 1\ \mathrm{sec}). \qquad (1.)$$

Thus a phase noise plot exhibiting a slope of $1/f^3$ at an offset frequency of 1 Hz represents a fairly reliable indication of underlying performance capability throughout the $\tau = 0.1$ to 100 second range. Since equipment capable of measuring phase noise very efficiently now exists, it is possible in this way to obtain a good indication of oscillator performance in a matter of hours rather than days.

II. Oscillator Design

A block diagram of the basic oscillator used for these studies is shown in Figure 1. The device is designed to allow relatively easy access for removal/replacement of quartz resonators. It utilizes a double oven enclosure for the quartz resonator, with the oscillator circuitry contained within the outer oven. Although the device is not intended for use in space, care has been taken to utilize only oscillator circuit components which are generally acceptable for use in space applications. A photograph of the device is shown in Figure 2.

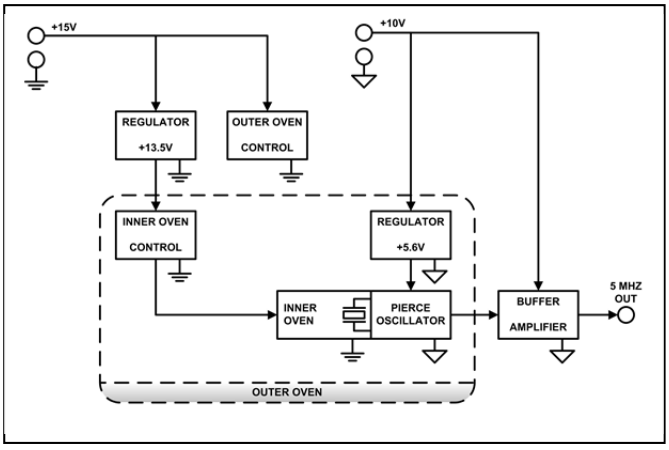


Figure 1. Block diagram of oscillator used for resonator tests

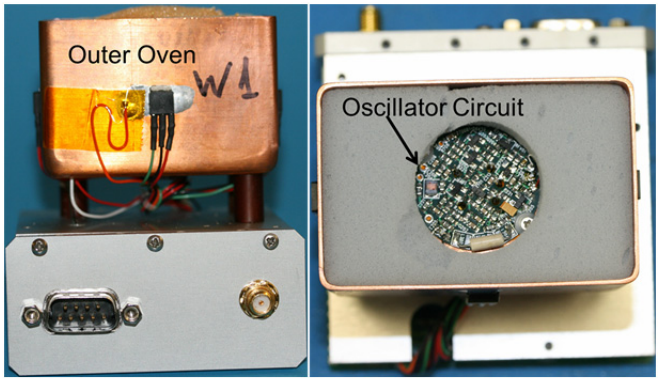


Figure 2. Oscillator used to test quartz resonators

III. SC-Cut Resonators

The work reported here has so far involved the testing of over one hundred 5 MHz, fifth overtone, SC-cut quartz resonators in oscillators like the one shown in Figure 2. These resonators have been made to a variety of specifications in which φ angle, contour, electrode diameter, and electrode metallization are varied as a means of optimizing performance. In order to draw meaningful conclusions from this work it is necessary to accumulate considerable statistics, as inevitably a number of un-controlled variables can obscure the real cause and effect associated with performance observations. This is an on-going effort, with initial results being reported here.

Although no definitive, simple recipe leading to improved noise performance has been identified to date, some useful preliminary conclusions are possible, and performance within 3 dB of what is generally considered to be the material limit of quartz has been observed in some resonators. The manufacturing yield for resonators with this level of performance is low, but reproducible.

A. Activity Dips – "Sweet Spot"

One working hypothesis at the start of this effort was that noise performance can be limited by the presence of activity dips, a problem more closely associated with AT-cut crystals. The idea is that at the oven set point temperature (typically the turn-over temperature or zero-slope of frequency vs. temperature) one or more competing modes may get partially excited in addition to the desired c-mode, and that this excitation is highly temperature dependent in comparison to the dominant c-mode [4]. It is postulated that it could be possible to improve the noise performance significantly in this case by offsetting the temperature slightly; thus strongly attenuating the competing temperature dependent modes, while affecting the dominant mode only slightly. In fact it has been suggested that "sweet spot" temperatures may exist at which the noise is dramatically improved over that observed at the presumed optimum (turn-over) temperature. In some cases it is indeed observed that SC-cut resonator series resistance shows significant variations near the turn-over temperature, and this could well be due to the presence of activity dips. The question is, in cases where little or no measured resistance variations near the operating temperature are observed, can performance be improved by slight changes

in oven operating temperature. Although documentation regarding this is sketchy, claims to this effect do exist, and therefore some effort was devoted to studying this idea.

Typical results are shown in Figure 3. , which shows phase noise at 1 Hz and 10 Hz offsets as a function of quartz resonator oven temperature, over a narrow window of 5°C around the turn-over temperature. Some small variations in phase noise are observed, but these are within the error bars of the measurement. This measurement has been repeated with a number of different resonators. In spite of concerted effort, no evidence of the existence of a "sweet spot" temperature has been found.

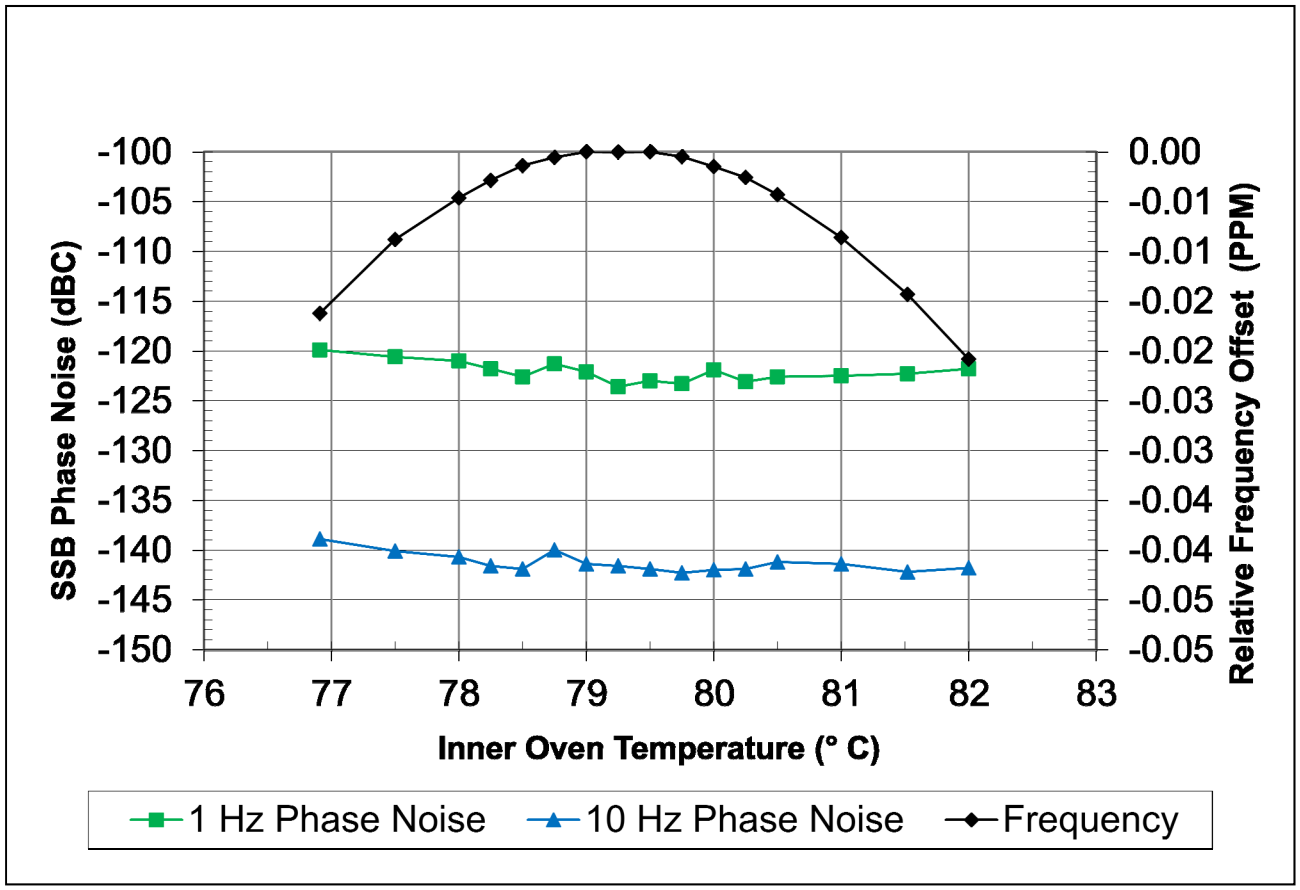


Figure 3. Single side band phase noise measured at 1 and 10 Hz offsets as a function of inner oven temperature near the resonator turn-over temperature

B. Edge Quality

The most consistent result to date has been the correlation between the visual quality of the quartz resonator edge (which is attached to the mounting structure) and noise performance. Those resonators which have edges relatively free from chips and scratches consistently outperform those with more noticeable edge defects.

This was first observed as a result of reprocessing various resonators. Reprocessing resonators is an excellent way to study the effects of different contours and electrode configurations, while keeping the quartz material involved constant. The influence of variations in crystal lattice defects and impurities on the measured noise characteristics is thereby reduced or eliminated.

The critical observation was that with a large number of reprocessed resonators, re-made with a variety of different electrode materials, diameters, and thicknesses, and different contours, including some resonators which were re-made with identical design to their previous configuration; uniformly worse noise performance was observed after reprocessing. A post-mortem of these devices showed that one consistent result was that the quality of the resonator edges had been degraded during the reprocessing activity.

As a result of this observation, resonators were fabricated with greater attention to the quality of the resonator edges. This yielded consistently better performing resonators, largely independent of other variables. Phase noise results from three such resonators are shown in Figure 4. In preparing the figure, pair-wise measurements were made, one comparing each resonator to each of the other two. The results were used to

perform a three-cornered hat calculation of the individual noise contributions from each resonator. The results of this calculation are displayed in Figure 4.

It is evident from the figure that the phase noise of all three resonators is well described by a $1/f^3$ function between 0.1 and 2 Hz; the best measuring -133 dBC/Hz at an offset of f = 1 Hz. From (1), this corresponds to:

$$\sigma_y(\tau) = 7.4 \times 10^{-14} \quad (2.)$$

This is within 3 dB of what is generally presumed to be the material limit for quartz [1], [2], [3].

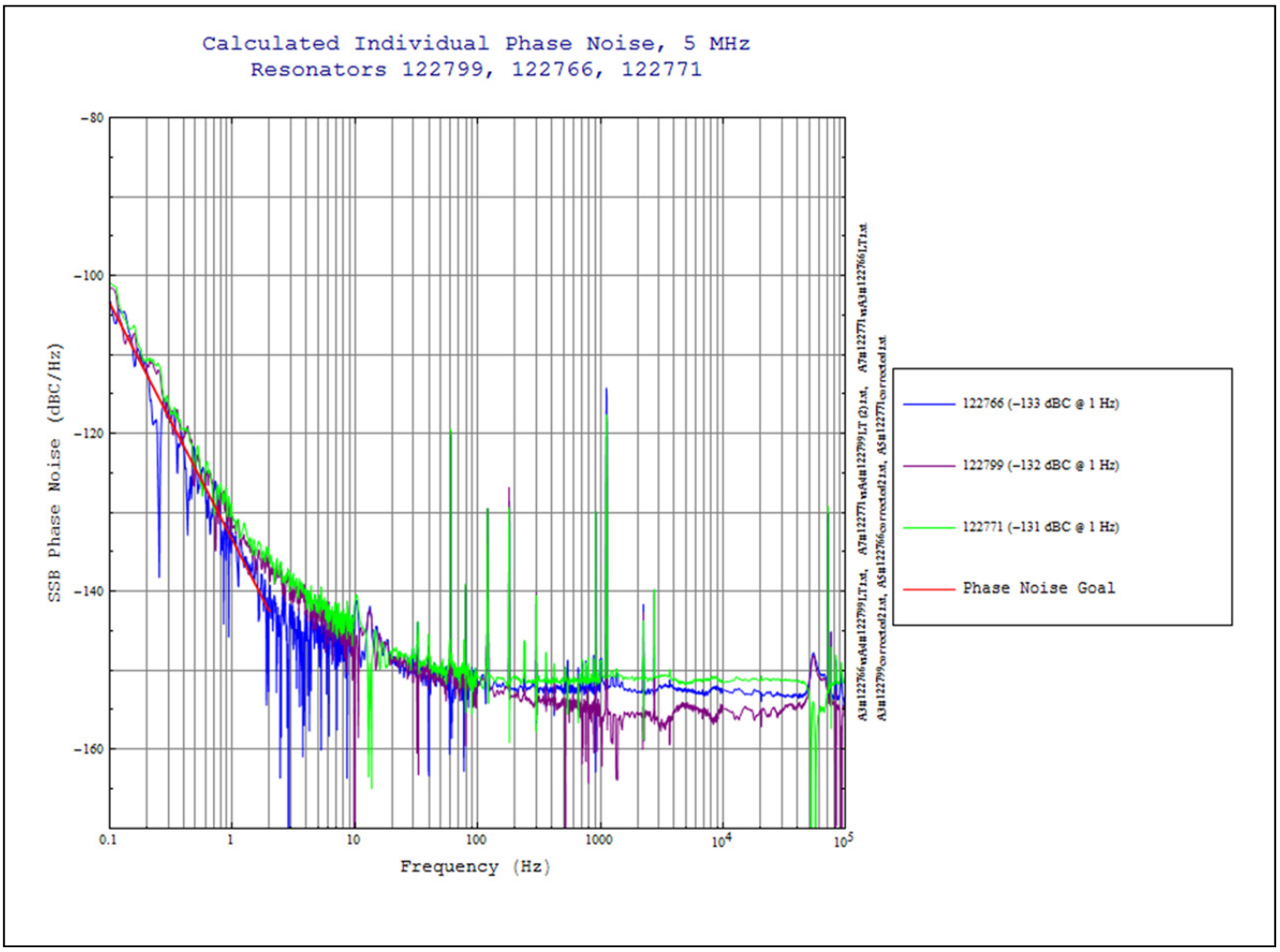


Figure 4. Single sideband phase noise calculated from individual measurements of three 5 MHz, fifth overtone, SC-cut quartz resonators. Displayed results are based on a three-cornered hat calculation of the individual noise contributions of each resonator, derived from pair-wise measurements.

IV. LD-CUT RESONATORS

The noise performance of quartz oscillators *outside* the bandwidth of the resonator is generally determined by thermal noise in the circuit elements [3]. This determines the white noise floor, for large offset frequencies, in the phase noise response. This noise can be reduced by increasing the power, or drive level in the oscillator. Unfortunately, this has the undesired effect of increasing the phase noise inside the resonator bandwidth because of the amplitude-frequency effect. This effect is the observed tendency of the frequency vs power function to exhibit an increasingly steep positive slope as power is increased. The result is that AM (white) noise in the oscillator circuit is converted into PM noise inside the resonator bandwidth.

It was discovered several years ago that for doubly rotated quartz resonators, the slope of the amplitude-frequency effect, as a function of φ angle, changes sign at approximately φ = 27° [5]. Therefore it is expected that resonators manufactured at this angle would not suffer from the undesirable effects described above when driven at higher power levels.

This comes at a price, however, since the stress compensation condition of the SC-cut (φ = 21.95°) is no longer met, and in addition the inflection temperature is increased from ~90°C for the SC cut to ~150°C. Operation

near 150°C is undesirable because the failure rates for semiconductor devices are much higher than for operation near 90°C. LD resonators can easily be manufactured with turn-over temperatures ~90°C, however, because this is further from the inflection temperature, the slope of frequency vs. temperature changes much more rapidly in the vicinity of this temperature than for an SC-cut resonator with the same turn-over temperature [6]. Therefore, oscillators using LD-cut resonators must maintain better temperature control than comparable oscillators using SC-cut resonators.

Nonetheless, the potential for eliminating the undesirable effects of the resonator sensitivity to drive level is worth investigating. To this end, a number of resonators with φ angles in the vicinity of the LD-cut have been manufactured and tested in oscillators. In Figure 5, the amplitude-frequency effect is plotted for several of these resonators, with a typical SC-cut resonator shown for comparison. Although these "LD-cut" resonators do not exhibit zero slope, they clearly exhibit reduced slope compared to SC-cut resonators.

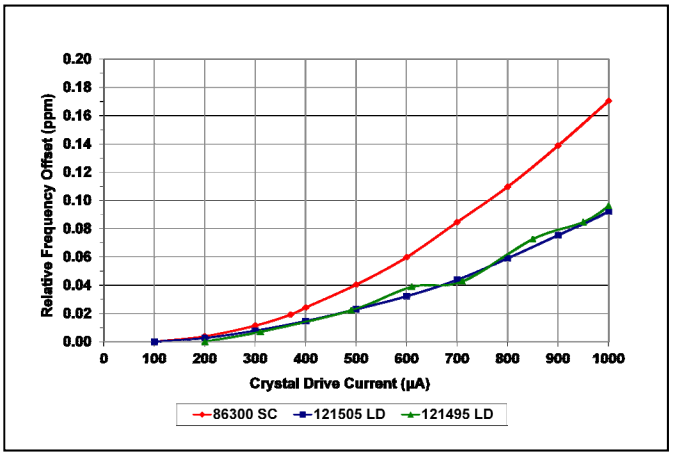


Figure 5. Amplitude-Frequency effect for SC-cut and LD-cut resonators. Slope for LD-cut resonators is significantly less than for SC-cut resonators.

In Figure 6, the phase noise at 1 Hz and 10 kHz offsets are plotted for one of these resonators as a function of resonator drive level. As expected, the noise floor, indicated by the phase noise at 10 kHz offset, improves at higher drive levels. Yet the close in phase noise, indicated by the phase noise at 1 Hz offset, is not appreciably degraded at these higher drive levels. This is the sought after reduction of the amplitude-frequency effect.

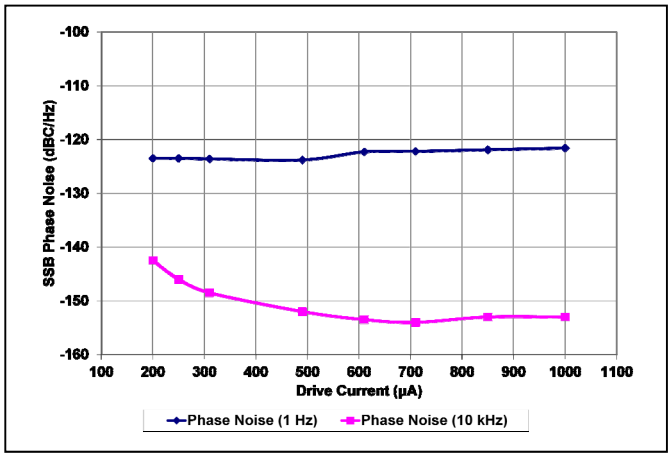


Figure 6. Phase Noise at 1 Hz and 10 kHz offsets as a function of drive current for an LD-cut resonator.

V. SUMMARY

An on-going effort, targeted at satellite applications, to improve frequency stability performance of precision 5 MHz quartz oscillators has been described. This effort has focused on the quartz resonator as the main performance limiting factor. Performance within 3 dB of the theoretical material limit of quartz has been obtained with several SC-cut resonators, where particular attention was paid to the quality of the resonator edges.

Preliminary attempts have also been made to manufacture and test LD-cut resonators. The results of these preliminary tests are promising, but further work is necessary in order to optimize the trade-off between stress compensation, temperature sensitivity, and frequency dependence on drive current.

REFERENCES

[1] J.R. Vig, and F.L. Walls, "Fundamental limits on the frequency instabilities of quartz crystal oscillators," Proceedings, IEEE International Frequency Control Symposium, 1994, p. 506.

[2] J.J. Gagnepain, J. Uebersfeld, and G. Goujon, "Relation between 1/f and Q-factor in quartz resonators at room and low temperatures, first theoretical interpretation," Proceedings, 35th Annual Frequency Control Symposium, 1981, p. 476.

[3] T.E. Parker, "Characteristics and sources of phase noise in stable oscillators," Proceedings, 41st Annual Frequency Control Symposium, 1987, p. 99.

[4] M. Bloch, J. Ho, C. Stone, A. Syed, and F. Walls, "Stability of high quality quartz crystal oscillators: an update," 43rd Annual Frequency Control Symposium, 1989, p. 80

[5] N. Gufflet, F. Sthal, J.J. Boy, R. Bourquin, and M. Mourey, "Doubly rotated quartz resonators with a low amplitude-frequency effect," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 48, no. 6, November, 2001, p. 1681.

[6] S. Galliou, F. Sthal, J.J. Boy, R. Bourquin, M. Mourey, "Recent results on quartz crystal LD-cuts operating in oscillators," Proceedings, IEEE Frequency Control Symposium, 2004, p. 475.

[7] M. M. Driscoll, "Oscillator AM-to-FM noise conversion due to the dynamic frequency-drive sensitivity of the crystal resonator," Proceedings, IEEE Frequency Control Symposium, 2008, p. 672.

[8] D.A. Howe, D.W. Allan, and J.A. Barnes, "Properties of signal sources and measurement methods," Proceedings of the 35th Annual Frequency Control Symposium, 1981, p 1.

Study of Fe^{3+}-Sapphire Maser above 4K

Karim Benmessai[1], Mohamad Mrad[2], Daniel Lloyd Creedon[1], Jean-Michelle Le Floch[1], Michael Edmund Tobar[1] , Pierre-Yves Bourgeois[2], Yann Kersalé[2], Vincent Giordano[2]

[1] Frequency Standards and Metrology, School of Physics, The University of Western Australia
Perth, WA, Australia

[2] Departement Temps Fréquence, Femto-st Institut
University of Franche Comté
BESANCON, FRANCE

Karim.benmessai@uwa.edu.au

Abstract— **The Whispering Gallery Mode Maser Oscillator (WHIGMO) uses paramagnetic Fe^{3+} ions in HEMEX sapphire at low temperature, and has been described in many previous publications [1-7]. The fractional frequency instability has been demonstrated to be as low as $\sigma_y(1s<\tau<100s) = 10^{-14}$. This instability is normally measured at the frequency-temperature turnover point of the sapphire where the effects of temperature fluctuations on frequency are nullified to first order [8-9]. Usually the temperature is determined by the relative concentrations of several paramagnetic ions such as Cr^{3+}, Mo^{3+}, and Ti^{3+}. In this paper, we show that Fe^{3+} ions have a strong effect on the turnover point. We compare the behavior of two crystals – "Pinocchio" (concentration of active Fe^{3+} ions = 10ppb) and "Geppetto" (100ppb). We also report the observation of an upper limit for maser operation above 30K. This limit is explained using a Boltzmann distribution for the active ion, and we show that the populations of the energy levels are so close at this temperature that the effect of the pump at 31 GHz used to create the maser at 12 GHz is ineffective.**

I. INTRODUCTION

To build an ultra-stable oscillator in the X-band, it is desirable to use crystals with low loss and high *Q*-factor. The best material for this purpose is HEMEX grade sapphire, which demonstrates *Q*-factors of order 10^9 when cooled to 4.2 K. The resonant frequency of the Whispering Gallery mode of interest is fixed when operated at a particular temperature, the 'turnover point', where temperature dependent frequency fluctuations are cancelled to first order. Many publications describe the application of this self-compensation technique. [8-17], and all show that the turnover temperature is dependent on the presence and relative concentrations of paramagnetic ions such as Cr^{3+} with an Electron Spin Resonance (ESR) at 11.4 GHz, Mo^{3+} (100 GHz) and Ti^{3+} (1 THz). The ions are substitutionally included in the crystal lattice and for purest sapphire result from the crystal growing process and occur unintentionally. We show in this paper how we can control the temperature annulment point of the Fe^{3+}-WHIGMO by manipulating the amount of ions involved in the maser process.

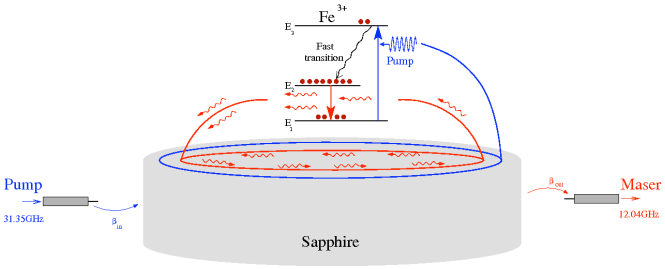


Figure 1. Principle of operation: We apply a signal at 31.35 GHz, coinciding with a WG mode, to pump the ions from the lowest energy level |1/2> to the third |5/2>. The ions then (non-radiatively) relax to the second energy level |3/2>. A population inversion is obtained between the two lowest energy levels, and stimulated emission can be obtained at 12.04GHz, which is enhanced by another WG mode coincident in frequency.

II. MASER DESCRIPTION

The maser scheme is based on the three energy levels of the Fe^{3+} ions at zero applied DC magnetic field in a WG sapphire resonator (Figure 1). The concentration of ions used to create the effect is about 10-100ppb, whereas the total concentration is on the order of 1ppm. The bandwidth of the

karim.benmessai@uwa.edu.au

ESR is $\Delta\nu_{Fe^{3+}} \approx 100$MHz, where the WG mode width is only $\Delta\nu_{WG} \approx 10$Hz at 4.2K.

The maser signal frequency ν_{op} is fixed at the frequency ν_{WG} of the WG mode involved in the process. It's well known that the frequency ν_{WG} is strongly sensitive to temperature variations and needs to be precisely controlled. In this kind of resonator, ν_{WG} possesses a temperature turnover point above 4.2 K. The frequency variations can be described by [8]:

$$\frac{\nu_{WG} - \nu_0}{\nu_0} = AT^4 + C(\mathrm{T}) \qquad (1)$$

where ν_0 is the mode frequency at 0 K without ions, and AT^4 is the thermal sensitivity due to the sapphire dilation ($A \approx 10^{-12}\mathrm{K}^{-4}$ for our modes). The other term $C(\mathrm{T})$ is thermal sensitivity due to the paramagnetic ions and is defined by (2):

$$C(\mathrm{T}) = \frac{1}{2}\sum_i \eta_i(\nu)\chi_0 \frac{\left(2\pi\tau_2^{(i)}\right)^2 \nu_i(\nu_i - \nu)}{1 + \left(2\pi\tau_2^{(i)}\right)^2 (\nu_i - \nu)^2} \qquad (2)$$

where η is the magnetic filling factor of the mode describing its magnetic field distribution compared to paramagnetic ions in the crystal, and χ_0 is the DC magnetic susceptibility of the ion.

III. Measurements

To characterize the maser signal we amplify it first (30dB) and then compare it to a signal from a synthesizer referenced to a hydrogen maser. The beat note is then amplified and band pass filtered before being counted (Figure 2). The accuracy of the temperature control at resonator cavity is about 2 mK.

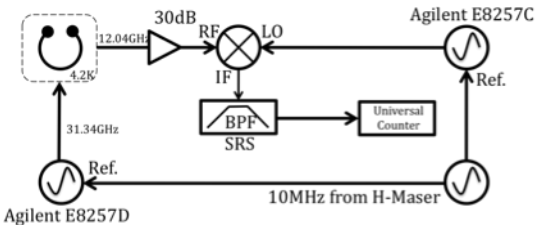


Figure 2. Experimental setup: The crystal is a cylindrical HEMEX sapphire, 30mm height and 50mm diameter, placed in a cylindrical cavity and cooled to 4.2K using a cryocooler .

A. "Pinocchio" behaviour

The maser signal characterized in this section is oscillating at the $WGH_{17,0,0}$ mode frequency of 12.038 GHz, and can be excited using a pump signal corresponding to any of several WG modes around 31.3 GHz. The frequency of the WG mode which enhances the maser signal is measured with a network analyzer and shows a turn over temperature T_{inv} around 7.8 K, with no dependence on the input power level applied to the crystal.

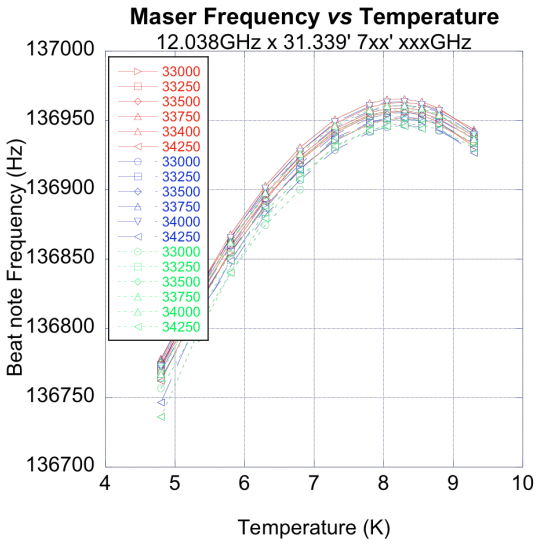


Figure 3. Evolution of the maser frequency (for "Pinocchio") at 12.037 GHz with temperature for different pump power levels. The turnover temperature stays constant around 8.3K The red curves are for a pum power at 17dBm, the blue at 16dBm and the green at 15dBm.

We vary the frequency of the pump signal used in the process, and observe the evolution of the maser frequency with temperature. For each pump we change its frequency in steps of 250 Hz and at varying power levels. As Figure 3 shows, T_{inv} is completely independent of the input power level, and the value of the pump frequency T_{inv} stays sensitively around 8.3K.

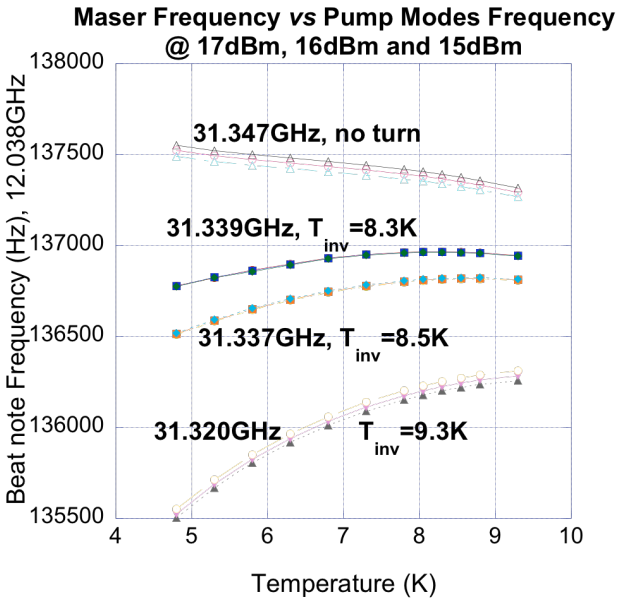


Figure 4. Evolution of the maser signal frequency with temperature for different pumps at 31.3xxGHz and different input power levels

Despite the fact that T_{inv} is independent of the input power and frequency of the pump signal (for a particular choice of pump mode), its value does change when a different pump WG mode is selected, as shown in Figure 3. The turnover point is inversely proportional to the pump frequency in the range 31.320–31.339 GHz. No turnover point is observed at higher frequencies. Note that for each pump mode, T_{inv} stays independent of power. The turnover point can be calculated for each pump by a simple second order polynomial equation.

B. "Geppetto" behaviour

For the second crystal, the maser operation is different. The concentration of the active ion is higher, resulting in a higher maser output power of –40dBm, compared to –56dBm for *Pinnochio*. In addition, maser signals are observed not only at the $WGH_{17,0,0}$ frequency, but also at different modes at frequencies within the ESR bandwidth[6]. Like *Pinnochio*, all the pump modes for which a maser signal can be excited in *Geppetto* have a temperature turnover point independent of the applied power.

Characterizing the turnover point for this crystal reveals more complex effects than *Pinocchio*. Figure 5 shows the signal at 12.037 GHz oscillating for a pump applied at 31.349 GHz. It represents the maser frequency dependence for pump frequencies in the range 31.349,782,015 – 31.349,782,715 GHz in steps of 100 Hz. Here, the maser frequency shows a strong dependence on the pump tuning. We note that when the frequency is in the range 31.349,782,115 – 31.349,782,315 GHz, there are three T_{inv} – one near 6 K, the second around 8 K and the third at approximately 10.5 K. For the range 31.349,782,415 – 31.349,782,715 GHz, there is only one T_{inv} around 8 K and the system behaves in the same classical fashion as the first crystal.

The behaviour of another maser signal oscillating simultaneously at 12.014 GHz when pumped at 31.310 GHz is shown in Figure 6. We again observe a strong dependence of T_{inv} on the pump frequency. When the pump frequency is swept in the range 31.310,667,885 – 31.310,678,485 GHz, we see T_{inv} moving from 9K to 8.7K and subsequently disappearing. The maser frequency also shows that more than one T_{inv} exists depending on the frequency tuning of the pump. For example, at 31.310,677,985 GHz the maser has two turnover points (T_{inv1} = 9 K, T_{inv2} = 11 K), at 31.310,678,185 GHz it has three (T_{inv1} = 8.7 K, T_{inv2} = 10.1 K and T_{inv3} = 11 K), and at 31.310,678,285 GHz there is no T_{inv} observed at all.

The temperature turnover point depends on the distribution of the ions and the population difference between their energy levels. When we measure T_{inv} for the $WGH_{17,0,0}$ mode using a network analyzer, no change is seen due to the small population difference at this frequency (5% of the total active population). However, when the maser is operating, i.e. being pumped, the configuration of the ions is completely different: population inversion exists between the |1/2> and |3/2> levels, and saturation is seen at 31 GHz where the population of the states |1/2> and |5/2> are equal. The population inversion rate is determined by the choice of pump mode near 31 GHz. For each pump, we can have strong or weak maser signal depending on its coupling and spatial field distribution in the crystal. Strong coupling allows an optimum exchange of energy with the ions, and field distributions occupying more volume in the crystal increase the interactions with different classes of ions. Considering the inhomogeneous broadening of the line at 31 GHz in addition to these factors, we can understand why T_{inv} is different from one pump to another. The Fe^{3+} spins act as individual packets selected by the pump WG mode corresponding to the applied pump signal. The pump signal selects a different quantity of ions depending on their ESR frequency.

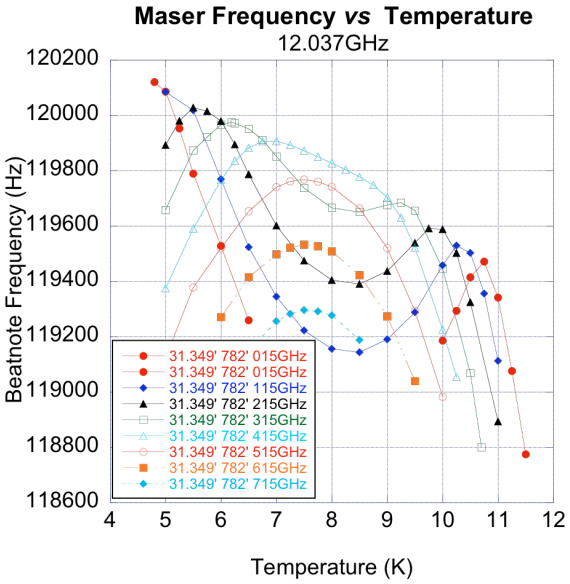


Figure 5. Evolution of the maser signal frequency at 12.037GHz with temperature for different value of the same pump at 31.349GHz

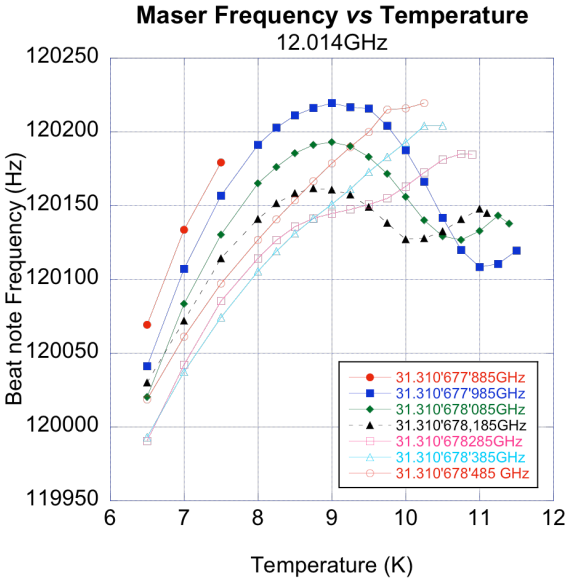


Figure 6. Evolution of the maser signal frequency at 12.014GHz with temperature for different value of the same pump at 31.310GHz

For *Geppetto* the case is more complex because T_{inv} is not only different from one pump to another for the same maser

signal, but is different for the same pump mode when the frequency is tuned. The maser frequency shows more than one temperature turnover point for some pump frequencies, and no inversion for others.

The characterization of the maser frequency with temperature was performed at a fixed pump frequency. However, the pump WG mode is temperature dependent as well as the WG modes near 12.04 GHz. To optimize the maser operation, it is necessary to tune the pump frequency for each operating temperature. This optimum frequency corresponds to maxima or minima of the maser frequency depending on the T_{inv} being measured. Figure 7 shows how the maser frequency varies with temperature for different pumps modes.

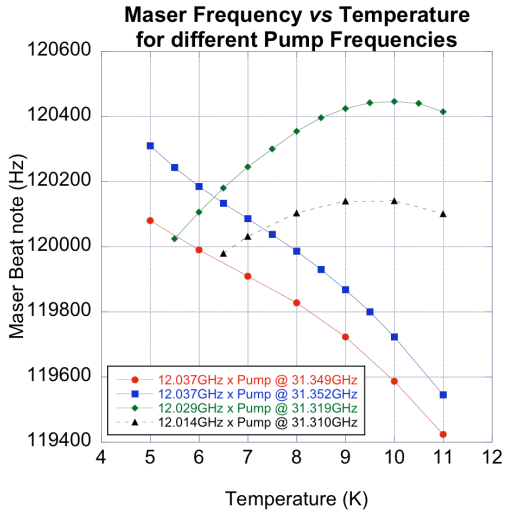


Figure 7. Evolution of the maser signal frequency with temperature for different pumps.

The signal at 12.14 GHz has a turnover point at 10K, the one at 12.029 GHz is at 10.3K, and the one at 12.037 GHz shows no turnover temperature at all. The measurements were performed at different power levels, which were found to have no effect on the turnover temperature.

IV. Maser Temperaure Operation Limit

In order to characterize the upper temperature limit for maser operation, we adjust the pump frequency for each measurement as per the last section. The dependence of the maser frequency is shown in Figure 8. The 12.037 GHz maser signal oscillates until 28 K when pumped at 31.349 GHz, and 30K when pumped at 31.352 GHz. The power of the maser signal is observed to decrease as the temperature is raised. This is due to the fact that, at thermal equilibrium, the populations of the Fe^{3+} energy levels follow a Boltzmann distribution. At higher temperatures, the difference in population between levels is reduced, resulting in less efficient masing.

At high temperature (20-30K) the number of ions pumped at 31.3 GHz is lower than at 4.2 K, leading to a smaller population inversion ratio. This directly affects the maser power and causes it to drop as temperature rises. In addition, the absorption at 31.3 GHz is a result of a mixing between the |1/2> and |5/2> ion states and occurs with very small probability [18-20]. The population difference at 31 GHz is high enough to create the maser signal with this probability at low temperatures (<30K), but not at high temperatures (>30K).

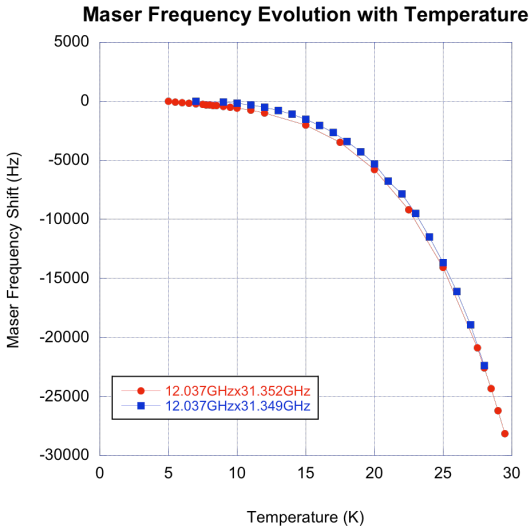


Figure 8. Evolution of the maser signal frequency with temperature for different pumps. The pump frequency is adjusted for each temperature in sort of having an extrema on the maser frequency.

At higher temperatures still, characterization of the WG modes around 12.04 GHz showed absorption effects up to 140 K. This strength of the absorption decreases when we increase the temperature. However no maser operation has been observed at these temperatures.

Conclusion

We have demonstrated how Fe^{3+} ions in a sapphire maser influence the frequency-temperature turnover point of the maser signal. We compared the behavior of two crystals, one with a low ionic concentration and the other with a high concentration. The first crystal showed classical behavior where T_{inv} is independent of the pump power and frequency, but was different between pump modes due to different packets of ions being accessed for each pump. The ions that create the maser are affecting a change in the intrinsic mode T_{inv}. The crystal with higher Fe^{3+} concentration showed a more complex behavior where T_{inv} was different not only from one pump mode to another, but showed many turnover points for some choices of pump, and in some cases none at all.

REFERENCES

[1] P.-Y. Bourgeois, M. Oxborrow, N. Bazin, Y. Kersalé, and V. Giordano, "Observation of a bistability effect in a cryogenic whispering gallery mode resonator," in *Proc. of 19th European Frequency and Time Forum*, Besançon, France, 2005, pp. 141-144.

[2] Bourgeois P-Y, Bazin N, Kersalé, Giordano V, Tobar M E, and Oxborrow, "Maser oscillation in a whispering gallery mode microwave resonator" Appl. Phys. Lett. 87 224104, 2005

[3] Bourgeois P-Y, Oxborrow M, Tobar M E, Bazin N, Kersal and Giordano, "Maser oscillation from electronic spin resonance in a cryogenic sapphire frequency standard", Int. J. Mod. Phys. B 20 1606, 2006

[4] Benmessai K, Bourgeois P-Y, Kersalé, Bazin N, Tobar M E, Hartnett J G, Oxborrow M and Giordano, "Frequency instability measurement system of cryogenic maser oscillator", Electron. Lett. 43 1436, 2005

[5] Benmessai K, Creedon D L, Tobar M E, Bourgeois P-Y, Kersalé and Giordano "Measurement of the fundamental thermal noise limit in a cryogenic sapphire frequency standard using bimodal maser oscillations", Phys.Rev. Lett. 100 233901, 2008

[6] D. L. Creedon, K. Benmessaï, M. E. Tobar *et al*, "High power solid-state sapphire Whispering Gallery mode maser" IEEE Trans. Ultrason., Ferroelectr., Freq. Control, 2009, Vol. 57, No. 3, pp. 282 - 285

[7] K. Benmessaï, D. L. Creedon, M. E. Tobar, P.-Y. Bourgeois, Y. Kersalé, and V. Giordano, "Measurement of the Fundamental Thermal Noise Limit in a Cryogenic Sapphire Frequency Standard Using Bimodal Maser Oscillations," *Phy. Rev. Lett.,* vol. 100, p. 233901, 2008.

[8] R P Kovacich, A G Mann and D G Blair, "Magnetic field tuning of paramagnetic frequency - temperature compensation in cryogenic sapphire dielectric microwave resonators", J. Phys. D: Appl. Phys. 30 3146–3152, 1997

[9] Mann A G, Giles A J, Blair D G and Buckingham M J, "Ultra-stable cryogenic sapphire dielectric resonators: mode frequency–temperature compensation by residual paramagnetic impurities", J. Phys. D: Appl. Phys. 25 1105–9, 1991

[10] John G. Hartnett, Michael E. Tobar, Jean-Michel Le Floch, Jerzy Krupka, Pierre-Yves Bourgeois, "Anisotropic paramagnetic susceptibility of crystalline ruby at cryogenic temperatures", PHYSICAL REVIEW B 75, 024415, 2007

[11] G. J. Dick et al., "Cryo-cooled sapphire oscillator with ultra-high stability", In Proceedings of the IEEE International Frequency Control Symposium and Exposition, pages 528–533, 1998

[12] G. J. Dick et al., "Cryo-cooled sapphire oscillator with ultra-high stability", In Proceedings of the IEEE International Frequency Control Symposium and Exposition, pages 528–533, 1998

[13] G. J. Dick et al., "Temperature-compensated sapphire resonator for ultra-stable oscillator capability at temperatures above 77 K", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency *Control*, **42**(5), 812–819, 1995

[14] R. T. Wang et al., "Cryo-cooled sapphire oscillator with mechanical compensation", *IEEE Interna*tional Frequency Control Symposium, pages 543–547, 2002

[15] Y. Kersalé et al., "Titanium doped cryogenic sapphire resonator oscillators", *European Frequency and Time Forum*, pages 368–372, 2005

[16] N. Boubekeur et al., "Frequency stability of Ti^{3+}-doped whispering gallery mode sapphire resonator oscillator at 34 K", *Electronics Letters*, **41**(9), 534–535, 2005

[17] M. E. Tobar et al., "Design of high-Q frequency-temperature compensated dielectric resonators", *Electronics Letters*, **35**(4), 303–305, 1999

[18] H. F. Symmons and G. S. Bogle, "On the Exactness of the Spin-Hamiltonian Description of Fe3+ in Sapphire," *Proc. Phys. Soc.,* vol. 79, pp. 468-472, 1962.

[19] [25] L. S. Kornienko and A. M. Prokhorov, "Electronic Paramagnetic Resonance of the Fe3+ Ion in Corundum," *Sov. Phys. JETP,* vol. 13, pp. 1120-1125, 1961.

[20] [26] G. S. Bogle and H. F. Symmons, "Paramagnetic Resonance of Fe3+ in Sapphire at Low Temperatures," *Proc. Phys. Soc.,* vol. 73, pp. 531-532, 1959.

Space Mini Passive Hydrogen Maser

- A Compact Passive Hydrogen Maser for Space Applications -

Marco Belloni, Marina Gioia, Simone Beretta
Selex Galileo
Milan, Italy
marco.belloni@selexgalielo.com

Pierre Waller
European Space Agency (ESTEC)
Nordwijk, Netherlands

Giovanni Busca
KyTime
Neuchâtel, Switerland

Fabien Droz, Pierre Mosset
SpectraTime
Neuchâtel, Switerland
mail@spectratime.com

Abstract— On-board atomic clocks represent the key technology for the success of any satellite navigation system mission and their development has been continuously supported by ESA. The PHM (Passive Hydrogen Maser) has been selected as master clock of the Galileo Navigation Payload and the almost three years of continuous observation on board of GIOVE-B satellite have confirmed the outstanding performance in terms of frequency stability and negligible drift [1]. The results coming from ground life testing have given important feedbacks on the PHM technology capability to comply with the required 12 years of lifetime. The above was the starting point for the development of a Miniaturized PHM (mPHM) carried out in the frame of the ESA European GNSS Evolutions Programme. The main target was to preserve the excellent PHM frequency stability performance with reduction of the overall mass and instrument sensitivity to the environmental conditions (in particular to temperature). This paper presents an overview of the main results achieved by the design, manufacturing and testing of a MPHM elegant breadboard (EBB); the 30% of mass reduction already achieved on the Physic Package gives the concrete possibility to have a complete instrument with a mass of 12Kg.

I. INTRODUCTION

The PHM is the best performing on-board clock ever produced for the navigation market. However the constraints associated to its state of the art performances in terms of mass and power consumption, represent a critical point to be investigated and possibly reduced in order to meet the requirements for the next generation navigation systems. Over the last year, various innovative techniques and technologies, especially for the Physic Package (65% of the PHM mass), have been identified for significant improvement of mass, volume and power consumption without affecting electrical performance, in particular frequency stability. In the frame of the ESA European GNSS Evolution Programme, several modifications have been introduced, with benefits in terms of mass and volume. In addition to Physic Package miniaturization, also a considerable dc power saving has been achieved, considering that almost the 40% of the PHM power consumption is presently used for heating and thermal regulation of some critical parts (H2 pressure, microwave cavity assembly, RF sensitive sub-modules). In parallel some activities have been focused for correction and improvement of some performance of the PHM present design. The excellent frequency stability demanded to the passive hydrogen maser makes it very sensitive to environmental parameters, first of them temperature, which leads to very severe operating constraints. The reduction of such sensitivity would reduce also the constraints, leading to the simplification of the navigation payload thermal design.

II. PHYSIC PACKAGE MINIATURISATION DESCRIPTION

The actual design, as developed in the frame of Galileo Programme, was the starting point of the activities (Figure 1). The overall Physic Package (PP) performs the function of a very narrow-band microwave amplifier and frequency discriminator, featuring an extremely stable central frequency. Main items composing the PP are the Microwave Cavity and Shield Assembly (MCSA), that includes the storage bulb, the microwave cavity, its magnetic shielding and the thermal control devices. A low vacuum enclosure is built over the MCSA to simulate actual in space thermal exchange conditions. Other fundamental items are the Hydrogen Beam Assembly, that includes the hydrogen purifier, the dissociator bulb, the beam collimator and the state selector. The hydrogen supply is assumed by the HSA (Hydrogen Storage Assembly), that is composed by a solid-state hydrogen storage vessel, with relevant pressure and temperature devices. The high vacuum is assured by the ion pump and the getters.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

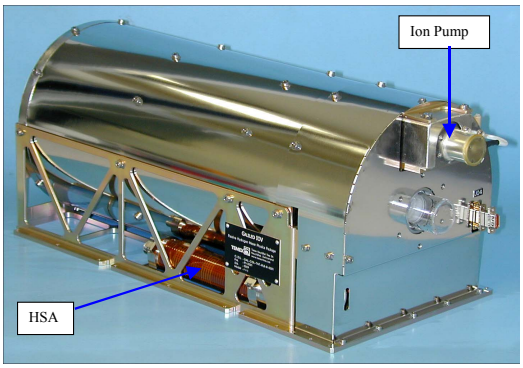


Figure 1. Galileo PHM Physic Package

The goal of the PHM improvement for the next Galileo generation was to reduce the mass by 30%, without affecting its frequency stability performance. As the core of the Physics Package, the microwave cavity is the key component for the mass, size and the electrical performance of the whole PP. The magnetron cavity structure has been maintained in order to preserve a correct distribution of the electric field. However several modifications have been simulated in order to achieve the best compromise between size reduction and quality factor.

The final solution has allowed a volume reduction of the cavity close to 30%; all structures enclosed the cavity such as magnetic shields, low vacuum chamber and mounting structure have been reduced by a similar ratio.

Other area of improvements have been explored and different solutions adopted. They can be resumed as following:

- Removal of the third thermal regulated stage
- Improvement of the magnetic field homogeneity
- Hydrogen tank thermal control simplification
- Getter high vacuum chamber redesign

The initial goal of 30% of mass reduction has been finally achieved (Figure 2), without affecting the storage bulb dimensions and hydrogen flux. This was, since the beginning, the necessary condition in order to minimize the risk of signal to noise ratio reduction. Of course degradation of the atomic gain quality factor was expected, as effect of the microwave cavity losses increasing. This has not been considered a limitation factor to achieve the frequency stability requirements, considering the margin existing in the PHM Physic Package and improvements planned on the RF receiver.

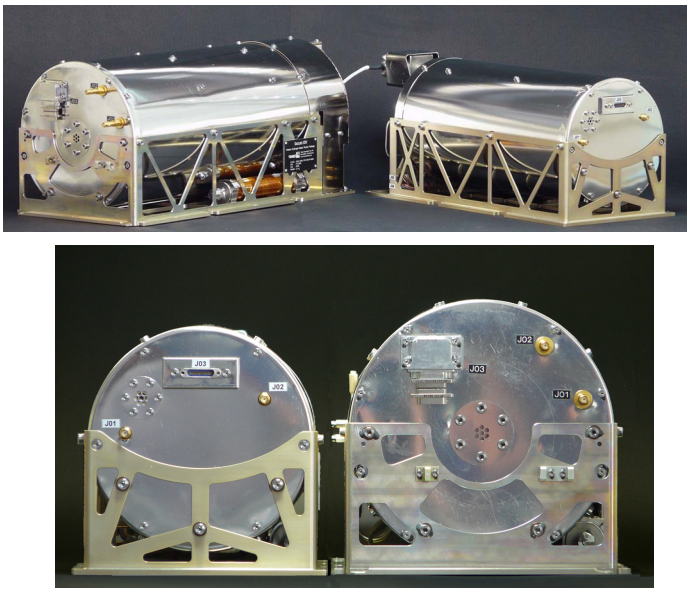


Figure 2. mPP vs PP comparison

III. ELECTRONIC PACKAGE

The LNA and Interrogation sub-module of the EP have been recognized as those with significant margin for improvement from an electrical standpoint and therefore to be submitted to an optimization activity, in the frame of the development phase. No mass reduction has been foreseen for the Elegant Bread Board (EBB) phase, since the electronic package gives a minor contribution to total mass. In addition it was preferred to proceed step by step, without modifying the electrical and mechanical interfaces with the PP.

Major objectives of the RF interrogation chain modification have been the following:

- Improvement of the phase modulation symmetry
- Reduction of the modulator thermal sensitivity
- Reduction of the phase noise contribution due to the multiplication chain (flicker PM in particular)

The first objective was related to the possibility to reduce the instrument atomic gain variation sensitivity. A parallel study has been carried out by KyTime in order to provide an assessment of this aspect. The outcomes have been translated into requirement specifications, which have been fully achieved with the adoption of a new design of the phase modulator.
Result of the redesign is a reduction of a factor 10 of instrument frequency sensitivity to the atomic gain variation.

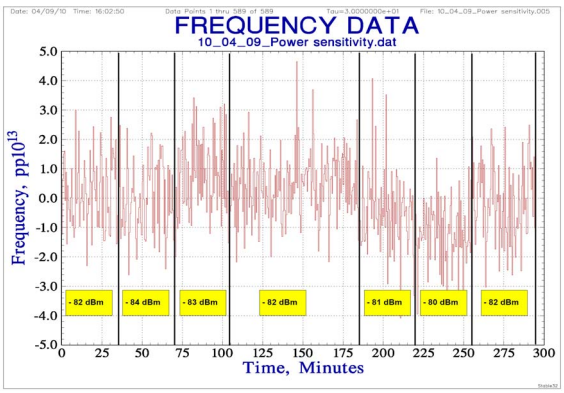


Figure 3. Frequency sensitivity to interrogation power with the new phase modulator

Figure 3 shows the results of a test giving the indirect verification of the sensitivity. The test consists in changing the power level of the interrogation signal, which simulates a variation of the atomic gain. The typical PHM sensitivity value is around $2*10^{-13}$/dB, now reduced to few parts per 10^{-14}.

The remaining two objectives (thermal sensitivity, phase noise) have been achieved reaching the reduction of a factor of 10 of the modulator thermal sensitivity ($<1*10^{-14}$/°C without temperature compensation) and improvement of more than 10 dB on the added phase noise.

Activity has been also devoted to the receiver noise figure (NF) improvement. Moreover matching technique has been adopted to have a tunable coupling between microwave cavity and RF receiver. The final result is the availability of an additional degree of freedom in order to obtain the best compromise between reduction of the white frequency noise (due to receiver NF) and minimization of the cavity pulling due to receiver load variation.

Finally a general review of the Electronics Package structure has been carried out for mass reduction and optimization/adaptation of the Hydrogen Dissociator Oscillator (HDO) housing to the new Mini PP interface. Tailoring of the current thermal control circuits to the new thermal interface requirements of the miniaturized physics package has been implemented as well.

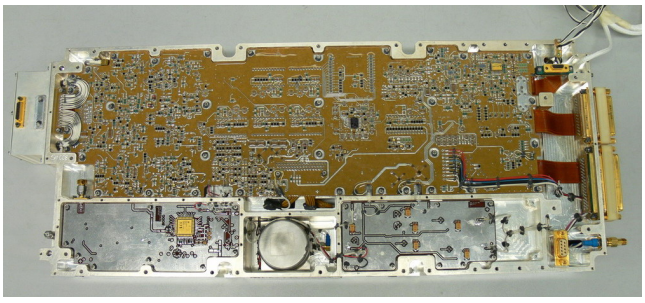


Figure 4. Main Electronic Module

IV. MPHM EBB INTEGRATION RESULTS

Both the Electronic Package (EP) and the Physic Package (PP) have been independently verified using reference PP and reference EP. This development philosophy has given sufficient degree of confidence that no issues were present before integration. The Programme objectives have been carefully verified by test.

A. REDUCTION OF TEMPERATURE SENSITIVITY

The EBB-MPHM thermal sensitivity has been verified in vacuum conditions applying several temperature transitions between +15°C and -5°C. Very low sensitivity to temperature variation has been observed, giving confirmation of the promising results obtained at EP and PP level. A detail of the instrument actual temperature sensitivity is provided in Figure 6.

The measured temperature coefficient is well below $1*10^{-14}$/°C, which is enough to comply with the Galileo Allan Deviation (ADEV) requirement, even considering a 20°C peak to peak temperature variation (refer to Figure 7).

Figure 5. Thermal Vacuum chamber

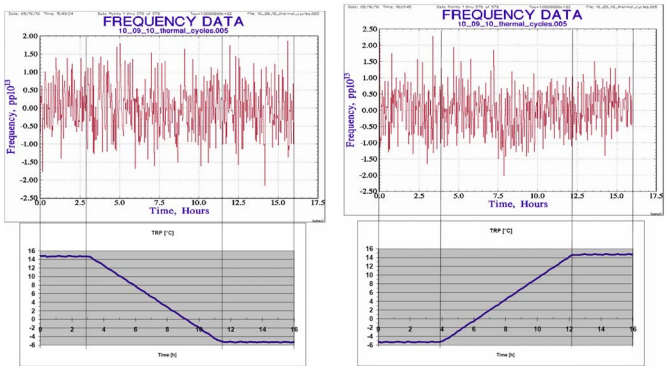


Figure 6. EBB-MPHM output frequency vs basplate temperature transition

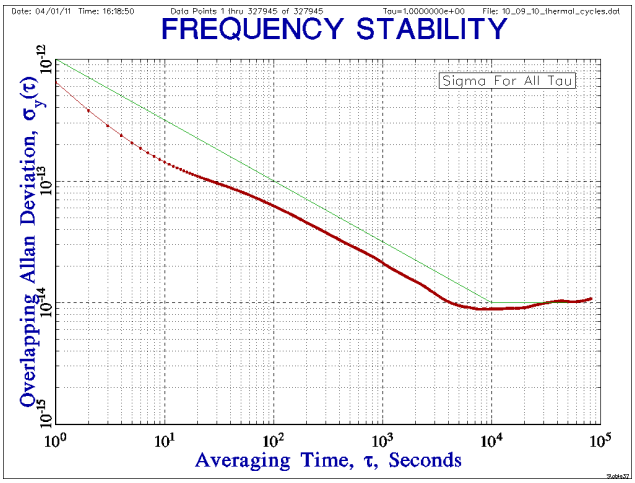


Figure 7. ADEV using frequency data during thermal cycle. In green the GALILEO requirement

B. FREQUENCY STABILITY PERFORMANCES

Measurements of frequency stability has been performed with a total duration of almost 1 month. The measured frequency drift is lower than $2*10^{-15}$/day. This value is perfectly in line with the typical PHM frequency drift [2] [3]. The measured frequency stability is reported in figure 8

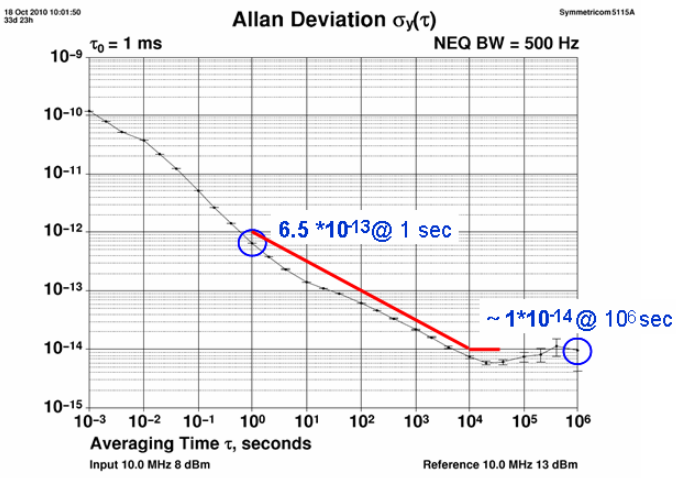


Figure 8. ADEV over 28 days of measurement. In red the GALILEO requirement.

It is worth noticing that the above measurement has been obtained after a few days of stabilization time and at the first attempt. The reduction of the added phase noise in the RF interrogation chain is noticeable from the good ADEV value achieved also at short term (i.e. $6.5*10^{-13}$ @ 1 sec). The long term measurement has given the opportunity to evaluate, from the EBB mPHM telemetry, the cavity drift. As expected the redesign of the microwave cavity and C-field coil has reduced the microwave cavity drift [2] by one order of magnitude with respect to the PHM typical cavity drift for the same observation time.

C. MASS AND POWER CONSUMPTION

The initial mass breakdown of the PHM is reported in the following table.

TABLE I. PHM MASS BREACKDOWN

Item	Mass (Kg)	Percentage
Electronic Package	6.2	33.33%
Physic Package	12.4	66.67%
Total	18.6	100%

The mass reduction achieved at PP level has reduced the relevant mass of 30%. The EBB mPHM resulting mass breakdown is detailed hereafter.

TABLE II. EBB MPHM MASS BREACKDOWN

Item	Mass (Kg)	Percentage
Electronic Package	6.2	43.42%
Physic Package	8.08	56.58%
Total	14.28	100%

The next step will be the reduction of the Electronic Package mass in order to achieve the target at instrument level of 12 Kg.

Power saving with respect to the PHM declared consumption has been obtained and further area of improvements have been identified.
10W power saving along the operating temperature range is expected starting from the actual measurements on the EBB mPHM.
The following table summarizes the expected improvements of the mPHM with respect to the PHM. Most of the parameters have been already verified on the EBB, providing a good degree of confidence on their feasibility.

TABLE III. PHM VS MPHM SUMMARY TABLE. IN GREEN THE IMPROVED PARAMETERS

	PHM	mPHM	Remark
Parameter			
Mass (Kg)	18.6	**12**	with mEP
Power Consumption (W)	64	**54**	@ 10°C
ADEV (sec)			
1	$1.8*10^{-12}$	**$6.5*10^{-13}$**	Measured on EBB
10	$3.2*10^{-13}$	**$1.4*10^{-13}$**	Measured on EBB
100	$7*10^{-14}$	**$6.3*10^{-14}$**	Measured on EBB
1000	$2.2*10^{-14}$	$2.2*10^{-14}$	
10000	$7*10^{-15}$	$7*10^{-15}$	
50000	$< 1*10^{-14}$	$<1*10^{-14}$	
Frequency drift (/day)	$< 1*10^{-14}$	$<1*10^{-14}$	
Temperature coefficient (/°C)	$2*10^{-14}$	**$1*10^{-15}$**	Measured on EBB

V. CONCLUSIONS

The outputs from the EBB Phase provide good confidence that a lighter PHM, with same frequency stability performances, but reduced constraints from integrator point of view, will be available in the close future.

It is worth noticing that the Mini-PHM shall be considered an evolution of the present PHM, because no new technology is developed, preferring an optimization or refining of existing solutions. This is basically the reason that makes the mPHM one of the most interesting clocks for future Space application, where outstanding frequency stability requirements are demanded in combination with stringent reliability figure.

Figure 9. Elegant Bread Board of the mPHM

REFERENCES

[1] P.Waller, F.Gonzalez, S.Binda, D.Rodriguez, G.Tobias, A.Cernigliaro, I.Sesia, P.Tavella, "Long-Term Performance Analysis of GIOVE Clocks" – PTTI 2010.

[2] M. Belloni, M. Gioia, F. Droz, P. Mosset et al. " Space Passive Hydrogen Maser - Performances, lifetime data and GIOVE-B related telemetries –" EFTF 2010.

[3] F. Droz, P.Mosset, Q.Wang, P.Rochat, M.Belloni, et al. "Space Passive Hydrogen Maser – Performances and lifetime data-", Proceeding of EFTF – IFCS 2009, 20-24 Apr. 2009, Besançon (France), pp 393-398.

Dick effect and cavity pulling on HORACE compact cold atom clock

N. Rossetto[1], F.X. Esnault[2], D. Holleville[1],
J. Delporte[3], M. Lours[1] and N. Dimarcq[1]
(1) : SYRTE - Observatoire de Paris
75014 Paris – France

(2): NIST – 325 Broadway,
Boulder, 80305 CO, USA
(3): CNES – 18 avenue Edouard Belin
31055 Toulouse Cedex - France

***Abstract*— HORACE is a compact cold caesium (Cs) atom clock using isotropic light cooling technique directly inside the spherical microwave cavity. It has been designed for onboard and space applications. Ultimate stability as low as 2.2 10^{-13} $\tau^{-1/2}$ had been demonstrated, limited by atomic shot noise and quantum projection noise. Influence of Dick effect has been studied and a compact and simple frequency synthesizer using off the shelf quartz has been realized and characterized. As atoms remain in the cavity during Ramsey sequence, long term stability evaluation focused on cavity pulling effect. This proceeding presents both the microwave synthesizer development and cavity pulling study using a tricky technique to tune the microwave cavity.**

I. INTRODUCTION

HORACE is a compact cold atom clock designed at SYRTE for onboard and space applications. In this clock all the interactions between atoms and light or RF magnetic field are performed inside the microwave cavity. Technical details about the setup can be found in Figure 1 and in [1].

The clock sequence is divided into four successive steps:

- Cooling of up to 2 10^8 atoms using the isotropic light technique.
- Preparation of the atomic sample in a pure hyperfine state.
- Ramsey interrogation performed while atoms are falling down inside the cavity.
- Optical detection of atomic population with a signal to noise ratio (SNR) up to 1000.

With this sequence, the clock cycle duration is about 85 ms. Comparison between atomic transition frequency and clock signal generated by microwave synthesizer is then performed every 85 ms. During time between two consecutive comparisons, the frequency fluctuations of the local oscillator (LO) induce a noise on the clock signal frequency, called Dick effect [2].

During all the interrogation time, the atoms remain still inside the microwave cavity. This difference with usual fountain clock justifies a detailled study of cavity pulling identified as one of the main systematics on HORACE clock.

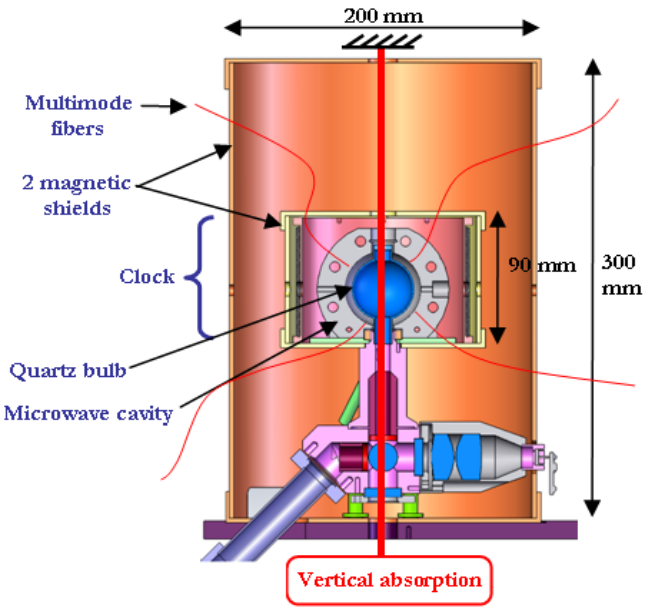


Figure 1: Cut-scheme of the experiment with the vacuum chamber surrounded by the cavity and the 2 magnetic shields. Inner walls of the cavity are optically polished to act as an integrating sphere. This cooling zone is fed by 6 multimode fibres. Detection of atomic populations is made with a retro-reflected vertical beam. The time of flight zone above the cavity is used for atomic temperature diagnostic.

II. NEW MICROWAVE SYNTHESIZER

In previous measurements presented in [1], the interrogation microwave signal was based on a 100 MHz very low noise signal generated from a Cryogenic Sapphire Oscillator (CSO) [3]. This CSO is 1 m^3 size and for obvious reasons is not adapted for onboard and space applications. In order to have a stand alone working of the clock, we need to use a portable LO. In addition, this new LO should not induce noticeable degradation of the short term stability of the clock. Therefore a new microwave synthesizer has been developed based on a suitable LO, designed to minimize the impact of Dick effect.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

In order to determine the characteristics of the new LO, some simulations have been performed, which estimate the impact of Dick effect on the short term stability directly from noise frequency spectrum of a LO. Assumptions made for these simulations are a cycle rate of 10 to 15 Hz and an interrogation time of 35ms, corresponding to typical values where interrogation time is limited by free fall of atoms in the cavity.

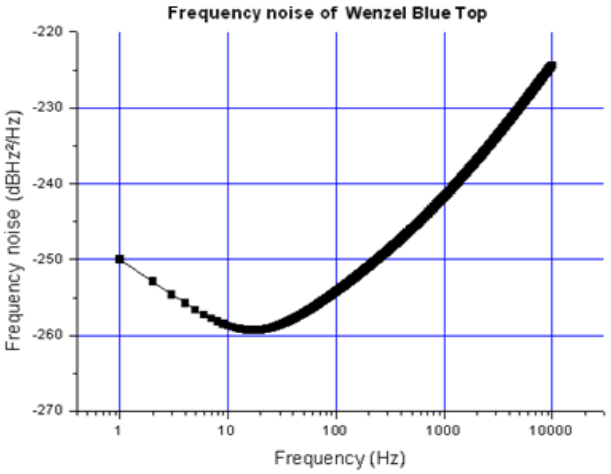


Figure 2: Frequency noise characteristics of the Wenzel Blue Top quartz. It presents a minimum around 15Hz, corresponding to the HORACE cycle rate.

With the objective to reduce impact of Dick effect above 2 10^{-13}, simulations lead to the selection of Wenzel Blue Top quartz. Its specifications are presented on Fig. 2 and show a frequency noise minimum around the clock cycle frequency. Considering the sequence parameters of HORACE and the characteristics of the Blue Top quartz, the expected contribution of the Dick effect is about 10^{-13}, which leads to a slight degradation of the short-term stability from 2.2 10^{-13} to the level of 2.4 10^{-13}.

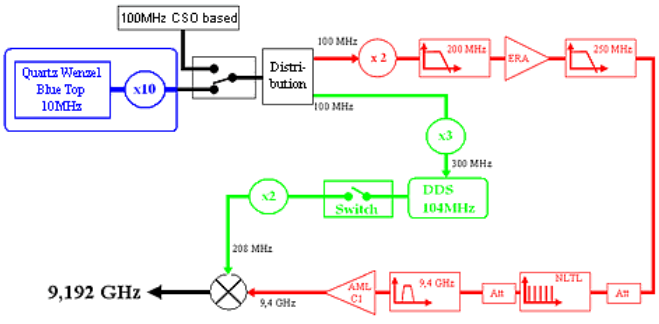


Figure 3: Scheme of the structure of the new synthesis. In red is the generation of the 9.4 GHz signal with a NLTL. In green is the RF signal produced by a DDS. To facilitate the characterization, the synthesis can be fed either by the Blue Top quartz, either by the reference based on CSO.

A new microwave synthesizer based on the Blue Top quartz has been build to generate the 9.192 GHz interrogation signal. The architecture of this synthesizer is shown on Fig. 3. The 10MHz signal generated by the Blue Top quartz is multiplied by ten to generate a 100 MHz reference signal. This signal feeds a NonLinear Transmission Line (NLTL), which the output RF frequency comb is filtered to keep the 94th harmonic. To obtain the clock signal, the 9.4 GHz signal is mixed with a 2x 104 MHz signal coming from a DDS synchronized with the 100 MHz reference signal.

The advantage of this new synthesizer is a simplicity that also characterises the whole clock:

- No secondary oscillator is used (like DRO or YIG). The output signal is generated from a NLTL [4].
- Instead of using a microwave switch, the RF component of the signal is toggled on/off to produce the 9.192 GHz signal.

Despite the use of RF filters in every stage, the output signal is polluted by spurious frequencies. Hopefully, the microwave cavity acts as a very narrow band filter that reduces by a factor of 40dB the parasite frequencies, as shown on Fig. 4.

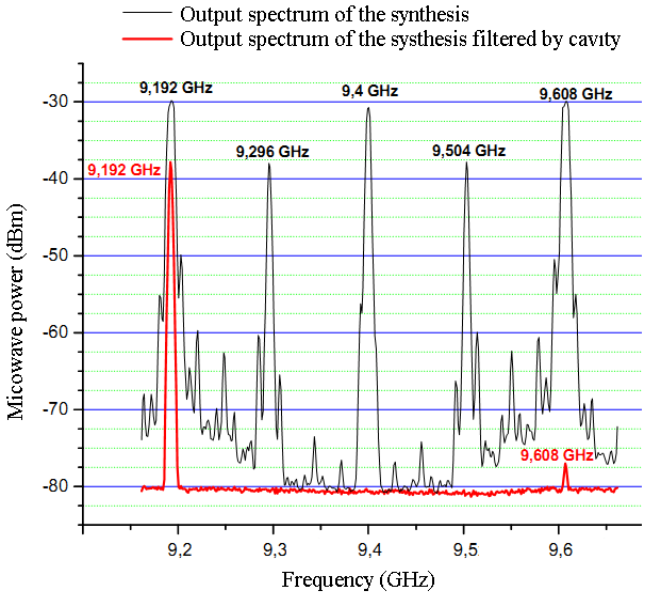


Figure 4: Filtering effect by the microwave cavity. The black line is the spectrum of the output signal of the new synthesis. The red line is the spectrum of the signal transmitted by the cavity.

The main objective of this new synthesizer is to reproduce the characteristics of the LO at 9.192 GHz. In other terms, the phase noise contribution of the synthesis must be negligible compared to the one of the Blue Top quartz, estimated at 9.192 GHz. The measurement has been performed using a similar synthesis, both fed by the same reference signal. Then the same measurement has been done using the Blue Top quartz as LO for the new synthesis, and the CSO signal for the other synthesis. The results are shown on Fig. 5.

To validate the new microwave synthesizer, short-term stability of HORACE clock has been measured using the CSO then the microwave synthesizer. Unfortunately, when these measurements have been performed, the signal to noise ratio was degraded due to extra instrumental noises, and the short-term stability with the CSO was of 3.3 10^{-13}. With the microwave synthesizer based on Blue Top quartz the short term stability increased to 3.7 10^{-13}, leading to a Dick effect induced by the synthesizer of 1.7 10^{-13}.

Although there is an extra noise compared to the expected 10^{-13} level Dick effect, this measurement validates the performances of this new microwave synthesizer as LO for an onboard HORACE clock.

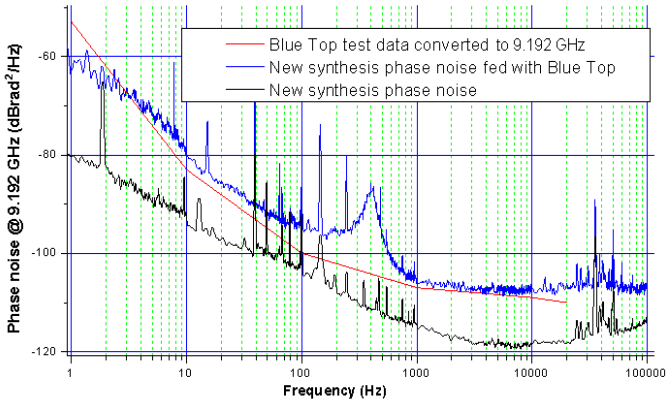


Figure 5: Phase noise of the new synthesis. The black line is the phase noise measurement of the new synthesis. The red line is the phase noise data of the Blue Top converted to 9,192 GHz. The blue line is the phase noise measurement of the new synthesis fed by the Blue Top against a similar synthesis fed by SYRTE reference signal. The excess of noise around 400 Hz is due to the comparison with CSO reference signal [5].

III. LONG TERM STABILITY: CAVITY PULLING

Contrary to the case of atomic fountain, in the case of HORACE the atoms remains still into the cavity during Ramsey interrogation, therefore the cavity pulling effect becomes the predominant systematic. As a first approximation model of cavity pulling in HORACE clock, we start from the one developed for atom fountain [6]:

$$\Delta\nu_{cav,pull} \propto \frac{\tau}{T}\frac{\mu_0\mu_B^2 N_{int} Q_{cav}}{2\pi\hbar V_{mode}} \times \left[\frac{2(\omega_{at}^2/Q)(\omega_{at}^2-\omega_c^2)}{(\omega_c^2-\omega_{at}^2)^2+(\omega_{at}\omega_c/Q)^2}\right] \quad (1)$$

In Eq. 1, the frequency shift depends on 3 parameters: the cold atoms number into the cavity (N_{int}) participating to the microwave interrogation, the ratio between the duration of each microwave pulse (τ) and the effective interrogation time (T), and the difference between resonant frequency of the cavity (ω_c) and the atomic frequency (ω_{at}).

In order to estimate the influence of the 3 parameters, some instrumental improvements have been done on the setup. First, electrical environment has been clean up. Then magnetic and thermal controls of the microwave cavity have been implemented to enhance the reliability of the experiment. With these upgrades, the long-term stability of HORACE clock has reached the 3 10^{-15} level within 2 10^4 seconds of integration.

We present an ingenious method to easily change the detuning of the microwave cavity. Instead of modifying the temperature of the cavity, as it used to be done in atomic fountains, we introduce some dielectric directly inside the microwave cavity. As shown on Fig. 1, there is a small air space between the quartz bulb and the inner walls of the cavity. Two small apertures allow the insertion of small rods: first one in copper to increase the frequency, second one in Pyrex to decrease it. In order to validate the method, the quality factor of the cavity and the transmitted RF power must remain the same. As presented on Fig. 6, the safe region allows a detuning of the cavity frequency of ± 600 kHz.

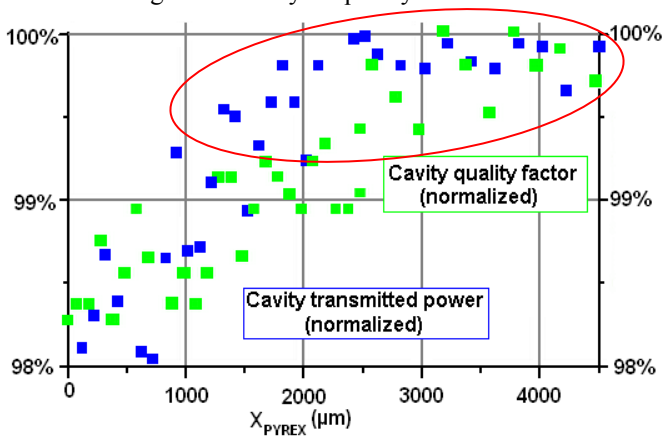


Figure 6: Variation of the normalized quality factor of the cavity (green squares) and of the normalized transmitted power by the cavity (blue squares), with the position of Pyrex cylinder. The same goes for the copper cylinder. The O position of the cylinder corresponds to the maximum of insertion into the cavity. The degradation of each factor does not excess 1%, for a position higher than 1500µm (corresponding to a 600 kHz tune). The useful zone is circled in red.

By fixing 2 of the 3 different parameters with a relative stability of 1% at one day for the cold atoms number, and uncertainty of 10 mK for the temperature of the cavity, we verified the influence of 2 of the 3 parameters:

- the τ/T dependency, with a slope of 1,9 10^{-12} which confirm the interest in using a Ramsey interrogation, instead of a Rabi interrogation, to minimize the amplitude of the effect (Fig. 7).

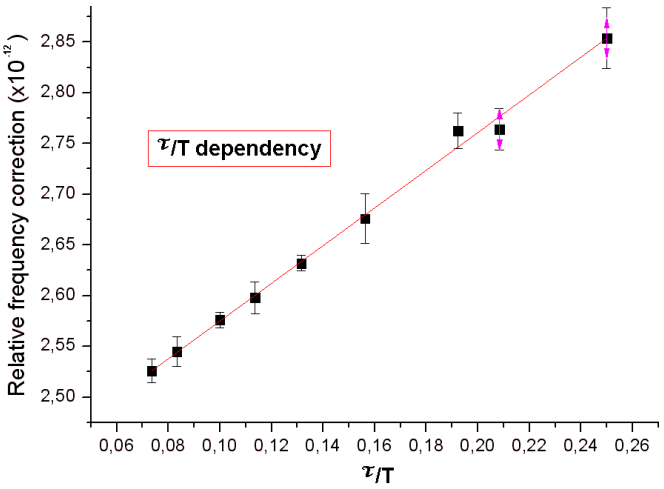


Figure 7: Linear dependency of the cavity pulling with the ration between the duration of a between the duration of each microwave pulse and the effective interrogation time.

- the cavity detuning dependency, with a maximum amplitude at $\nu_c = \nu_{at} \pm \Gamma/2$ with Γ as the line width of the resonant frequency of the cavity (Fig. 8).

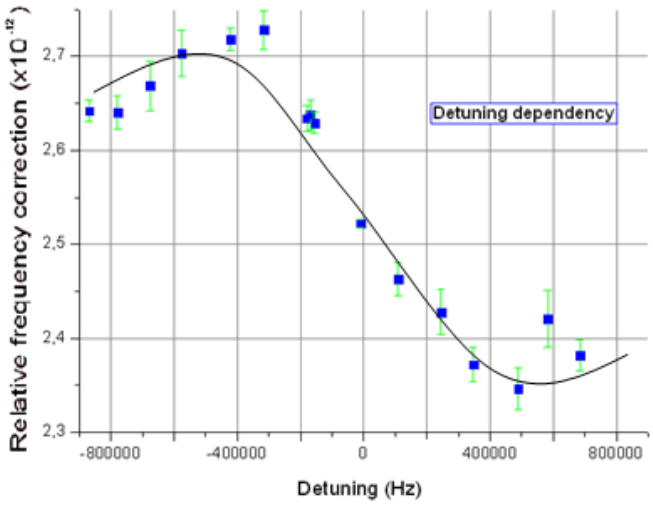


Figure 8: Relative clock frequency shift as a function of the cavity detuning, with τ = 2.5ms and T = 29ms.

Concerning the dependency with N_{int}, it appears that the effect decreases first with N_{int}, then increase again at low N_{int}. This can be related to one of the main issue in the experiment, concerning the lack of knowledge about the shape and the evolution of the cloud of cold atoms, and then the density of cold atoms and its evolution. The model used for the depedency need to be reviewed to adapt to the particularity of the clock.

IV. CONCLUSIONS AND PERSPECTIVES

With an expected contribution of Dick effect at 1 10^{-13} and a measured contribution at 1.7 10^{-13}, further investigations need to be done to understand this difference.

Concerning cold atoms density and its evolution, investigations are under progress, in order to refine the model of cavity pulling and to estimate the last main systematic measure: the shift due to cold collisions.

ACKNOWLEDGMENT

The authors would like to thank the French space agency (CNES) and Thales Electron Devices (TED) for supporting this work and for providing research studentship respectively. N. Rossetto would like to thanks R. Lambert and F. Chapelet for their essential help on the building and characterization of the synthesizer.

REFERENCES

[1] F.X. Esnault, D. Holleville, S. Guérandel, and N. Dimarcq, “Reaching a few 10^{-13} $\tau^{-1/2}$ stability level with the compact cold atom clock HORACE” ICSO 2008

[2] Special issue on the Dick effect. *IEEE Trans. Ultr. Ferr. Freq. Contr.*, vol. 45(4),

[3] Anthony G. Mann, Chang Sheng, and Andre N. Luiten, “Cryogenic Sapphire Oscillator with Exceptionally High Frequency Stability”, *IEEE Trans. Instr. Meas.*, 50 :2, 2001.

[4] R. Boudot, S. Guérandel, E. De Clercq. “Simple-Design Low-Noise NLTL-Based Frequency Synthesizers for a CPT Cs Clock” *IEEE Transactions on Instrumentation and Measurement* **58**, 10 (2009) 3659-3665

[5] Chambon, D. et al. "Design and realization of a flywheel oscillator for advanced time and frequency metrology," *Review of Scientific Instruments* , vol.76, no.9, pp.094704-094704-5, Sep 2005 doi: 10.1063/1.2018567

[6] S. Bize, C. Mandache Y. Sortais and, A. Clairon, and C. Salomon. Cavity frequency pulling in cold atom fountains. *IEEE Trans. Instrum. Meas.*, 50(2):503–506, 2001

Photonically Generated 10 GHz Microwaves with Close-to-Carrier Phase Noise < -100 dBc/Hz

T.M. Fortier, M. S. Kirchner, F. Quinlan, J.A. Taylor, N. Lemke, A. Ludlow, Y. Jiang, T. Rosenband, J. C. Bergquist, C.W. Oates and S.A. Diddams

Time and Frequency Division, National Institute of Standards and Technology
325 Broadway, MS 847 Boulder CO, 80305
fortier@boulder.nist.gov

***Abstract*—We demonstrate a 10 GHz photonic oscillator based on optical frequency division of a high-stability optical reference with a modelocked fs laser. Characterization with a second independent photonic oscillator reveals a close-to-carrier phase noise that is < 100 dBc/Hz, reducing to a shot noise floor of -156 dBc/Hz at a 1MHz offset.**

I. Introduction

A photonic approach to microwave generation that uses high quality factor ($Q \sim 10^{11}$) optical cavities and optical frequency division can enable high-stability and low timing-jitter signals without requiring cryogenic temperatures [1-3]. Stabilization of the light from a continuous wave (CW) laser to a high-finesse, passive optical cavity can yield a fractional frequency instability of $\Delta f/f \leq 1\times10^{-15}$ for short averaging times [4, 5]. Ideally, transfer of this stability to the microwave domain with a modelocked fs laser preserves the frequency stability of the reference while reducing the optical phase noise by the division factor squared. Using this technique, we demonstrate a 10 GHz microwave signal with a fractional frequency instability $< 1\times10^{-15}$ and with a phase noise that is < 100 dBc/Hz at a 1 Hz offset. The integrated timing jitter of the phase noise spectrum from 1 Hz to 1 MHz is < 1 fs.

II. Generation of 10 GHz Micorwave Signals Via Optical Frequency Division

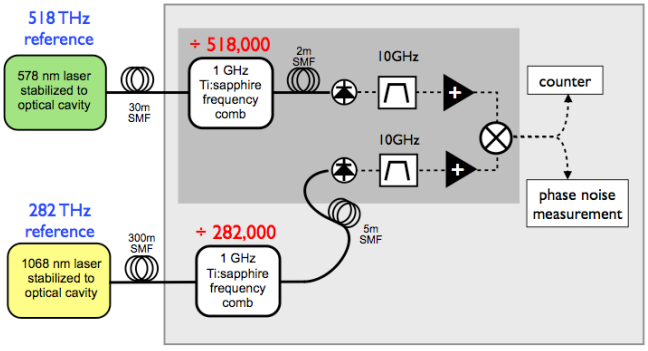


Figure 1. Generation and characterization of 10 GHz signals via optical frequency division. SMF denotes single mode fiber.

We generate 10 GHz microwave signals via photodetection and of the tenth harmonic of the repetition rate, f_r, of a 1 GHz Kerr-lens modelocked Ti:Sapphire laser [6] that is phase-locked to an high-Q optical reference [4, 5]. Stabilization of an optical frequency comb divider (OFCD) to an optical reference transfers the stability of that reference to a timing stability in the OFCD laser pulse train. Photodetection of this pulse train produces photocurrent that yields a microwave spectrum of harmonics of f_r up to the photodetector (PD) cut-off bandwidth. A spectrally pure 10 GHz microwave signal is extracted via electronic filtering and amplification of the photocurrent signal. As seen in Figure 1, a second similar but fully independent photonic oscillator is used to characterize the frequency stability and phase noise of the generated 10 GHz microwave signals. Details of the measurement methods can be found in [1].

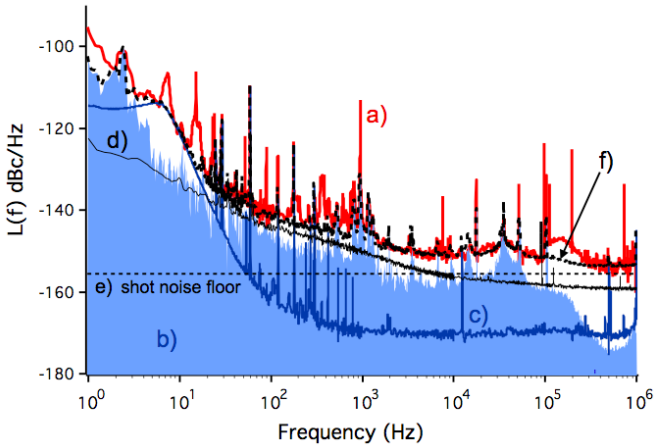


Figure 2. Single sideband phase noise on a 10 GHz carrier contributed by two oscillators. a) red solid trace – phase noise resulting from the comparison of two independent photonic oscillators. b) blue filled trace - phase noise contribution, scaled to 10 GHz, of the optical references, and the optical frequency comb dividers. c) blue solid trace – phase noise of the digital phase noise measurement system. d) black solid trace – phase noise contributed by the electronic amplifiers. e) Shot noise floor from two photonics oscillators. f) black dashed line – sum of the noise contributions from b), c), d) and e)

Figure 2 shows the phase noise on a 10 GHz carrier that is generated via comparison of two photonic oscillators. Also shown is the optical phase noise, scaled to 10 GHz, which is contributed by the optical references and the residual noise of

the OFCDs. Fig 2b represents the microwave phase noise that would be possible ignoring the noise that is contributed by the measurement electronics and photodetection of the laser repetition rate. Figure 2c is the calculated noise that results from contributions of the electronic amplifiers, the phase noise measurement system, the shot noise floor, the optical references and the residual noise of the OFCD. The good agreement between the calculated and the measured phase noise demonstrates that we have properly identified the current limitations to the purity in the 10 GHz microwave signal.

In conclusion, we have produced 10 GHz microwaves with sub-femtosecond timing jitter. The demonstrated phase noise produced is comparable to, or better than those achieved with cryogenic dielectric oscillators [7-9]. Future work will aim at reducing the phase noise contributions of the electronics and photodetection. At frequencies > 10 kHz photodetector nonlinearities result in a saturation of the photocurrent at -8 dBm for 12 mW of incident optical power [10]. In the future, the shot noise limited floor can be decreased with more optical power provided the high-speed photodetector remains linear. Additionally saturation effects in the PD can be mitigated by repetition rate filtering with a Fabry-Perot cavity [Kirchner]. A still lower noise floor would require photodetectors with increased power handling or an alternative hybrid approach with a dielectric sapphire oscillator locked to our photonic oscillator.

The fundamental limits to the achievable phase noise of our photonic approach are given by thermally-excited noise in the glass spacer, mirror substrates and optical coatings of the FP reference cavity [Numata], as well as the photodetector shot noise mentioned above. The lowest reported thermal noise for a FP reference cavity corresponds to approximately L(f) = -117 dBc/Hz at 1 Hz offset when converted to a 10 GHz carrier.

III. Acknowledgment

We thank A. Hati, L. Hollberg, D. Howe, C. Nelson, N. Newbury, and S. Papp for their contributions and comments on this manuscript, and A. Joshi and S. Datta of Discovery Semiconductor for providing the 10 GHz InGaAs photodiodes. This work was supported by NIST. It is a contribution of an agency of the US government and is not subject to copyright in the US.

References

[1] Fortier T.M. et al., "Generation of Ultrastable Microwaves via Optical Frequency Division," Nature Photonics, in press.

[2] Bartels, A. et al., "Femtosecond-laser-based synthesis of ultrastable microwave signals from optical frequency references," Opt. Lett. **30**, 2005, pp. 570-572.

[3] Millo J. et al., "Ultralow noise microwave generation with fiber-based optical frequency comb and application to atomic fountain clock," Appl. Phys. Lett. **94**, 2009, pp.141105-141103.

[4] Jiang Y.Y. et al. "Making optical atomic clocks more stable with 10-16-level laser stabilization," Nature Photonics **5**, 2011, pp.158-161.

[5] Young B.C. et al., "Visible Lasers with Subhertz Linewidths," Physical Review Letters **82**, 1999, pp. 3799-3802.

[6] Fortier T.M., Bartels A. and Diddams S.A., "Octave-spanning Ti:sapphire laser with a repetition rate of 1 GHz for optical frequency measurements and comparisons," Opt. Lett. **31**, 2006, pp. 1011-1013.

[7] Mann, A. G., Sheng, C. & Luiten, A. N. Cryogenic Sapphire Oscillator with Exceptionally High Frequency Stability. *IEEE Trans. Inst. Meas.* **50**, 519-521 (2001).

[8] Locke, C.R., Ivanov, E.N., Hartnett, J.G., Stanwix, P.L. & Tobar, M.E. Invited Article: Design techniques and noise properties of ultrastable cryogenically cooled sapphire-dielectric resonator oscillators. *Review of Scientific Instruments* **79**, 051301-051312 (2008).

[9] Grop, S. *et al.* ELISA: A cryocooled 10 GHz oscillator with 10^{-15} frequency stability. *Review of Scientific Instruments* **81**, 025102-025102-025107 (2010).

[10] Taylor J.A. et al. "Characterization of Power-to-Phase Conversion in High-Speed P-I-N Photodiodes," IEEE Photonics Journal, 2011, pp. 140-151.

[11] Diddams et al. "mproved signal-to-noise ratio of 10 GHz microwave signals generated with a mode-filtered femtosecond laser frequency comb," Opt. Exp. **17**, 2009, pp. 3331-3340.

[12] Numata, K., Kemery, A. & Camp, J. Thermal Noise Limit in Frequency Stabilization of Lasers with Rigid Cavities. *Phys. Rev. Lett.* **93**, 250602 (2004).

High Frequency Thermal-Piezoresistive MEMS Resonators for Detection of Organic Gases

Arash Hajjam, Andrew Logan, Jagadeesh Pandiyan and Siavash Pourkamali
Department of Electrical and Computer Engineering, University of Denver, Denver, CO, USA
Email: ahajjam@du.edu

***Abstract*— This paper reports on high frequency micromechanical resonant silicon microbalances capable of detection and concentration measurement of volatile organic compounds in gas phase. In this work, regular gasoline has been used as the source of organic compounds in gas phase and 1813 photoresist has been used as the absorbent coating on the resonator surfaces. A new technique for formation of thicker polymer coatings on resonator surfaces has been utilized allowing frequency shifts as high as 2200ppm to be achieved for the resonators upon exposure to gasoline vapor. This is equivalent to an effective thickness of ~40nm for the absorbed organic compounds. Measured response time constants are around 40 seconds. Faster and more sensitive responses are expected to be achievable by maximization of the absorbing surface area using porous resonator structure coatings.**

I. INTRODUCTION

Sensors capable of detection of organic compounds in gas phase could have several applications in oil and gas industry among others. Examples of such applications include rapid estimation of oil content of oil sand samples and early detection of hazardous leaks. Such sensors can help save plenty of time and resources during oil exploration by avoiding the costly and time consuming process of using off-site laboratory analysis. Sensors for detecting volatile organic compounds using surface acoustic wave (SAW) resonators and quartz crystal microbalances (QCM) have previously been demonstrated [1-4]. The detection mechanism in such sensors is based on a frequency shift caused by the added mass from the absorbed molecules.

MEMS resonators, as smaller versions of such resonant devices, are interesting candidates for development of more sensitive and highly integrated chemical and biological sensors. Arrays of such sensors can be batch fabricated at high quantities at a very low cost. As a result, a variety of microscale electromechanical resonators have been considered for sensory applications lately [5,6].

MEMS resonators mostly make use of electrostatic (capacitive) [7] or piezoelectric [8,9] electromechanical transduction with both methods having their respective advantages and disadvantages. Electrostatic excitation suffers from relatively small actuation forces available from electrostatic actuators and the need for very thin transduction gaps complicating fabrication of such devices. Furthermore, narrow transduction gaps are also very vulnerable to contaminants in environmental sensing applications.

Another popular type of MEMS resonators are the piezo-electrically actuated resonators which have better robustness in comparison to electrostatic resonators but piezoelectric transduction requires integration of piezoelectric thin films and metallic electrodes with the micromechanical elements which makes the fabrication somewhat challenging. Such devices usually have lower quality factors in comparison to capacitive devices.

Thermally actuated resonators are the third group of MEMS resonators [10-12]. In addition to having a very simple fabrication process, they are very suitable for sensory applications. Their monolithic structures have shown excellent robustness when in direct contact with the surrounding environment Frequencies up to 60MHz have been demonstrated for thermally actuated micromechanical resonators [13] and much higher frequencies can be achieved by shrinking the resonator dimensions [14]. Power consumption is usually a concern for thermally actuated devices but it has been shown that if the devices are small enough, power consumption can be in the sub-milliwatt range [14]. In addition, power consumption makes such devices active components capable of self sustained oscillation without the need for additional electronic circuitry [15].

II. RESONATOR FABRICATION

The standard single mask SOI process was used for fabrication of the thermal-piezoresistive resonators used in this work [15]. This process starts by growing a thin (~200 nm) layer of silicon dioxide on the substrate. The silicon dioxide layer is patterned to define the resonator structures. The silicon structures are then carved into the SOI device layer by deep reactive ion etching (DRIE) of silicon all the way down to the buried oxide layer (BOX). Finally, the underlying buried oxide is removed in hydrofluoric acid (HF). At the same time the remaining oxide mask on top of the structures is also etched away. The whole procedure is shown in steps a to c of Figure 1. The result is suspended single crystalline silicon

This work was supported by National Science Foundation (NSF) under grants MRI-RAPID #1061489 and MRI #0923518.

resonant structures with integrated thermal actuator/piezoresistive sensors.

In a previous work, polymer coatings were formed by dipping the resonator sample in a highly diluted polymer solution in acetone leading to very thin polymer coatings [16]. The resonators were spun and hard baked in order to get rid of the excess solvent (step d of Figure 1). Even though the results in [16] showed that the coated resonators were responding to organic vapors, the sensitivity was not very promising and a maximum frequency shift of only 200 ppm was obtained for a highly saturated gas sample. The reason for such low frequency shift was suspected to be the small polymer thickness on the surface of the devices. Much higher sensitivities are expected to be achievable for thicker polymer coatings.

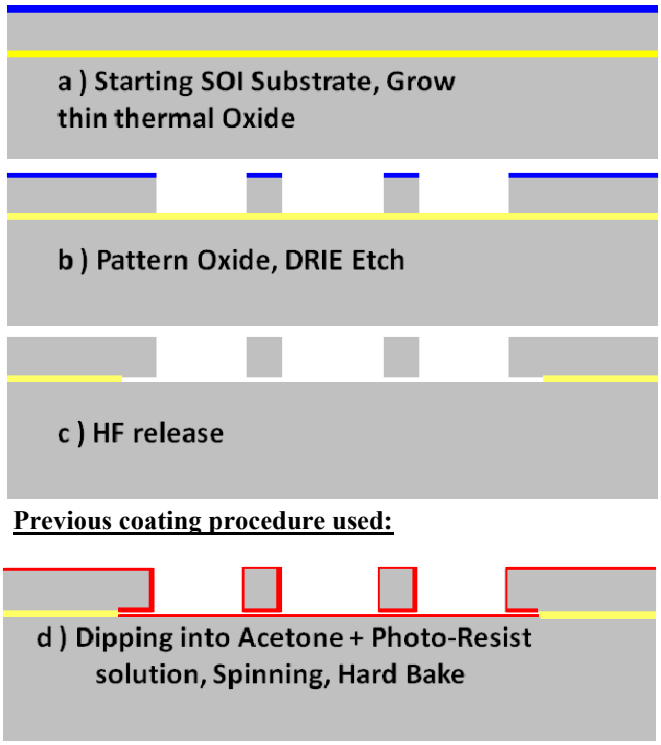


Figure 1. a-c) Thermally actuated resonator fabrication process flow and d) polymer coating.

Using higher concentration polymer solutions in this technique to achieve higher polymer thickness generally results in excessive polymer residues in the gap between the resonators and the SOI handle layer leaving the devices non-operational as can be seen in the SEM picture in Figure 2.To overcome this problem, a new procedure was utilized in this work for polymer coating of the devices (Fig. 3). Steps a and b which are not shown in Fig. 3 are identical to the fabrication process shown in Figure 1. However, in this approach the polymer solution (a much higher concentration solution) is added and spin coated before the devices are undercut. Undercut is done in hydrofluoric acid after coating with polymer (steps c and d of the fabrication process shown in Figure 3. Therefore, using this approach there will be no polymer residue underneath the device structures.

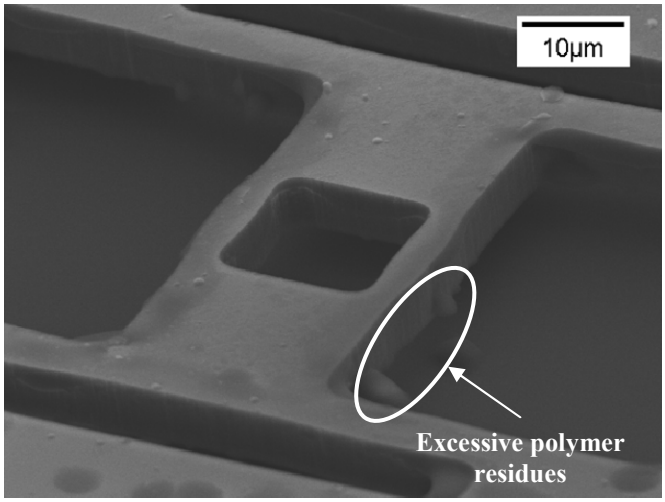


Figure 2 SEM view of a device in which high polymer thickness has resulted in excessive polymer residues in the gap between the resonators and the SOI handle layer which has left the devices non-operational.

SEM view of a 24MHz I-shaped bulk acoustic resonator polymer coated using the abovementioned technique is also shown in Fig. 3. This resonator has been fabricated on a SOI substrate with both device layer and BOX thickness of 5µm.

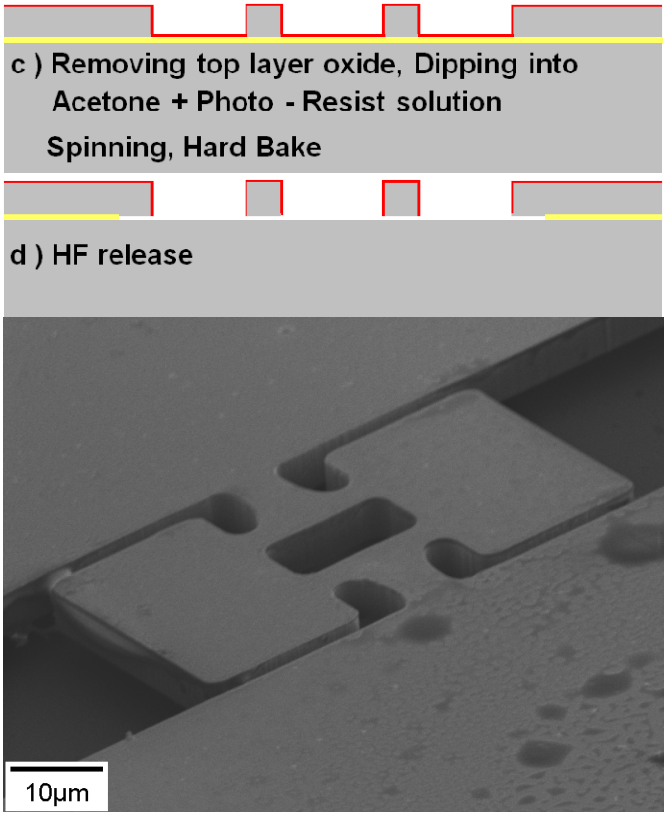


Figure 3. New procedure for coating of the thermally actuated resonators in which the polymer is added before the devices are undercut and undercut is done in hydrofluoric acid after coating with polymer. Also, an SEM picture of a 24MHz IBAR resonator used in this work along with its electrical connections is shown.

III. Resonator Description

Figure 4a shows the schematic view of the resonant structures used in this work referred to as dog bone or I-shaped resonators. They can cover a wide range of frequencies from a few MHz to tens of MHz. Such devices are very suitable for thermal actuation and their operating principle has thoroughly been explained in [13]. The resonators were operated in a one-port configuration with the thin actuator beams of the structures acting simultaneously as both thermal actuators and piezoresistive sensors. Actuation occurs by passing a fluctuating electrical current through the actuator beams in the middle of the structure.

Figure 4b shows the in-plane extensional resonant mode of an I-shaped resonator. In this mode the masses on the two ends of the beams vibrate back and forth in opposite directions. At resonance, the resistance of the actuator beams is modulated by the resulting alternating mechanical stress due to the piezoresistive effect. This results in a detectable small signal motional current in the device.

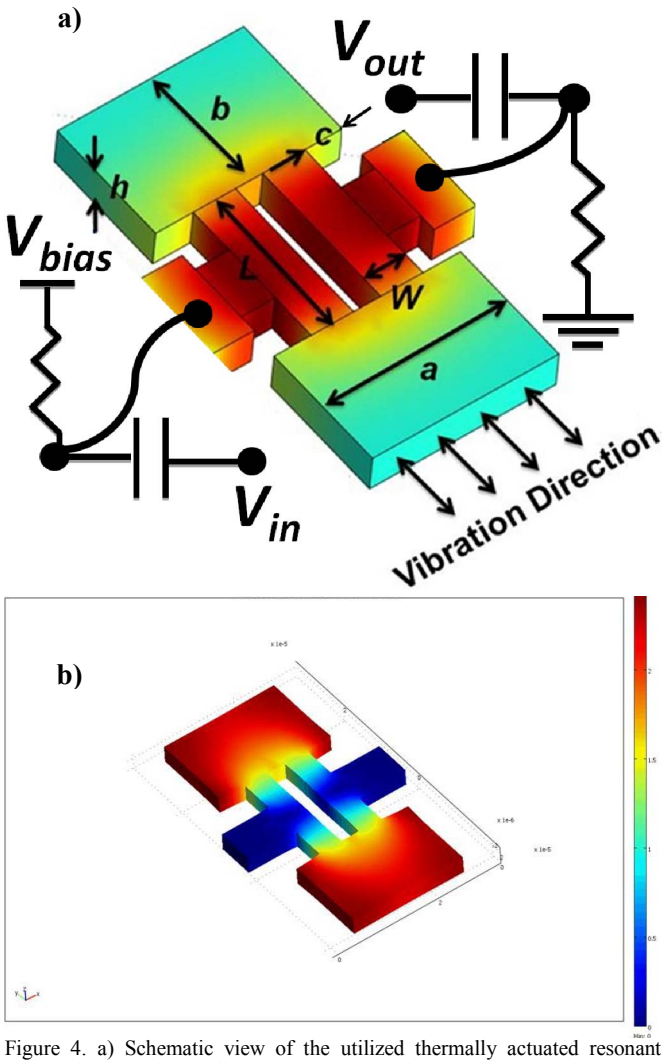


Figure 4. a) Schematic view of the utilized thermally actuated resonant sensor showing the qualitative distribution of AC temperature fluctuation amplitude (red being the maximum and blue the minimum). The electrical connections required for one-port operation of the resonator are also shown; b) COMSOL eigen frequency analysis results showing the fundamental in-plane resonance mode shape for an I-shaped structure.

IV. Measurement Setup And Results

Figure 5 shows the camera picture of the setup used in order to characterize the resonator sensory response as well as its schematic diagram. In this setup, nitrogen bubbled through a gasoline container is directed towards the resonator under test.

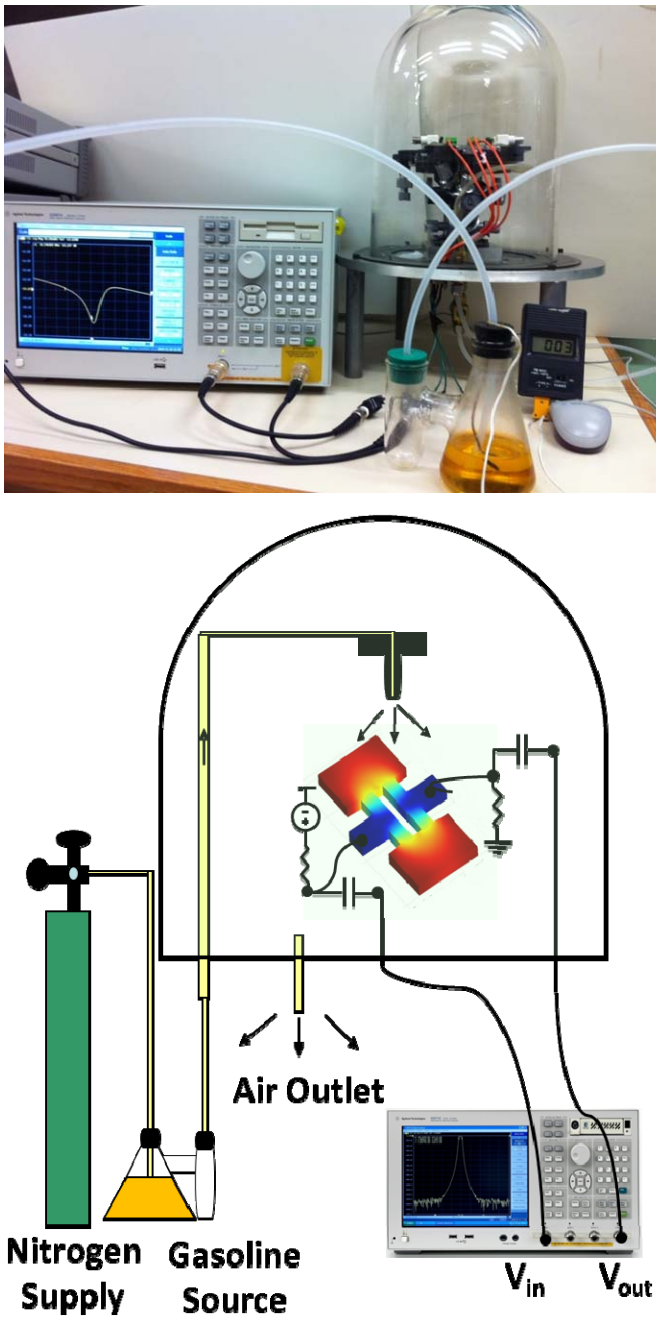


Figure 5. Schematic diagram and camera picture of the test setup used to characterize the resonator sensory behavior in which nitrogen, bubbled through a gasoline container, is directed towards the resonator under test.

Having a flow of the gas sample passing over the resonator surfaces maximizes the contact between potential molecules of interest in the gas sample and the resonator surface, hence, maximizing the absorption probability.

Figure 6 shows the change in the measured frequency versus time for the resonator shown in Fig. 3 upon exposure to nitrogen flow rich with gasoline vapor at room temperature. As can be seen, upon exposure, the resonance frequency reduces over time with a time constant of ~40sec. Restoration of the resonant frequency to its initial value happens very quickly as soon as the source is removed. It can be seen that more than 10 times freq shift (in ppm) is achieved in comparison to ~200ppm frequency shift obtained using the previous coating procedure [16].

The mass sensitivity of the resonators can be theoretically calculated as follows:

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \Rightarrow \frac{\partial f}{\partial m} = -\frac{f}{2m} \qquad (1)$$

where k, m, and f are the effective stiffness, effective mass and resonant frequency of the resonator respectively. Knowing the dimensions and therefore the mass of the resonators, the mass of the deposited organic molecules (∂m) was estimated from the measured frequency shifts to be ~35pg. Assuming an approximate density of 720 kg/m^3 for the organic compound, the volume and consequently the effective thickness of the absorbed organic compounds (assuming a uniform distribution over all resonator surfaces) was calculated. The observed maximum frequency shift of 2200ppm (0.22%) is equivalent to an absorbed organic layer of ~40nm thick.

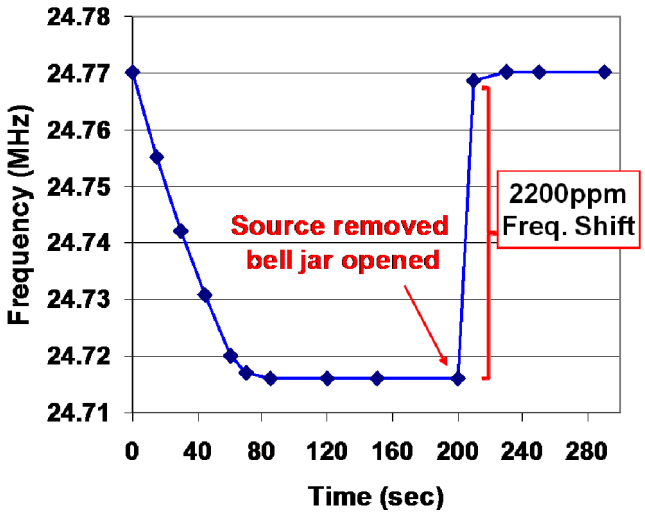


Figure 6. Change in the measured resonance frequency for the 24MHz IBAR resonator of Fig. 4 as a function of the cumulative exposure time showing an overall frequency shift of ~54KHz (2200 ppm). The source was removed after 200 sec. The return to a frequency very close to the initial frequency is almost instantaneous.

The vapor pressure which is an indication of a liquid's evaporation rate increases exponentially with temperature according to the Clausius–Clapeyron relation. In order to see how the sensor responds to different gas concentrations its frequency shift was measured while varying the temperature of the gasoline source. Fig. 7 shows the resonator frequency shift as a function of the temperature of the gasoline sample. As expected, at lower temperatures the lower organic vapor pressure leads to smaller frequency shifts. The shift increases exponentially with temperature (as the vapor pressure increases exponentially with temperature). At about 5ºC, the sensor response starts to saturate and stays constant for temperatures above ~15 ºC .

Gasoline is a mixture of different hydrocarbons with Toluene being one of the main ones. Therefore, in order to approximate the concentration of the detected gasoline vapor, our calculations are based on the saturated vapor pressure of Toluene. Based on the vapor pressure taken from [17], assuming a temperature of -10C°, the minimum gasoline concentration detected corresponds to ~4000 ppm whereas; the minimum LOD measurable would be ~5.5ppm.

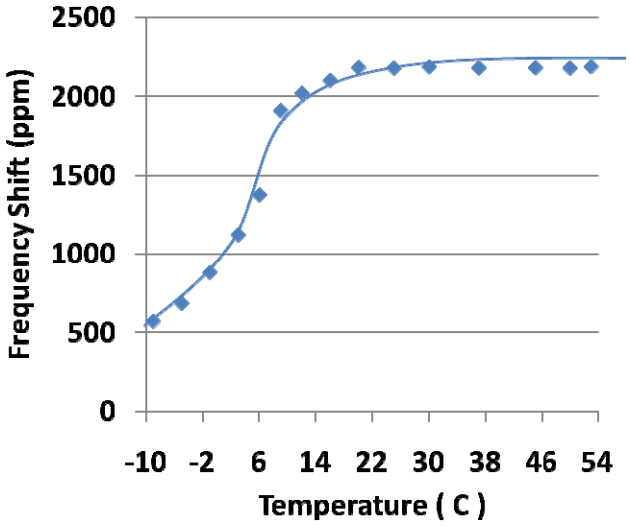


Figure 7. The resonator frequency shift as a function of the temperature of the gasoline. The frequency shift increases with temperature as the organic vapor pressure increases. The increase is exponential at lower temperatures, until the sensor reaches saturation.

V. CONCLUSION

A thermally actuated high frequency in-plane silicon resonator was successfully used for estimation of concentration of gasoline vapor in gas samples. Such a device has a monolithic crystalline silicon structures which makes it very robust towards contaminants and can be fabricated using a very simple CMOS compatible process. The observed maximum frequency shift of 54kHz (2200ppm) is approximately 10 times the frequency shift measured using the previous resonator coating procedure. This measured frequency shift corresponds to a thickness of 40nm deposited organic compound on the surface of resonator. Minimum LOD was estimated to be ~5.5 ppm. Future work includes further design optimization of the resonators and narrowing down their thickness to achieve higher mass sensitivities. Using more suitable absorbent polymer coatings and maximization of the absorbing surface area using porous resonator structure coatings could significantly improve the sensitivity.

ACKNOWLEDGMENT

Authors would like to thanks the staff at Georgia Tech nanotechnology research center and Dr. Ashwin Samarao for their help with silicon DRIE (deep reactive ion etching of silicon).

REFERENCES

[1] K.H Clifford, E. R. Lindgren, K. S. Rawlinson, L. K. McGrath and J. L.Wright "Development of a surface acoustic wave sensor for In-Situ monitoring of volatile organic compounds," J. Sensors, 3 , pp. 236-247, 2003.

[2] K.H Clifford., and R.C. Hughes, 2002, *In-Situ* Chemiresistor Sensor Package for Real-Time Detection of Volatile Organic Compounds in Soil and Groundwater, *Sensors*, 2, 23-34

[3] Frye, G.C., 2000, In situ field screening of volatile organic compounds using a portable acoustic wave sensor system, in Vadose Zone Science and Technology Solutions, Looney, B.B. and R.W.Fallta (eds.), Battelle Press, Columbus, 580-589.

[4] I.Sugimoto, M.Seyama and M.Nakamura, "Detection of petroleum hydrocarbons at low ppb levels using quartz resonator sensors and instrumentation of a smart environmental monitoring system," J.Environ.Monit., 1, pp. 135-142,1999.

[5] Y. Hwang, F. Gao, and R. N. Candler, "Porous silicon resonator for sensitive vapor detection," Solid State Sensors and Actuators Workshop, Hilton Head, pp. 150-153, 2010.

[6] A. Hierlemann, O. Brand, C. Hagleitner, H. Baltes, "Microfabrication techniques for chemical/biosensors," Proc. IEEE, vol. 91, pp. 839-863, 2003.

[7] S. Pourkamali, G. K. Ho and F Ayazi, "Low-impedance VHF and UHF capacitive silicon bulk acoustic wave resonators," *IEEE Trans.Electron Devices*, vol. 54, no. 8, pp. 2017-2023, Aug. 2007.

[8] G.K. Ho, R. Abdolvand, A. Sivapurapu, S. Humad, F. Ayazi, "Piezoelectric-on-Silicon Lateral Bulk Acoustic Wave Micromechanical Resonators," *J. Microelectromech. Syst.,* vol. 17, no. 2, pp. 512 – 520, Apr. 2008.

[9] B. Harrington, M. Shahmohammadi, and R. Abdolvand, "Toward Ultimate Performance in GHz MEMS Resonators: Low Impedance and High Q," IEEE MEMS 2010, pp. 707 – 710.

[10] J. H. Seo and O. Brand, "High Q-Factor In-Plane-Mode Resonant Microsensor Platform for Gaseous/Liquid Environment," *J. Microelectromech. Syst.*, vol. 17, no. 2, pp. 483-493, Apr. 2008.

[11] A. Hajjam, A. Rahafrooz, J.C. Wilson, and S. Pourkamali, "Thermally Actuated MEMS Resonant Sensors for Mass Measurement of Micro/Nanoscale Aerosol Particles," in *Proc. 8th IEEE Sensors*, Christchurch, New Zealand, Oct. 2009, pp. 707-710.

[12] A. Hajjam, J.C. Wilson, A. Rahafrooz and S. Pourkamali, "Fabrication and characterization of thermally actuated micromechanical resonators for airborne particle mass sensing: II. Device fabrication and characterization," *J. Micromech. Microeng.,* vol. 20, 12501, Dec 2010.

[13] A. Rahafrooz, A. Hajjam, B. Tousifar, and S. Pourkamali, "Thermal actuation, a suitable mechanism for high-frequency electromechanical resonators," in Proc. 23rd IEEE MEMS, Hong Kong, Jan. 2010, pp. 200–203.

[14] A. Rahafrooz and S. Pourkamali, "Thermo-electro-mechanical modeling of high frequency thermally actuated I^2-BAR resonators," in Proc. Tech. Dig. Solid-State Sens., Actuator, Microsyst. Workshop, Hilton Head Island, SC, Jun. 2010, pp. 74–77.

[15] A. Rahafrooz and S. Pourkamali, "Fully micromechanical piezo-thermal oscillators," to be published, in *IEEE International Electronic device meeting (IEDM)*, San Francisco, Dec. 2010.

[16] A. Hajjam, J. Pandiyan, A. Rahafrooz, and S. Pourkamali, "MEMS Resonant Sensors for Detection of Gasoline Vapor," IEEE Sensors Conference, Nov. 2010, pp. 1538-41.

[17] Data sheet, CRC Handbook of Chemistry and Physics 44th ed, http://en.wikipedia.org/wiki/Toluene_(data_page)

Receptor-Coated Porous Silicon Resonators for Enhanced Sensitivity of Vapor Detection

Yongha Hwang, Sungmin Kim, and Rob N. Candler
Electrical Engineering Department
University of California, Los Angeles
Los Angeles, USA
Email: hwangyongha@ucla.edu

***Abstract*—We present microscale silicon resonators with nanoscale pores for increased surface area, which are capable of providing enhanced sensitivity in vapor sensors. Increased mechanical stability and detection performance are also achieved by keeping parts of the resonating device nonporous and adding a receptor coating. We demonstrate improvements up to 165% and 654% in resonator sensitivity to isopropyl alcohol (IPA) for partially-porous silicon resonators and receptor-coated partially-porous silicon resonators, respectively, as compared to nonporous silicon resonators.**

I. INTRODUCTION

MEMS/NEMS resonators are commonly used as chemical or biological sensors, as surface-bound analytes change the mechanical stress and total mass of the resonator, leading to measureable shifts in resonant frequency. The sensitivity to adsorbates can be enhanced by scaling devices down to the nanoscale. Ultrasmall NEMS resonators have demonstrated resolution on the order of attograms or less [1-2]. These extremely sensitive resonant sensors, however, have additional challenges, such as fabrication process and signal detection [3-4].

Rather than fabricating nanoscale devices, our approach uses widely available microscale fabrication techniques to achieve porous structures with nanoscale features, for the possibility of microscale devices with nanoscale sensitivity. By building resonators that have a porous silicon surface, we take advantage of the increased area for molecular binding and improve the sensitivity of the resonators to the vapor concentration of interest. In addition, while achieving high sensitivity by shrinking resonators to the nanoscale results in the decrease of sensing area, the microscale sensors with nanoscale porous of this work can simultaneously achieve high sensitivity and a large capture cross section. This enables the porous resonators to detect low concentrations of analytes without the sample preconcentration that may be required for nanoscale resonators.

We previously demonstrated initial results for silicon resonators that were fully porous [5]. This work focuses on silicon resonators with targeted regions of porosity and functionalization, combining the advantages of increased surface area of porous silicon with the structural strength of nonporous silicon.

II. FABRICATION OF RECEPTOR-COATED POROUS SILICON RESONATORS

The nonporous silicon resonator consists of whole polycrystalline silicon. Regions of the nonporous resonators can be converted to porous by electroless metal-assisted etching. We designed a fully-porous silicon resonator and a partially-porous silicon resonator in order to compare their performance. While both the center paddle and torsional beams of the fully-porous structures consist of porous-etched silicon, only the center paddle is porous-etched, and the torsional beams remain nonporous-etched silicon for the partially porous device as shown in Fig. 1(h). This is done in order to maintain the stiffness of the torsional beams and improve yield during the release process.

We have fabricated each type of resonator using surface micromachining process as shown in Fig.1. Porous and nonporous silicon resonators were fabricated on the same wafer in order to quantify the enhancement in sensitivity. The resonators were fabricated on 4-inch (100) silicon wafers as the substrates, followed by depositions of 5.5-μm-thick silicon dioxide serving as a sacrificial layer and 470-nm-thick polycrystalline silicon by LPCVD. After the resonator structures were defined by photolithography and deep reactive ion etched, a standard photoresist (AZ 5214) was patterned to serve as a mask for selective porous etching. The surface was made porous using an electroless metal-assisted etching technique, which involves randomly depositing gold clusters as chemical catalysts [6]. A cathode reaction happens with hydrogen peroxide on the top surfaces of gold clusters, and an anode reaction happens with silicon and hydrofluoric acid on the bottom surfaces. These reactions supply hole injection into silicon and etch silicon molecules underneath gold clusters. By changing the fabrication conditions, such as thickness of gold clusters and etch time of porous silicon, porosification

This work was supported by the BRIGE program of the National Science Foundation under award # 0926228 and the Center on Functional Engineered Nano Architectonics (FENA).

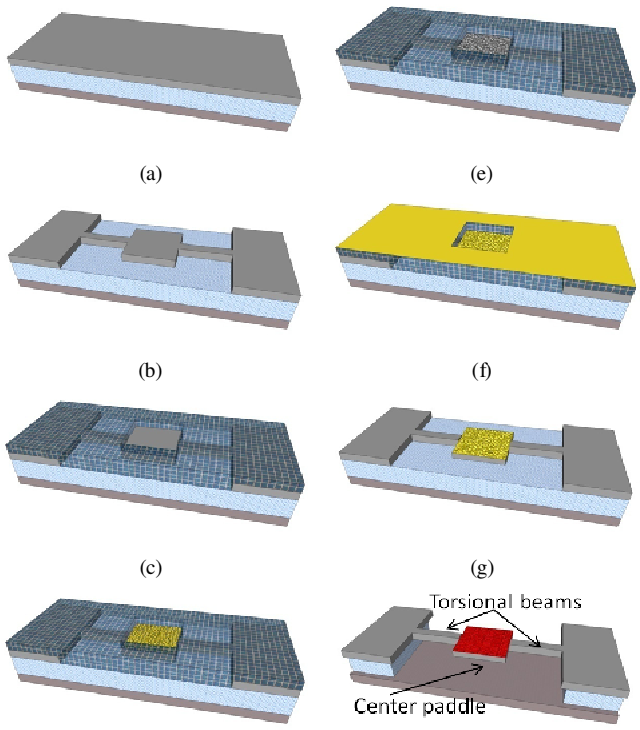


Fig. 1. Fabrication process for partially-porous silicon resonators. (a) Deposition of SiO_2 and polycrystalline silicon by LPCVD. (b) Photolithography and polycrystalline silicon etching by DRIE for defining body of the resonator. (c) Photolithography for opening of selective porous region. (d) Random gold cluster coating. (e) Porous silicon etching and gold layer removal. (f) Cr/Au deposition as linker of receptor material. (g) Liftoff of Cr/Au. (h) Sacrificial layer etching and receptor coating.

can be controlled. The remaining gold clusters were subsequently removed. After this, 25 nm-thick chromium and gold films were deposited for adhesion of the receptor coating. The electroless metal-assisted porous etching steps were omitted for the nonporous silicon resonators and the second photolithography for partial opening of the center paddle was omitted for fully-porous silicon resonators. Followed by liftoff of metal layers and sacrificial layer etch, the samples were coated with receptor solution. 4-mercaptobenzoic acid was used as a receptor material since it interacts with alcohol through hydrogen bonds [7].

Fig. 2 shows the change in roughness due to the porous etching process; this allows the porous silicon resonators to have a larger surface area than the non-porous resonators. According to the AFM root-mean-square surface roughness measurements, the surface area of the nonporous silicon and the porous silicon are 28.5 μm^2 and 158 μm^2, respectively, for a scan range of 5 µm ×5 µm. Depending on the porous etch time, the silicon surface becomes rough, and the total thickness gets reduced, as shown in Fig. 3.

III. RESULTS AND DISCUSSION

The frequency spectra of the resonant vapor sensors were measured by an optical beam deflection setup that consists of a laser Doppler vibrometer (LDV) system (Polytec OFV-5000), a custom-built vacuum chamber, and a network

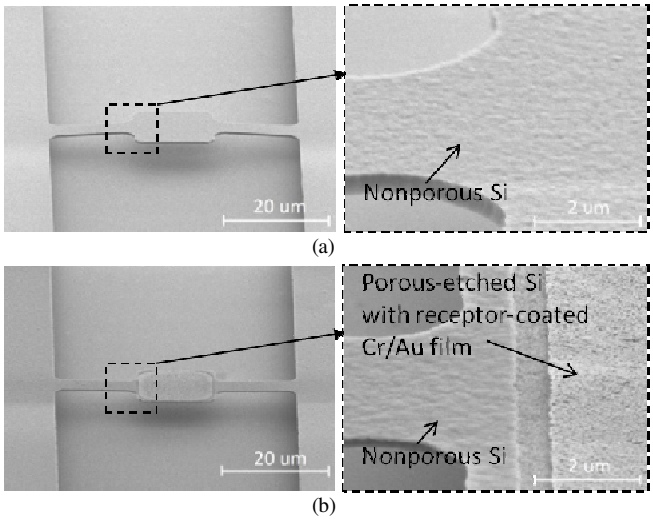


Fig. 2. SEM images of (a) the nonporous silicon resonator and (b) the partially-porous silicon resonator with the receptor coating.

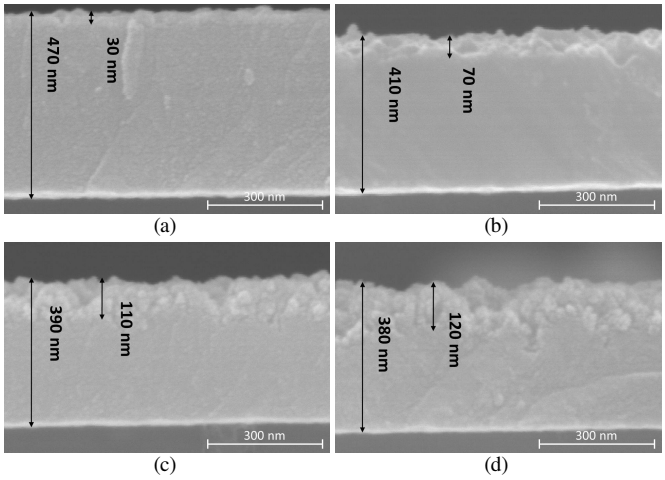


Fig. 3. Cross sectional SEM images of porous-etched silicon layer for different etch times: (a) 0 second (i.e. nonporous silicon), (b) 10 seconds, (c) 20 seconds, (d) 30 seconds.

analyzer (HP 8753D) as shown in Fig. 4. To induce mechanical vibration, the resonator sample was mounted on an external piezoelectric actuator (Thorlabs AE0203D08F) controlled by the stimulus signal from the network analyzer. Measurements were performed in a vacuum chamber in order to avoid energy losses by air damping. A stainless-steel reservoir that was filled with isopropyl alcohol (IPA) was connected to the vacuum chamber via a metering valve to control the vapor concentration. Resonator vibration was detected by reflection of the laser. The reflected beams were converted to an electrical signal by the LDV controller, and the network analyzer recorded the resonant frequency of the resonator mass sensor. Silicon resonators can detect IPA vapor since IPA adsorb to silicon surface via oxygen atoms [8].

The first two modes of oscillation for the resonator are a translational mode, corresponding to motion of a center paddle normal to a substrate, and a torsional mode, corresponding to rotation of the center paddle around axis of the torsional beams. The LDV system measures the difference of motion at

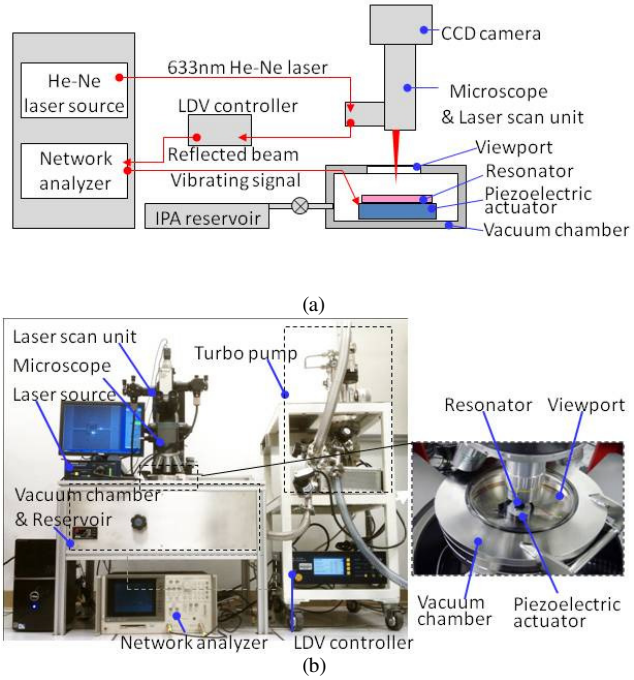


Fig. 4. (a) Schematic diagram. (b) Photos of the experimental setup for mechanical actuation and readout using a laser Doppler Vibrometer.

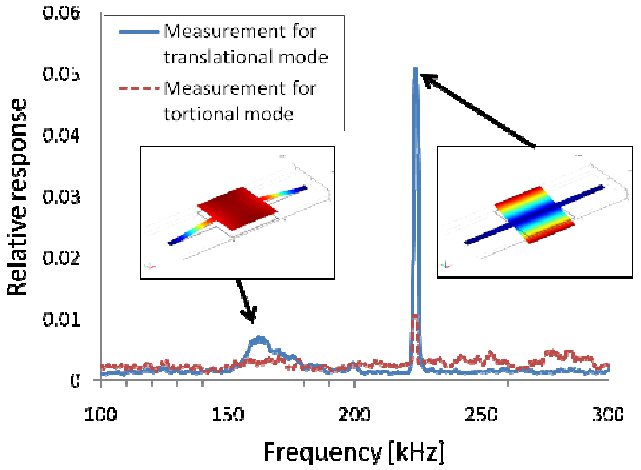


Fig. 5. A comparison of relative displacements of the translational mode and the torsional mode of the nonporous silicon resonator. The insets show schematics of the translational mode (left) and the torsional mode (right).

different two positions. One reflected beam for reference is positioned on the nonmoving pad structure, and the other is on the center of the paddle or edge to see the translational mode or torsional mode. As shown in Fig. 5, experimental results show the torsional mode has greater peaks than the translational mode. It is because torsional mode minimizes environmental damping, as the resonator undergoes less volume change [9]. In addition, torsional mode is known to be relatively insensitive to mechanical nonlinear effects due to the increase in drive amplitude [10].

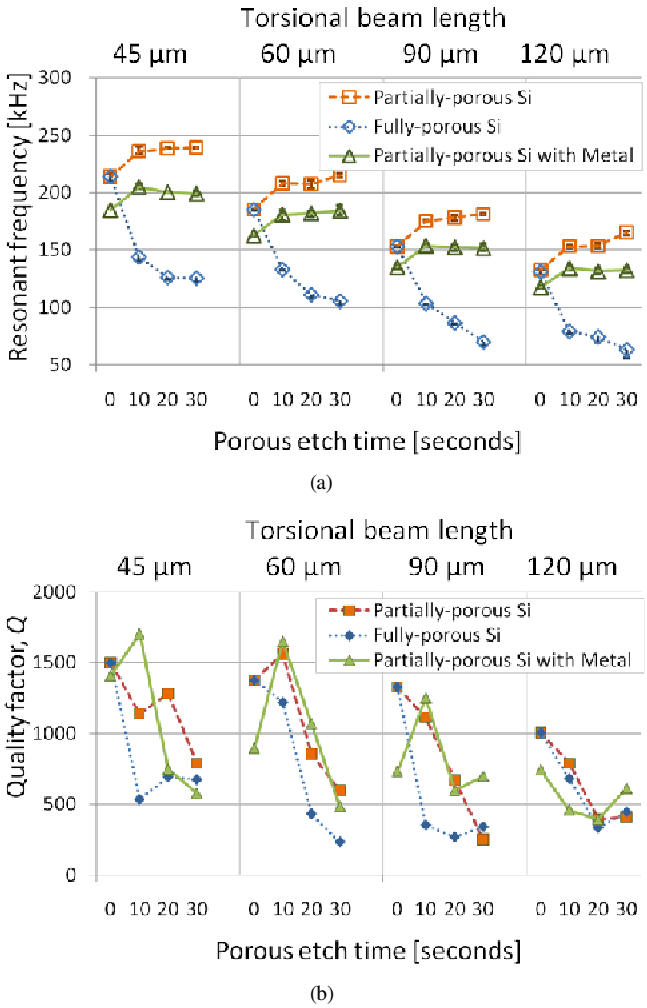


Fig. 6. Measured (a) resonant frequencies and (b) quality factors for different etch times and torsional beam lengths. The values of zero–second-etched condition indicate the case of nonporous silicon resonators.

Fig. 6 shows the dependence of resonant frequency and quality factor on the etch time for porosification (0-30 seconds) for resonators with different lengths of torsional beams (45-120 µm). As porous etch times increase, the resonant frequencies of fully-porous and partially-porous silicon structures have opposite tendencies because the partial porous structures primarily have a decrease in mass, whereas the fully-porous structures also have a large decrease in stiffness due to the porous beams. Partially-porous silicon structures with metal and receptor coating have the same trends with uncoated partially-porous silicon structures except having a frequency offset due to additional coating of metal and receptor. The quality factor decreases due to the increase in surface to volume ratio of resonators due to further porous etch. Also, the quality factors of these devices are lower than desired at ~1000. Enhancement of quality factors of these devices is an area of ongoing research.

Fig. 7 illustrates the different responses of each type of structure by the relative resonant frequency shift from different IPA concentration. The resonator sensitivities are calculated from the slopes of the normalized frequency shifts.

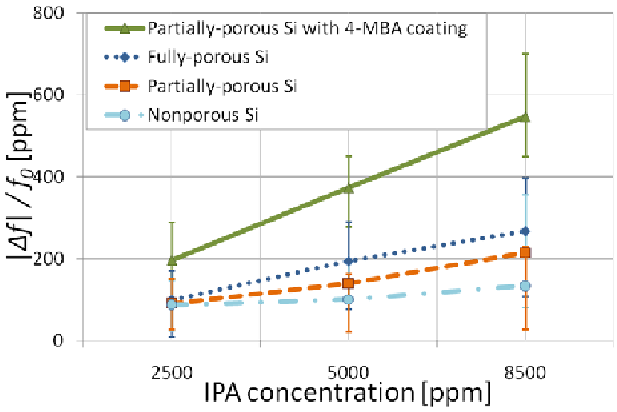


Fig. 7. A comparison of relative frequency shifts vs. gas concentration. The error bars indicate the standard deviation in the data, which comprised five measurements at five minute intervals for each concentration.

The sensitivities of partially-porous, fully-porous, and receptor coated partially-porous structure are improved to 261%, 165%, and 654%, respectively. The partially-porous structures have a slightly smaller improvement in sensitivity when compared to the fully-porous structure because the fully-porous structures have higher surface area. Another reason is the initial mass of the partially-porous silicon resonator is greater than the fully-porous resonator, which is relevant because sensitivity is dependent upon fractional mass change of adsorbed gas molecules.

Relative resolution is inversely proportional to the product of sensitivity and quality factor [11-12]. We demonstrate that the resolution was reduced to 44% of nonporous silicon structure in the case of the partially-porous silicon structure, and 26% for the receptor-coated partially-porous silicon structure. In addition, torsional beams of the fully-porous silicon structure which are already weakened by porosification limit further porosification for larger surface area, whereas porous paddle of the partially-porous silicon structure can be porous-etched more deeply in order to further enhance the sensitivity.

We have demonstrated enhanced chemical vapor sensitivity using partially porous functionalized silicon resonators by using a microscale device with nanoscale features. By keeping critical parts of the resonator non-porous, the mechanical stability of these resonators is improved over their fully porous counterparts. The combination of porous surface and functionalization for resonant sensing can be extended to many other adsorption-based sensing applications.

ACKNOWLEDGMENT

The authors gratefully thank Prof. C. J. Kim and C. Y. Lee for helping us to process the porous silicon etching. We would also like to thank Tom Lee of the Nanoelectronics Research Facility at UCLA for assistance in fabrication of the test chamber. Device fabrication was performed in the Nanoelectronics Research Facility (NRF) at UCLA.

REFERENCES

[1] Y. T. Yang, C. Callegari, X. L. Feng, K. L. Ekinci, and M. L. Roukes, "Zeptogram-scale nanomechanical mass sensing," *Nano Lett.,* vol. 6, no. 4, pp. 583-586, Apr. 2006.

[2] K. Jensen, K. Kim, and A. Zettl, "An atomic-resolution nanomechanical mass sensor," *Nature Nanotech.,* vol. 3, no. 9, pp. 533-537, Sep. 2008.

[3] B. Ilic, Y. Yang, K. Aubin, R. Reichenbach, S. Krylov, and H. G. Craighead, "Enumeration of DNA Molecules Bound to a Nanomechanical Oscillator," *Nano Lett.,* vol. 5, no. 5, pp. 925-929, 2005.

[4] K. L. Ekinci, Y. T. Yang, and M. L. Roukes, "Ultimate limits to inertial mass sensing based upon nanoelectromechanical systems," *J. Appl. Phys.,* vol. 95, no. 5, pp. 2682-2689, Mar. 2004.

[5] Y.-H. Hwang, F. Gao, and R. N. Candler, "Porous silicon resonators for sensitive vapor detection," in *Proc. Hilton Head Workshop*, Jun. 2010, pp. 150-153.

[6] C. Lee and C.-J. Kim, "Maximizing the Giant Liquid Slip on Superhydrophobic Microstructures by Nanostructuring Their Sidewalls," *Langmuir,* vol. 25, no. 21, pp. 12812-12818, Jul. 2009.

[7] J. Wang, W. G. Wu, Y. Huang, and Y. L. Hao, "< 100 > n-type metal-oxide-semiconductor field-effect transistor-embedded microcantilever sensor for observing the kinetics of chemical molecules interaction," *Appl. Phys. Lett.,* vol. 95, no. 12, p. 124101, Sep. 2009.

[8] J. A. Glass, E. A. Wovchko, and J. T. Yates, "Reaction of Methanol with Porous Silicon," *Surf. Sci.,* vol. 338, no. 1-3, pp. 125-137, Sep. 1995.

[9] D. Z. Jin, X. X. Li, H. H. Bao, Z. X. Zhang, Y. L. Wang, H. T. Yu, and G. M. Zuo, "Integrated cantilever sensors with a torsional resonance mode for ultraresoluble on-the-spot bio/chemical detection," *Appl. Phys. Lett.,* vol. 90, no. 4, p. 041901, Jan. 2007.

[10] S. Evoy, D. W. Carr, L. Sekaric, A. Olkhovets, J. M. Parpia, and H. G. Craighead, "Nanofabrication and electrostatic operation of single-crystal silicon paddle oscillators," *J. Appl. Phys.,* vol. 86, no. 11, pp. 6072-6077, Dec. 1999.

[11] A. Orsini and A. D'Amico, "Chemical sensors and chemical sensor systems: fundamentals limitations and new trends," in *Advances in Sensing with Security Applications*. vol. 2, J. Byrnes and G. Ostheimer, Eds., ed: Springer Netherlands, 2006, pp. 69-94.

[12] P. S. Waggoner and H. G. Craighead, "Micro- and nanomechanical sensors for environmental, chemical, and biological detection," *Lab Chip,* vol. 7, no. 10, pp. 1238-1255, Oct. 2007.

Effect of a mass layer on SH Waves in Piezomagnetic/Piezoelectric Material Structures

Jing Cui, Jianke Du, Ji Wang
Piezoelectric Device Laboratory, Department of Mechanics and Engineering Science
School of Engineering, Ningbo University, Ningbo, Zhejiang 315211, China
E-mail: dujianke@nbu.edu.cn

***Abstract*-An analytical approach is taken to investigate SH wave propagating in layered piezomagnetic/ piezoelectric structure where a thin piezomagnetic film with a mass layer bonded to a semi-infinite piezoelectric substrate. The dispersion relations are obtained and the effect of the thickness and the density of the mass layer on the properties of the SH waves are figured and presented.**

I. INTRODUCTION

It is well known that the composites made of piezoelectric/piezomagnetic materials exhibit magneto-electric effect that is not present in single-fields. Consequently, the coupling nature of piezoelectric phase and piezomagnetic phase has led to wide applications in adaptive material systems, and they are extensively used as sensors and actuators, acoustic devices, and transducers, etc. The development of micro-acoustic sensors provide the need for further investigations of the SH wave propagating in layered piezomagnetic/piezoelectric structure with a mass layer.

Soh and Liu [1] obtained the existence of the interfacial SH waves in transversely isotropic piezoelectric/ piezomagnetic structure, gave the conditions that the materials should be satisfied when the SH waves exist. Du et al. [2] introduced the SH waves propagating in a circular cylindrical layered piezoelectric structure with initial stress. Wang [3]and Liu [4] discussed a horizontal shear surface wave in magneto-electro-elastic materials at the same time.

In this paper, the effect of the thickness and the density of the mass layer on the properties of the SH waves propagating in layered piezomagnetic/piezoelectric structure is investigated analytically.

II. PROBLEM FORMULATION

Consider a layered piezomagnetic/piezoelectric structure where a thin piezomagnetic film with a mass layer bonded to a semi-infinite piezoelectric substrate. The piezomagnetic and piezoelectric materials are polarized along the z-axis direction. We here only consider the so-called anti-plane wave propagation problem. The displacement components, the electric potential and the magnetic potential are assumed as

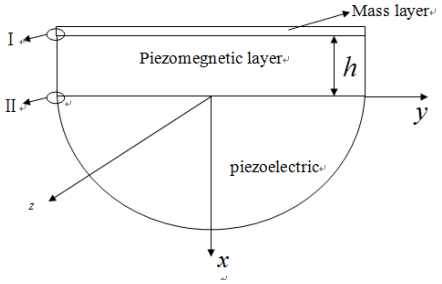


Fig.1 A layered magneto-electric structure

$$u(x,y)=0,\ \ v(x,y)=0,\ \ w=w(x,y,t),$$
$$\varphi=\varphi(x,y,t),\ \ \psi=\psi(x,y,t), \tag{1}$$

where u,v,w are the displacement components in x,y,z direction, respectively; ϕ is electric potential, and φ is magnetic potential.

For linearly magneto-electric materials, the equilibrium equations of elasticity without body forces and the Gauss's law of magneto-electrostatics without free charge are given as follows

$$\sigma_{ji,j}=\rho\ddot{u}_i,\ \ D_{i,i}=0,\ \ B_{i,i}=0,\ \ (i,j=1,2,3), \tag{2}$$

where σ_{ij},D_i,B_i are the stress, electric displacement and magnetic induction components, respectively. ρ is the mass density. The subscript comma denotes a partial derivative with respect to the coordinates, and a superimposed dot represents the derivative with respect to the time.

For an anisotropic and linearly magneto-electro-elastic solid, the coupled constitutive relation can be written as

$$\begin{aligned}\sigma_i &= c_{ik}\gamma_k - e_{ki}E_k - f_{ki}H_k,\\ D_i &= eik\gamma_k + \varepsilon_{ik}E_k + g_{ik}H_k,\\ B_i &= f_{ik}\gamma_k + g_{ik}E_k + \mu_{ik}H_k,\end{aligned} \tag{3}$$

where γ_i, E_i and H_i are the strain, electric field, and magnetic field components, respectively; c_{ij}, ε_{ij} and μ_{ij} are the elastic, dielectric, and magnetic permeability coefficients, respectively; e_{ij}, f_{ij} and g_{ij} are the piezoelectric, piezomagnetic, and magneto-electric coefficients, respectively. For a linearly transversely isotropic magneto-electric-elastic medium with the poling direction along the *z*-axis, the coupling wave equations can be obtained as

$$c_{44}^{*}\nabla^2 w = \rho\frac{\partial^2 w}{\partial t^2},\ \nabla^2\xi = 0,\ \nabla^2\zeta = 0, \tag{4}$$

where

$$c_{44}^{*} = c_{44} + \frac{e_{15}}{\mu_{11}\varepsilon_{11} - g_{11}^2}(\mu_{11}e_{15} - g_{11}f_{15}) + \frac{f_{15}}{g_{11}^2 - \varepsilon_{11}\mu_{11}}(g_{11}e_{15} - \varepsilon_{11}f_{15}),$$

and $\xi = e_{15}w - \varepsilon_{11}\varphi - g_{11}\psi,\ \zeta = f_{15}w - g_{11}\varphi - \mu_{11}\psi.$

$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the two-dimensional Laplacian operator in the Cartesian coordinates. The electric potential and magnetic potential can be written as

$$\begin{aligned}\varphi &= \frac{1}{\mu_{11}\varepsilon_{11} - g_{11}^2}\left[\left(\mu_{11}e_{15} - g_{11}f_{15}\right)w - \mu_{11}\xi + g_{11}\zeta\right],\\ \psi &= \frac{1}{g_{11}^2 - \varepsilon_{11}\mu_{11}}\left[\left(g_{11}e_{15} - \varepsilon_{11}f_{15}\right)w - g_{11}\xi + \varepsilon_{11}\zeta\right].\end{aligned} \tag{5}$$

The displacement, electric potential, and magnetic potential in the substrate tend to zero far from the layer along the positive *x*-direction

$$x \to +\infty,\ w_m = 0,\ \varphi_m = 0,\ \psi_m = 0, \tag{6}$$

The continuity conditions at the interface between the thin piezomagnetic layer and the piezoelectric substrate are written as

$$\begin{aligned}&w(0,y) = w^{(m)}(0,y),\ \tau_{zx}(0,y) = \tau_{zx}^{(m)}(0,y),\\ &\varphi(0,y) = \varphi^{(m)}(0,y),\ D_x(0,y) = D^{(m)}(0,y),\\ &\psi(0,y) = \psi^{(m)}(0,y),\ B_x(0,y) = B^{(m)}(0,y),\end{aligned} \tag{7}$$

where the superscript m indicates the quantities in the substrate. There are two typical cases being considered in this study, namely, the magneto electrically open and the shorted surface conditions. The magneto-electrically open conditions are written as

$$\tau_{zx}(-h,y) = \rho' h'\ddot{w},\ D_x(-h,y) = 0,\ B_x(-h,y) = 0. \tag{8}$$

The magneto electrically shorted conditions are given as

$$\tau_{zx}(-h,y) = \rho' h'\ddot{w},\ \varphi(-h,y) = 0,\ \psi(-h,y) = 0, \tag{9}$$

where ρ', h' are the density and the thickness of the mass layer, respectively.

III. SOLUTION OF THE EQUATIONS

A. Solutions in the piezomagnetic layer

We consider the following solution of (4)

$$\begin{aligned}w(x,y,t) &= w(x)\exp[ik(y-ct)],\\ \xi(x,y,t) &= \xi(x)\exp[ik(y-ct)],\\ \zeta(x,y,t) &= \zeta(x)\exp[ik(y-ct)],\end{aligned} \tag{10}$$

where $k = 2\pi/\lambda$ is the wave number, $i = \sqrt{-1}$, c is phase velocity. Substituting (10) into (4), we can obtain

$$\begin{aligned}&w''(x) + k^2b^2w(x) = 0,\\ &\xi''(x) - k^2\xi(x) = 0,\\ &\zeta''(x) - k^2\zeta(x) = 0,\end{aligned} \tag{11}$$

where $b^2 = \rho c^2/c_{44}^{*} - 1 > 0$. The solutions are

$$\begin{aligned}w(x) &= Ae^{-ikbx} + Be^{ikbx},\\ \xi(x) &= Ce^{-kx} + De^{kx},\\ \zeta(x) &= Ee^{-kx} + Fe^{kx}.\end{aligned} \tag{12}$$

So for the piezomagnetic layer, we can summarize

$$\begin{aligned}&w(x,y,t) = [A\cos(kbx) + B\sin(kbx)]\cos(ky - \omega t),\\ &\varphi(x,y,t) = \frac{1}{\mu_{11}\varepsilon_{11} - g_{11}^2}\begin{Bmatrix}(\mu_{11}e_{15} - g_{11}f_{15})[A\cos(kbx) + B\sin(kbx)]\\ -\mu_{11}(Ce^{-kx} + De^{kx}) + g_{11}(Ee^{-kx} + Fe^{kx})\end{Bmatrix}\cos(ky - \omega t),\\ &\psi(x,y,t) = \frac{1}{g_{11}^2 - \varepsilon_{11}\mu_{11}}\begin{Bmatrix}(g_{11}e_{15} - \varepsilon_{11}f_{15})[A\cos(kbx) + B\sin(kbx)]\\ -g_{11}(Ce^{-kx} + De^{kx}) + \varepsilon_{11}(Ee^{-kx} + Fe^{kx})\end{Bmatrix}\cos(ky - \omega t),\\ &D_x(x,y,t) = k(-Ce^{-kx} + De^{kx})\cos(ky - \omega t),\\ &B_x(x,y,t) = k(-Ee^{-kx} + Fe^{kx})\cos(ky - \omega t),\end{aligned} \tag{13}$$

where A, B, C, D, E, F are unknowns to be determined.

B. Solutions in the piezoelectric substrate

The displacement, electric potential and magnetic potential are given as

$$\begin{aligned}w^{(m)}(x,y,t) &= A_1e^{-kb^m x}\exp[ik(y-ct)],\\ \xi^{(m)}(x,y,t) &= C_1e^{-kx}\exp[ik(y-ct)],\\ \zeta^{(m)}(x,y,t) &= E_1e^{-kx}\exp[ik(y-ct)],\end{aligned} \tag{14}$$

where $b^{m2} = 1 - \rho^m c^2/c_{44}^{m^8} > 0$, A_1, C_1, E_1 are unknown constants. So for the piezoelectric substrate, we can summarize

$$w^{(m)}(x,y,t) = A_1e^{-kb^m x}\cos(ky - \omega t),$$

$$\phi^{(m)}(x,y,t) = \frac{1}{\mu_{11}^m\varepsilon_{11}^m - g_{11}^{m2}}\begin{bmatrix}A_1(\mu_{11}^m e_{15}^m - g_{11}^m f_{15}^m)e^{-kb^m x}\\ -C_1\mu_{11}^m e^{-kx} + E_1 g_{11}^m e^{-kx}\end{bmatrix}\cos(ky - \omega t),$$

$$\psi^{(m)}(x,y,t)=\frac{1}{g_{11}^{m2}-\varepsilon_{11}^{m}\mu_{11}^{m}}\begin{bmatrix}(g_{11}^{m}e_{15}^{m}-\varepsilon_{11}^{m}f_{15}^{m})A_1e^{-kb^m x}\\ -g_{11}^{m}C_1e^{-kx}+\varepsilon_{11}^{m}E_1e^{-kx}\end{bmatrix}\cos(ky-\omega t),$$

$$D_x^{(m)}(x,y,t)=-C_1ke^{-kx}\cos(ky-\omega t),$$

$$B_x^{(m)}(x,y,t)=-E_1ke^{-kx}\cos(ky-\omega t). \tag{15}$$

IV. SOLUTION OF THE PROBLEM

From (7) we can obtain

$$\begin{aligned}
&A-A_1=0,\\
&Bm_1b-Cm_2+Dm_2-Em_3+Fm_3\\
&+A_1b^m m_1^m+C_1m_2^m+E_1m_3^m=0,\\
&A(\mu_{11}e_{15}-g_{11}f_{15})(\mu_{11}^m\varepsilon_{11}^m-g_{11}^{m2})-C\mu_{11}(\mu_{11}^m\varepsilon_{11}^m-g_{11}^{m2})\\
&-D\mu_{11}(\mu_{11}^m\varepsilon_{11}^m-g_{11}^{m2})+Eg_{11}(\mu_{11}^m\varepsilon_{11}^m-g_{11}^{m2})+\\
&Fg_{11}(\mu_{11}^m\varepsilon_{11}^m-g_{11}^{m2})-A_1(\mu_{11}\varepsilon_{11}-g_{11}^2)(\mu_{11}^m e_{15}^m-g_{11}^m f_{15}^m)\\
&+C_1\mu_{11}^m(\mu_{11}\varepsilon_{11}-g_{11}^2)+E_1g_{11}^m(\mu_{11}\varepsilon_{11}-g_{11}^2)=0,\\
&-C+D+C_1=0,\\
&A(g_{11}e_{15}-\varepsilon_{11}f_{15})(g_{11}^{m2}-\varepsilon_{11}^m\mu_{11}^m)-Cg_{11}(g_{11}^{m2}-\varepsilon_{11}^m\mu_{11}^m)\\
&-Dg_{11}(g_{11}^{m2}-\varepsilon_{11}^m\mu_{11}^m)+E\varepsilon_{11}(g_{11}^{m2}-\varepsilon_{11}^m\mu_{11}^m)\\
&+F\varepsilon_{11}(g_{11}^{m2}-\varepsilon_{11}^m\mu_{11}^m)-A_1(g_{11}^m e_{15}^m-\varepsilon_{11}^m f_{15}^m)(g_{11}^2-\varepsilon_{11}\mu_{11})\\
&+C_1g_{11}^m(g_{11}^2-\varepsilon_{11}\mu_{11})-E_1\varepsilon_{11}^m(g_{11}^2-\varepsilon_{11}\mu_{11})=0,\\
&-E+F+E_1=0,
\end{aligned} \tag{16}$$

1) The magneto-electrically shorted conditions:

$$\begin{aligned}
&(\mu_{11}e_{15}-g_{11}f_{15})[A\cos(kbh)-B\sin(kbh)]\\
&-\mu_{11}(Ce^{kh}+De^{-kh})+g_{11}(Ee^{kh}+Fe^{-kh})=0,\\
&(g_{11}e_{15}-\varepsilon_{11}f_{15})[A\cos(kbh)-B\sin(kbh)]\\
&-g_{11}(Ce^{kh}+De^{-kh})+\varepsilon_{11}(Ee^{kh}+Fe^{-kh})=0,\\
&m_1[Ab\sin(kbh)+Bb\cos(kbh)]+m_2(-Ce^{kh}+De^{-kh})\\
&+m_3(-Ee^{kh}+Fe^{-kh})=\rho' h'kc^2[-A\cos(kbh)+B\sin(kbh)].
\end{aligned} \tag{17}$$

In order to obtain the nontrivial solutions of the above-mentioned unknown constants A,B,C,D,E,F,A_1,C_1,E_1, the determinant of the coefficient matrix of liner algebraic equations (16), (17) have to be equal to zero.

2) The magneto-electrically opened conditions:

$$\begin{aligned}
&k(-Ce^{kh}+De^{-kh})=0,\\
&k(-Ee^{kh}+Fe^{-kh})=0,\\
&m_1[Ab\sin(kbh)+Bb\cos(kbh)]+m_2(-Ce^{kh}+De^{-kh})\\
&+m_3(-Ee^{kh}+Fe^{-kh})=\rho' h'kc^2[-A\cos(kbh)+B\sin(kbh)],
\end{aligned} \tag{18}$$

In order to obtain the nontrivial solutions of the above-mentioned unknown constants A,B,C,D,E,F,A_1,C_1,E_1, the determinant of the coefficient matrix of liner algebraic equations (16), (18) have to be equal to zero.

V. NUMERICAL RESULTS AND DISCUSSION

TABLE I. MATERIAL COEFFICIENTS OF THE PIEZOELECTRIC $BaTiO_3$

$c_{44}/10^9 N\cdot m^{-2}$	$f_{15}/N\cdot A^{-1}\cdot m^{-1}$	$\varepsilon_{11}/10^9 C^2\cdot N^{-1}\cdot m^{-2}$
45.3	550	0.08
$\mu_{11}/10^{-6} N\cdot S^2\cdot C^{-2}$	$\rho/10^3 kg\cdot m^{-3}$	$g_{11}/10^{-9} Ns/VC$
-590	5.3	-30.67

TABLE II. MATERIAL COEFFICIENTS OF THE MAGNETOSTRICTIVE $CoFe_2O_4$

$c_{44}/10^9 N\cdot m^{-2}$	$e_{15}/C\cdot m^{-2}$	$\varepsilon_{11}/10^9 C^2\cdot N^{-1}\cdot m^{-2}$
43	11.6	11.2
$\mu_{11}/10^{-6} N\cdot S^2\cdot C^{-2}$	$\rho/10^3 kg\cdot m^{-3}$	$g_{11}/10^{-9} Ns/VC$
5	5.8	-30.67

We assume that the thickness of the piezomagnetic layer is 0.1mm in calculation. It is supposed that the unknown constant $A=10^{-9}m$ to make the displacement results understood easily.

Figs.2-5 are presented here in order to illustrate the effect of the thickness of the mass layer on the phase velocity and the displacements. We assume the density of the mass layer is $3.6\times10^3 kg/m^3$. We can find that both velocity and the displacements decrease with increase of thickness of the layer.

Figs.6-9 are presented here in order to illustrate the effect of the density of the mass layer on the velocity and the displacements. We assume the thickness of the mass layer is $h'=0.0002mm$. The results show that the velocity and the displacements also decrease with increase of the density of the layer.

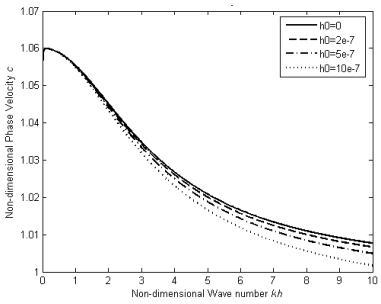


Figure 2. Phase velocity for magneto-electrically shorted case

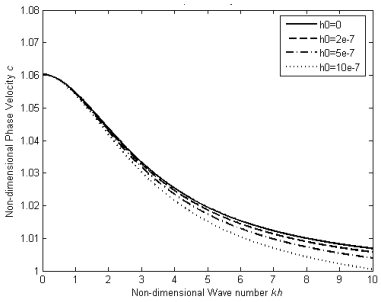


Figure 3. Phase velocity for magneto-electrically open case

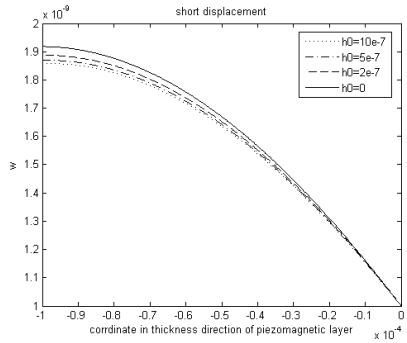


Figure 4. Displacement for magneto-electrically shorted case

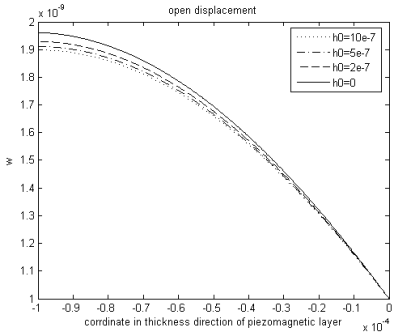


Figure 5. Displacement for magneto-electrically open case

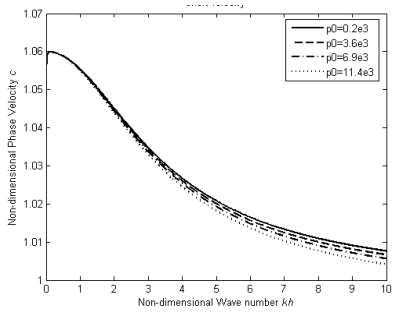


Figure 6. Phase velocity for magneto-electrically shorted case

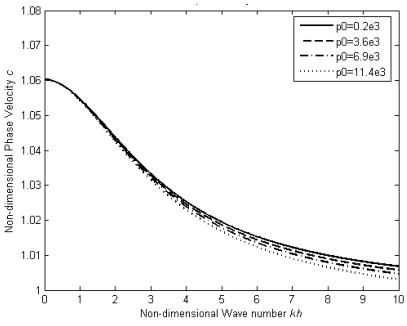


Figure 7. Phase velocity for magneto-electrically open case

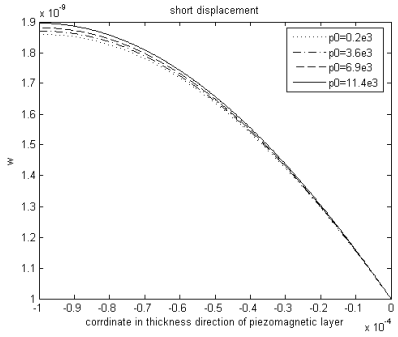


Figure 8. Displacement for magneto-electrically short case

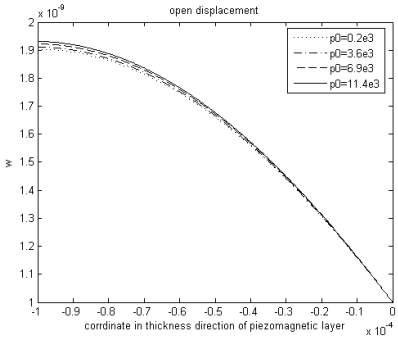


Figure 9. Displacement for magneto-electrically open case

VI. CONCLUSIONS

An analytical approach is taken to investigate SH wave propagating in layered piezomagnetic/ piezoelectric structure where a thin piezomagnetic film with a mass layer bonded to a semi-infinite piezoelectric substrate. The results show that the thickness and the density of the mass layer have significant effect on the phase velocity of SH waves and the distribution of the displacements. The results show that the phase velocity and displacements decrease with the increase of the thickness and the density of the mass layer for both electrically opens and shorted cases.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Nos. 11072116 and 10932004), Key Industrial Project of Bureau of Science and Technology, City of Ningbo (No.2005B100015).

REFERENCES

[1] Soh. A K and Liu J X, "Interfacial shear horizontal waves in a piezoelectric-piezomagnetic bi-material," J. Philosophical Magazine Letters, 2006, pp.31-35.

[2] Jianke Du, Ji Wang, and Xiaoying Jin, "SH Waves in a Circular Cylindrical Layered Piezoelectric Structure with Initial Stress," IEEE International Frequency Control Symposium, June 5-7, Miami, FL, USA, 2006, pp.93-96.

[3] Wang B L, Mai Y W, and Niraula O P, "A horizontal shear surface wave in magnetoelectroelastic materials," J. Philosophical Magazine Letters, 2007, pp.53-58.

[4] Liu J X, Fang D N, and Liu X L. "A shear horizontal surface wave in magnetoelectric materials," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2007, pp.1287-1289.

Progress on an Optical Link for Ultra-Stable Frequency Dissemination using a Public Telecommunication Network

O. Lopez, B. Chanteau, V. Roncin, Ch. Chardonnet, A. Amy-Klein
Laboratoire de Physique des Lasers,
CNRS, Université Paris 13,
93430 Villetaneuse, France
anne.amy-klein@univ-paris13.fr

A. Haboucha, G. Santarelli

LNE-SYRTE,
Observatoire de Paris, CNRS, UPMC,
75014 Paris

***Abstract*— We transfer the frequency of an ultra-stable laser over a 470-km optical link comprising 400 km of public telecommunication fiber network, simultaneously carrying data traffic. A dense wavelength division multiplexing scheme is used. The ultrastable signal is inserted in and extracted from the communication network using bidirectional optical add-drop multiplexers. The link passes through two important nodes of the telecommunication network and an amplification stage. The phase noise on the optical link is compensated with the usual round-trip technique. The 470-km multiplexed link shows an Allan deviation of 5×10^{-15} at one second and 1.3×10^{-18} after 1 hour. This work paves the way to a wide dissemination of ultra stable optical clock signals between distant laboratories via the Internet network.**

I. Introduction

The transfer of ultra-stable frequency signal between distant laboratories is required by many applications in time and frequency metrology, fundamental physics, particle accelerators and astrophysics [1-8]. Stable radio and microwave frequencies transmission over an optical link has already been demonstrated [2-5,9-12]. Frequency transfer using the optical phase of an ultra-stable laser over a dedicated fiber link was reported on distances up to about 200 km by several groups [13-19]. The present challenge is to extend this technique of frequency dissemination on longer distances in order to connect laboratories of different countries.

For this purpose, we have recently developed a novel dissemination approach over non-dedicated fibers [20, 21]. We take advantage of the existing Internet fiber network, already connecting every laboratories via the National Research Networks. The ultra-stable frequency signal is propagating simultaneously with the Internet traffic in the same fiber using one dedicated wavelength in a dense wavelength division multiplexing (DWDM) approach.

With Internet fibers, we have a very limited control on the fiber network and the attenuation and noise is likely to be higher than with dedicated fibers, which can limit the transfer to a few hundreds of km. For longer distances, we have foreseen the segmentation of the link into several cascaded sections. In that case, a repeater station should be used between the different segments of the link. This multiple sub-link approach allows for an increased correction bandwidth and robustness regarding attenuation.

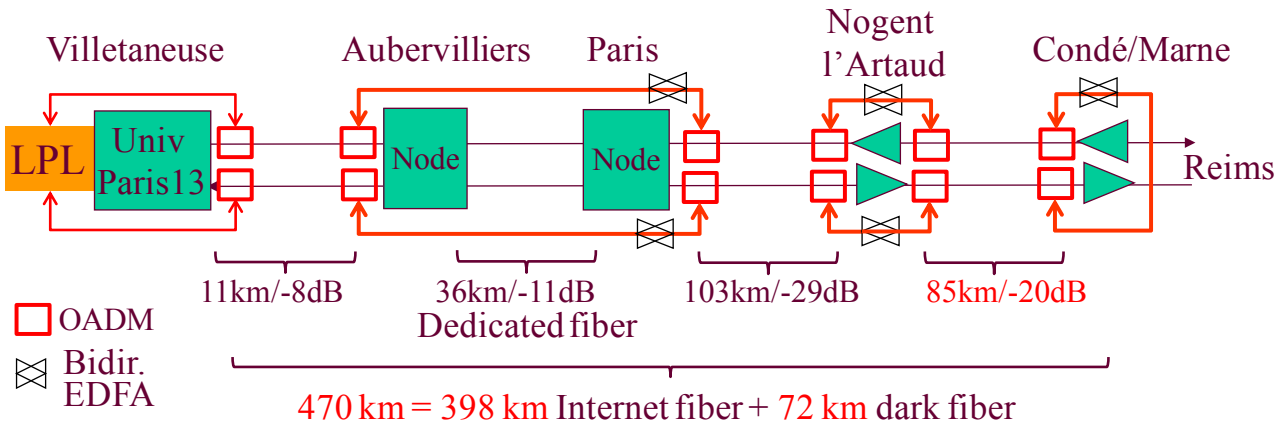


Figure 1. Scheme of the 470 km multiplexed optical link

II. MULTIPLEXED OPTICAL LINK

We have demonstrated a first implementation of a 470-km multiplexed optical link using the French National Telecommunication network for Technology, Education and Research (RENATER).

The overall scheme of the 470 km-long optical link (LPL-Condé sur Marne-LPL) is depicted in Fig. 1. Concerning the ultrastable signal transmission, round-trip propagation of the optical signal is required over the same fiber for noise compensation. For that purpose we use bidirectional off-the-shelf optical add-drop multiplexers (OADM). Such three-port component can insert or extract the 1542.14 nm wavelength from the other wavelengths, with isolation better than 25 dB for an adjacent channel (100 GHz) and better than 40 dB for other channels. Twelve OADMs are used along the link to insert and extract the ultrastable signal.

This link starts and ends at LPL laboratory, at Université Paris 13, in the north surroundings of Paris. It is composed of three different fiber spans. The first span is composed by two 11 km-long fibers connecting the information service and technology center of Université Paris 13 to a Data Centre Facility located in Aubervilliers (Interxion1). The digital stream between Université Paris 13 and Aubervilliers is encoded over an optical carrier on the channel #34 (1550.12 nm) of the International Telecommunication Union grid (ITU) whereas the ultrastable signal is carried by the channel #44, at 1542.14 nm. The second span is composed of two 36 km-long dedicated dark fibers which connect the two Data Centre Facilities of Interxion 1 and TeleHouse 2, downtown Paris. The third span is composed of two 188 km longhaul intercity fibers simultaneously carrying internet data traffic. In this part the digital data signal are transferred over two channels #42 and #43, only 0.8 nm and 1.6 nm away from the ultrastable signal. Moreover, this span includes two amplification stages for the data transfer, at Nogent l'Artaud. We have installed two OADMs linked with a bidirectional amplifier to bypass this amplification stage, enabling a selective bidirectional amplification of the ultrastable signal. At Condé sur Marne the reference signal arriving from one fiber is sent into the other 235 km-long fiber back to Université Paris 13. Figure 2 depicted the OADMs and amplifiers modules installed at Condé sur Marne.

Figure 2. OADMs and bidirectional amplifiers inside the shelter of Condé sur Marne

Total attenuation along the link is more than 136 dB due to the large number of connectors and the OADMs. In order to overcome the round-trip link losses, we used five bidirectional Erbium-doped fiber amplifiers.

III. EXPERIMENTAL RESULTS

The optical phase noise power spectral density of the 470-km optical link is shown in Fig. 3, without and with compensation. The phase noise reduction is around 40 dB at 1 Hz.

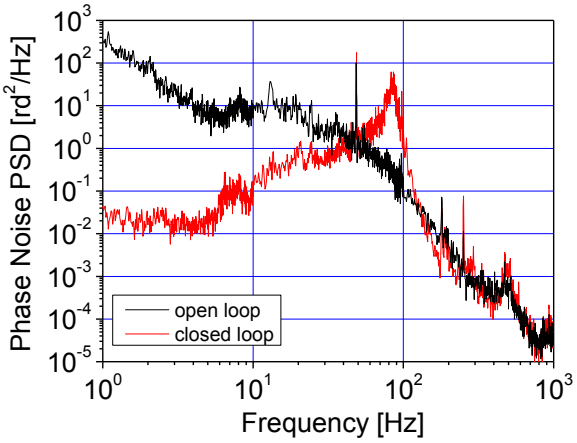


Figure 3. Phase noise power spectral density of the free running (black line) and compensated 470-km link (red line)

Fig. 4 shows the fractional frequency stability (Allan deviation) of the 470 km link for about 15 hours of continuous operation, measured with a Π-type frequency counter. The free-running fiber frequency noise (black circles) is measured simultaneously using the compensation signal applied to the AOM. For a measurement bandwidth of 10Hz, the Allan deviation is 5×10^{-15} at 1 s averaging time and scales down as $1/\tau$ from 1 s to 8000 s reaching 7×10^{-19}. With the full bandwidth the Allan deviation is 5 times larger.

IV. CONCLUSION

We demonstrate the ultra stable transfer of an optical frequency over 470 km of installed optical fibers. This ultrastable optical link uses an optical telecommunication network simultaneously carrying Internet data and goes through one amplification stage and two Data Center Facilities using multiplexers and bidirectional erbium-doped fiber amplifiers. We have obtained link instability of 5×10^{-15} at 1 s

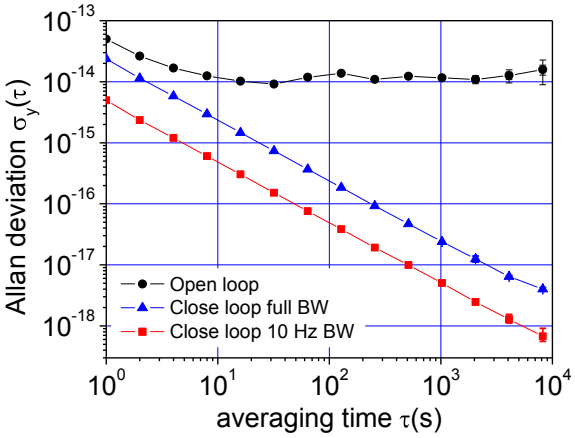


Figure 4. End-to-end fractional frequency instability of the 470-km free running optical link (black circles), and compensated link measured without (blue triangles) and with a 10 Hz filter (red squares)

which scales down to around 1.3×10^{-18} after 1 hour, in a measurement bandwidth of 10 Hz. These results are promising and represent an intermediate step for the future development of continental-scale frequency transfer.

Acknowledgment

The authors are deeply grateful to D. Vandromme, L. Gydé, T. Bono and E. Camisard from GIP RENATER for their support in using the Renater network. We acknowledge funding support from IFRAF, Université Paris 13 and LNE.

REFERENCES

[1] S. M. Foreman, K. W. Holman, D. D. Hudson, D. J. Jones, and J. Ye, "Remote transfer of ultrastable frequency references via fiber networks"Rev. Sci. Instrum. vol. 78, pp. 021101, 2007.

[2] C. Daussy et al, "Long-distance frequency dissemination with a resolution of 10^{-17}", Phys. Rev. Lett., vol. 94, pp. 203904, 2005.

[3] M. Calhoun, S. Huang, and R. L. Tjoelker, "Stable Photonic Links for Frequency and Time Transfer in the Deep-Space Network and Antenna Arrays", Proc. of the IEEE, Special Issue on Technical Advances in Deep Space Communications & Tracking, vol. 95 , pp. 1931-1946, 2007.

[4] B. Shillue, S. AlBanna and L. D'Addario, "Transmission of low phase noise, low phase drift millimeter-wavelength references by a stabilized fiber distribution system", Proc. IEEE Int. Top. Meet. Microw. Photon. (MWP 2004), pp. 201-204, 2004.

[5] R. Wilcox, J. M. Byrd, L. Doolittle, G. Huang, and J. W. Staples, "Stable transmission of radio frequency signals on fiber links using interferometric delay sensing," Opt. Lett. vol. 34, pp. 3050-3052, 2009

[6] S. G. Karshenboim, "Fundamental physical constants: looking from different angles", Can. J. Phys., vol. 83, pp. 767-813, 2005.

[7] V. V. Flambaum, "Enhanced effect of temporal variation of the fine-structure constant in diatomic molecules", Phys. Rev. D, vol. 69, pp. 115006, 2004.

[8] J. P. Uzan, "The fundamental constants and their variation: observational and theoretical status", Rev. Mod. Phys. vol. 75, pp. 403-455, 2003.

[9] O. Lopez et al, "86-km optical link with a resolution of 2×10^{-18} for RF frequency transfer", Eur. Phys. J. D, vol. 48, pp. 35-41, 2008.

[10] M. Fujieda, M. Kumagai, and S. Nagano "Coherent Microwave Transfer Over a 204-km Telecom Fiber Link by a Cascaded System", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, pp. 168-174, 2010.

[11] M. Kumagai, M. Fujieda, S. Nagano, and M. Hosokawa, "Stable radio frequency transfer in 114 km urban optical fiber link," Opt. Lett. vol. 34, pp. 2949-2951, 2009.

[12] O. Lopez, A. Amy-Klein, M. Lours, Ch. Chardonnet, and G. Santarelli, "High-resolution microwave frequency dissemination on an 86-km urban optical link", Appl. Phys. B, vol. 98, pp. 723-727, 2010.

[13] N. R. Newbury, P. A. Williams, W. C. Swann, "Coherent transfer of an optical carrier over 251 km", Opt. Lett., vol. 32, pp. 3056-3058, 2007.

[14] P. A. Williams, W. C. Swann, and N. R. Newbury, "High-stability transfer of an optical frequency over long fiber-optic links", J. Opt. Soc. Am. B, vol. 25, pp. 1284-1293, 2008.

[15] M. Musha, F. Hong, K. Nakagawa, and K. Ueda, "Coherent optical frequency transfer over 50-km physical distance using a 120-km-long installed telecom fiber network", Opt. Express, vol. 16, pp. 16459-16466, 2008.

[16] H. Jiang et al, "Long-distance frequency transfer over an urban fiber link using optical phase stabilization", J. Opt. Soc. Am. B, vol. 25, pp. 2029-2035, 2008.

[17] G. Grosche et al, "Optical frequency transfer via 146 km fiber link with 10^{-19} relative accuracy", Opt. Lett., vol. 34, pp. 2270-2272, 2009.

[18] O. Terra et al, "Phase-coherent comparison of two optical frequency standards over 146 km using a telecommunication fiber link", Appl. Phys. B, vol. 97, pp. 541-551, 2009.

[19] H. Schnatz, et al. "Phase-Coherent Frequency Comparison of Optical Clocks Using a Telecommunication Fiber Link", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, pp. 175-181, 2010.

[20] F. Kéfélian, O. Lopez, H. Jiang, Ch. Chardonnet, A. Amy-Klein and G. Santarelli, "High-resolution optical frequency dissemination on a telecommunication network with data traffic", Opt. Lett, vol. 34, pp. 1573-1575, 2009.

[21] O. Lopez, A. Haboucha, F. Kéfélian, H. Jiang, B. Chanteau, V. Roncin, C. Chardonnet, A. Amy-Klein and G. Santarelli, "Cascaded multiplexed optical link on a telecommunication network for frequency dissemination", Opt. Expr., vol. 18, pp 16849-16857, 2010.

Active Detection of Propagation Delay Variations in Single Way Time Transfer Utilizing Dual Wavelengths in an Optical Fiber Network

Sven-Christian Ebenhag*, Per Olof Hedekvist* and Kenneth Jaldehag
Measurement Technology, Communications Lab
SP Technical Research Institute of Sweden
Borås, Sweden
*also with Chalmers University of Technology
Göteborg, Sweden
Email: sven-christian.ebenhag@sp.se

Abstract— **Several communication systems of today rely on the real time accessibility of accurate time and frequency measures and there is an increasing demand for the development of new and redundant methods for the distribution of these measures. The classical two-way method is able to compensate for the inevitable variations in the time and frequency propagation delay. The two-way method is used for time transfer in free space, electrical or optical domain, but has the disadvantage of often using two different paths for transmitting back and forward. The paths may be of equal length and have equal propagation delay, but nevertheless there is often a remaining asymmetry in the propagation paths. The inevitable asymmetry between the paths in the time transfer delay must be detected and compensated for, if an accuracy better than µs-level is needed for transmission distance exceeding a few km. Furthermore, if the number of users is high, there will be a complex and large network of two-way time signal transmissions. Therefore, a solution using one-way broadcasting would be more desirable, and would be possible if the variations in transmission time could be estimated from the received data at the far (user) end. The one-way method uses only one path of transmission and is possible to implement in existing Wavelength Division Multiplexing-networks.**

Proof of concept and results of this one-way time transfer technique based on transmission of a repetitive signal, modulated on two lasers at different wavelengths 8 nm apart and transmitted through an optical fiber, has been presented previously. These data showed a strong correlation between a change in transfer time at one wavelength, and the transfer time difference for the signals at the two wavelengths.

In this paper, the setup and the measurement results have been improved and new data is collected which shows improvement in the reliability and quality of this technique. The stability is improved through component analysis and minimizing error sources. The distance is improved from 38km to 160km.

I. INTRODUCTION

The development of new communication systems and evolution of existing ones has increased the need for precise time and frequency transfer. These transfers often exceed baselines of 100km and GPS is the common technical solution. The main benefit with GPS is that it is easy to install, but it is a single point of failure. Time and frequency transfer utilizing optical fiber is a good alternative to GPS since there in many countries already are existing networks and it does not rely on transmission over radio waves in open air. Several transfer methods using optical fibers has been developed or are under development **[1-4]**, using dedicated fibers or already existing fiber networks. Performance is often the limiting aspect in the choice of transfer method, where the long term stability relies on an accurate estimate of the variations of transfer delay. A common way for high performance transfer is the two-way method, which is an excellent method when the user has easy access to the whole system and when both transmission paths are equal. For the best results both directions in the transfer should operate in the same transmission line to be able to cancel out transmission path delays.

The method can be implemented very efficiently when

978-1-61284-110-6/11/$26.00 ©2011 IEEE

operating on existing data traffic **[5-6]**, however when data format or other limitations require that the time and frequency transfer must operate on dedicated communication channels, the techniques' administration scales with the number of users and becomes capacity consuming.

This paper presents a technique for one-way time transfer utilizing dual wavelengths and the difference in group velocity to estimate the delay variation of the timing signal in one of the wavelength channels. The first proof-of-concept was previously presented **[7-8]**, using modulated lasers at 1310nm and 1550nm and **[9]** wavelengths at 1535 and 1553nm. Those results showed that it is possible to perform a one-way time and frequency transfer with two wavelengths and by evaluate these two against each other, create a correction signal for compensation for influences along the transmission path.

To enable long distance baselines, the optical signal must be possible to amplify. It is also desirable to cover both wavelengths inside the optical C-band (1530 – 1560nm) and for the purposes of using the existing fiber network operate within one band, i.e 8nm.

This evaluation is therefore performed with two wavelengths 8nm apart and the presented result show that it is possible to perform a one way time and frequency transfer if two wavelengths within the same transmission band are used. By evaluation of these two against each other, is it possible to create a correction signal for compensation of transmission time variations using a 160km fiber including Erbium doped fiber amplifiers (EDFAs).

II. BACKGROUND

The occurrence of asymmetry in many two-way transfers is due to many components such as ageing, connector connections and different length in the transfer paths and equipment in and between the nodes. With this classic transfer method as reference point, this paper will discuss a different method based on a single path.

Standardization of single way time transfer is done in the IRIG-B system, and there are commercial fiber based frequency transfer systems available. Neither of these does however take variations in transfer time into account, mainly because they are supposed to work over short distances and not with high time transfer accuracy.

When the distance increase and a large amount of the transmission media is outdoors the temperature will vary along the path as well as the propagation properties of the material. Even fiber below ground is known to experience temperature variations of 20°C over the year **[10]**, and if some sections of the fiber distance are installed along the power grid through air-lines, the propagation delay variation will be even worse, as the variations will follow the air temperature. However, if that is the case, not only the transfer time changes, but also the dispersion, which can be measured at the receiving end of the fiber, and the correlation between these two parameters can be used. The measurement technique for fiber dispersion is well known **[11]** and the variation with respect to temperature have been studied previously **[10, 12].** Nevertheless, the correlation to the time transfer delay has not yet been fully investigated or utilized for time transfer until the first presentations of this work **[7]**.

In a fully operational solution, the time from the Master clock can be distributed to a network of Slave clocks, with a precision better than what would be the case if a single signal was transmitted.

For the one-way time transfer illustrated in Figure 1, there is a time varying part associated with ΔT_A as also is included in Equation (1). Even though the average transfer time can be taken into account through a calibration, the time varying part will pose as an uncertainty unless it continuously is measured and compensated for.

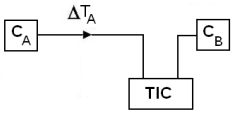


Figure 1. The figure is a sketch of a one-way time transfer between clock A (C_A) and clock B (C_B).

$$C_A\text{-}C_B=TIC\text{-}\Delta T_A(t) \quad (1)$$

A proposed method to solve for $\Delta T_A(t)$ is by transmitting two wavelengths in the same fiber and detect the propagation delay differences between these **[13-16]**. The difference is then applied as a correction term for the delay variations in $\Delta T_A(t)$ **[7-9]**. Development of this one-way transfer method with enhanced stability has resulted in transfer quality well comparable with well established two-way methods.

Furthermore it is operating within the C-band of commercial Dense Wavelength Division Multiplexing-networks (DWDM-network) for distances >100km. The reason for using the C-band is that the commonly used Erbium Doped Fiber Amplifier (EDFA) has a flat gain spectrum between 1530nm and 1560nm. Previously presented experiments **[7-9]**used two wavelengths more than 18nm apart, while this evaluation operates 8nm apart. The used wavelengths corresponds to ITU channel 21 (1558.983nm/192.3THz) and 41 (1550.918nm/193.3THz) in a DWDM system.

III. THEORY

The theory for single way dual wavelength optical fiber time and frequency transfer is based on the transit time τ for propagation of a single mode in a fiber **[17]** expressed as the group velocity for a certain distance L and the wavelength λ.

$$\tau = \frac{L}{c}\left(n - \lambda \frac{dn}{d\lambda} \right) \quad (2)$$

Where n is the refractive index and c is the speed of light in vacuum. The transfer time τ, sometimes known as the group delay time in a fiber, is thus dependent on the refractive index and the wavelength. This means that two different

wavelengths will propagate at different velocity in the same fiber. A standard single mode fiber is temperature dependent, to an extent shown in previous studies **[10]**, and the most important factor to include in the calculations. By calculating the derivative of the transfer time with respect to temperature, both wavelength and refractive index will be taken into account **[9]**. Polarization mode dispersion is omitted since it is small compared to chromatic dispersion over the wavelength span. The variation in transfer time as a function of temperature can thus be calculated **[9]** and that expression shows how the refractive indices of the two wavelengths are influenced by temperature, and based on this the variations in the time transfer delay can be calculated. The time transfer technique uses the property that the variations are different, but correlated, which also is supported by experimental results.

The difference in transfer time through the fiber is dependent on the variation of length, L, and the variation in refractive index, n. Both these effects will affect the chromatic dispersion of the fiber, but through different properties.

The refractive index of the fiber can be described by the Sellmeier equation **[18, 19]**

From the Sellmeier equation and empirical data for fused Silica (SiO2) **[19]** the material dispersion can be calculated with the notation that it may be slightly offset in communication fibers where the silica is doped with small amount of other substances. Nevertheless the overall behavior is comparable. It is possible to estimate the amount of time transfer delay variations with respect to temperature. Assuming a fiber where material dispersion is dominant (as is the case in standard single mode fiber), at a length of 20 km and measurement at 1530 nm and 1560 nm **[9].** The slope of such calculated dispersion is -0.0016 ps/(nm·km·°C), which is comparable to previously reported results -0.0025 ps/nm/km/°C for non-zero dispersion shifted fiber (NZDSF) and -0.0038 ps/nm/km/°C for large core fiber**[10]**.

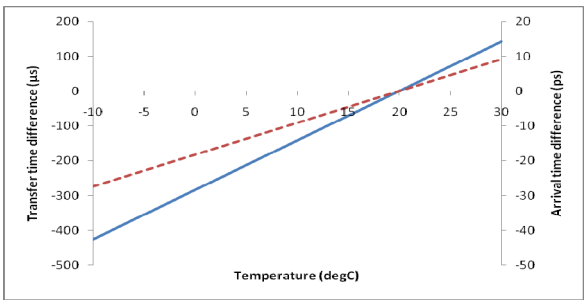


Figure 2. Temperature dependence of transfer time (solid blue, left axis) and arrival time difference (dashed red, right axis)

The solid curve (left axis) shows the calculated time transfer delay for a signal at 1530 nm, and the dashed curve shows the arrival time difference for two signals at 1530 nm and 1560 nm. Both curves are normalized with respect to the value at 20°C, and it is apparent that the time transfer delay within a single, 20 km long fiber varies with almost 0.5 ms when affected by 40°C temperature difference. This variation can be detected and compensated for, using transmission at two wavelengths and a measurement system that can measure time variations on picoseconds level with sufficient precision.

This evaluation assumes that the cabling or mounting will stretch the fiber at increasing temperature, however leaving the volume intact. The variations in dimensions of the glass are assumed to be negligible. If the core of the fiber is modeled as a glass cylinder, of length L and diameter d, a geometrical approach gives that the variation in temperature will change the length with $\Delta L(T\text{-}T_0)$ and the diameter with $\Delta d(T\text{-}T_0)$, such that

$$\frac{\Delta d(T-T_0)}{d} = -\frac{\Delta L(T-T_0)}{2L} \tag{3}$$

where T is the temperature and T_0 is the reference temperature.

This change in diameter will change the dispersion according to the variation in waveguide dispersion **[20, 21]**:

$$D_W(\lambda) = -\frac{n_2\Delta}{c\lambda} V \frac{d^2(Vb)}{dV^2} \tag{4}$$

where n_2 is the refractive index of the cladding and Δ is the relative difference of refractive index in the core and in the cladding. V and b are the normalized frequency and the normalized propagation constant, respectively, and can be found through:

$$V = ka\sqrt{n_1^2 - n_2^2} \cong kan\sqrt{2\Delta} \tag{5}$$

$$b = \frac{(\beta/k)^2 - n^2}{n_1^2 - n_2^2} \tag{6}$$

where k is the free-space propagation constant, β is the propagation constant and $a = d/2$ is the fiber core radius. From these equations, it is apparent that fibers with notable waveguide dispersion, e.g. dispersion shifted fibers, dispersion compensating fibers, etc, will have different response to a change in diameter d, than standard fibers where material dispersion is dominant. However, this response must be evaluated for each fiber design, since the term $V(d^2(Vb)/dV^2)$ is between 0 and 1.2 with a maximum at V≈1.2. These equations show nevertheless that the system of detecting a variation in transfer time through a fiber with substantial waveguide dispersion is possible, but must be optimized for the actual fiber parameters.

IV. EXPERIMENTAL AND REFERENCE SETUP

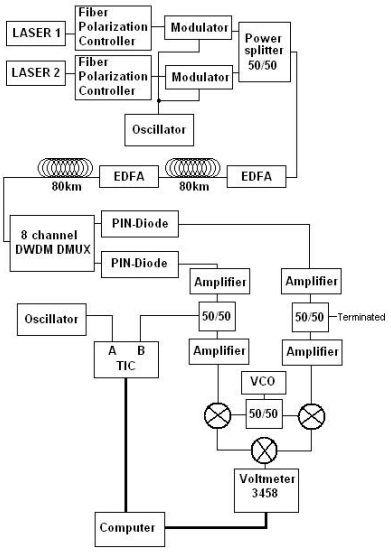


Figure 3. Experimental setup. The PIN-diode includes an electrical trans-impedance amplifier. Amplifiers are electrical amplifiers, 3458 a digital voltmeter and TIC is a time-interval counter

A combination of the experimental setup and the equipment for the verification of the proposed time and frequency transfer technique is shown in figure 3. Two CW lasers at wavelength 1551 nm and 1559 nm are after polarization controllers modulated in separate Mach-Zehnder modulators by a 10MHz reference oscillator, after which the modulated light waves are launched into the SMF through a 50/50 power combiner. The reference oscillator is frequency stabilized by an Auxiliary Output Generator (AOG) connected to a Hydrogen-maser as the Master clock. For this evaluation is the oscillator also used as reference to the measurement equipment, connected as an external oscillator to the Time Interval Counter (TIC), in order to evaluate the one-way technique. The total sum of the two spools of fiber that is in use was measured with an OTDR (optical time-domain reflectometer) to be 2x80km. Added to this length is 188m of transfer fiber between the lab and the outdoor fiber spools. The fiber path starts and ends in the laboratory for evaluation. The use of several fiber spools instead of one creates a similar case to a commercial link, which will be assembled of multiple fibers spliced by connectors, and there is no possibility to know the ageing of all optical fiber along this.

At the receiving end the two wavelengths are separated in an 8-channel Dense Wavelength-Division Multiplexing Demultiplexer (DWDM DMUX), and detected in two 10 Gb/s p-i-n receivers. The signals are amplified and connected to the Radio frequency (RF) ports of two double balanced mixers. One of the signals passes a power splitter that is connected to the reference TIC, which measures the total transfer time as reference. The output of the TIC is interpreted as the precision of an uncompensated single way time and frequency transmission.

By measuring the voltage changes from a third mixer connected to the other two mixers is a correction signal achieved and can be used for a real-time delay control of the uncompensated signal. Most of the equipment is housed in a laboratory with controlled environment, except the spools of SMF28 which is placed outdoors inside a box for emulation of an actual environmental condition.

The 8-channel DWDM DMUX has a built in optical filter function that is used for separation of the two wavelengths. Measurement equipment detects the two 10MHz sine waves after propagation through the fiber link, amplifies and compares with a reference signal, as shown schematically in Figure 4.

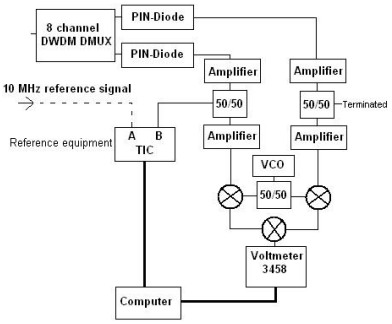


Figure 4. Schematics of the receiver and reference signal placed in the laboratory. The reference signal is compared to the information received from the transmitted wavelenghts

The output signals from the mixer connected to the voltmeter are calculated in Equation (8) and low pass filtered, which results in Equation (9).

$$\omega = \omega_1 = \omega_2 \tag{7}$$

$$M_v = \frac{A_1 A_2 A_L^2}{8}\left(\begin{array}{l}\cos\left(\left(\alpha_1 - \alpha_2\right)(t)\right) + \\ \cos\left(\left(2\left(\omega + \omega_L + \alpha_L\right) + \alpha_1 + \alpha_2\right)(t)\right)\end{array}\right). \tag{8}$$

Low pass filtering

$$M_v = \frac{A_1 A_2 A_L^2}{8}\left(\cos\left(\alpha_1(t) - \alpha_2(t)\right)\right). \tag{9}$$

This means that the measured result from the voltmeter must be corrected for according to Equation (9).

V. RESULTS

A major goal for this experiment was to resemble a fiber optical telecommunication network with a distance longer than 100km and these goals were fulfilled in the experimental setup. The experimental system consisted of a transmission side with lasers tunable within the optical C-band that was connected to fiber polarization controller and later on modulated with 10MHz. The transmission path consisted of a

preamplifier that is a EDFA followed by 80km of standard single mode fiber SMF28, then another EDFA followed by 80km SMF28. On the receiver side an 8-channel DWDM DMUX was used for separation of the two propagated wavelengths after which p-i-n-diodes were used for conversion to electrical domain. The transmitted wavelenghts are only 8nm apart to be able to fit within a single band in the proposed fiber optical university network of Sweden. After conversion to the electrical domain the improved mixer solution is used for extraction of delay difference in propagation, and time interval measurements are made as reference comparisons and verification. Transmitter and reciver are placed in laboratory enviroment as well as the EDFAs while the 160km of transmission fibers were placed outdoors in a container for enviromental impact. The temperature is also logged using an outdoor temperature sensor.

The results are extracted from a dataset of 10 days in the end of March and beginning of April 2011 and are displayed in figures 5 and 6. Time interval measurements in these figures are corrected for jumps as the output range is 0-100 ns, while the mixer solution is set to have a initial calibration at the start to be able to calculate the fiber path parameters that would not be possible to measure in a system spread over a large distance. This calibration is presumed necessary for every individual system. The mixer solution is in this experiment also set to have an initial startup period of 4 hours of which are spend to create a running average, aimed to reduce the noise. Furthermore, the mixer solution is scaled to be able to be evaluated versus the TIC measurements and a graph displaying these data is shown in Figure 5.

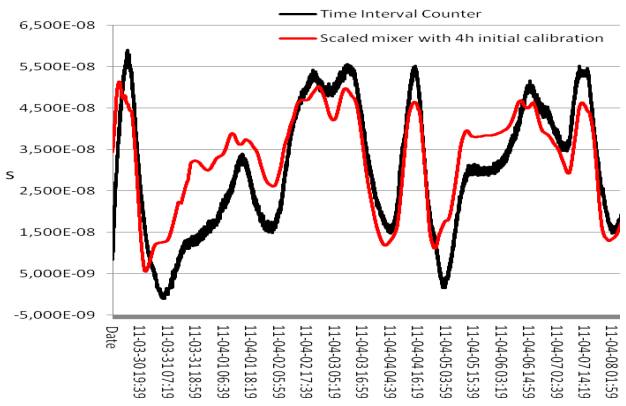


Figure 5. Data from the outputs of the mixer (red curve) and the time interval data represented by the reference TIC (black) is plotted from data during ten days of measurement

A similar behavior between the mixer solution and the reference TIC is read out from figure 5. The graphs in figure 6 include the curves in figure 5 which are plotted against the outside temperature. All three curves confirm a similar behavior. The mixer solution is scaled by a factor of 3 µs/V and has a offset of 15 ns. The standard deviation between the TIC graph and the Mixer graph is 8.45 ns.

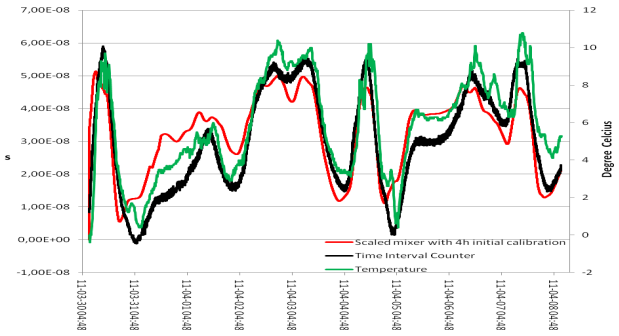


Figure 6. Data from the outputs of the mixer (red curve) and the time interval data represented by the reference TIC (black) is plotted togheter with the outside temperture

To be able to compare the to solutions properly, the original raw data without applicated smoothing should be plotted, and that is displayed in figure 7.

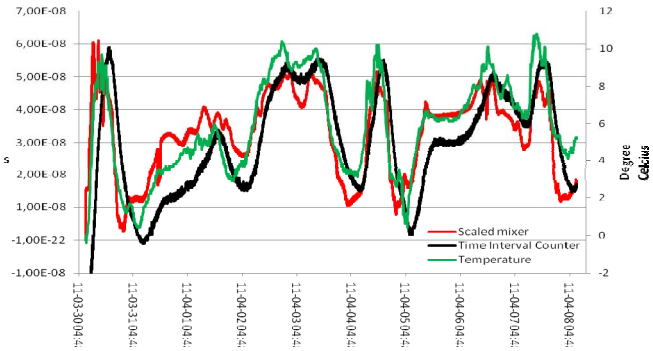


Figure 7. Rawdata from the outputs of the mixer (red curve) and the time interval data represented by the reference TIC (black) is plotted togheter with the outside temperture

By comparing TIC and Mixer data with the outside temperature there are evidence that the TIC add some delay and internal smoothing of the data. Further investigations of these equipment related issues have to be performed.

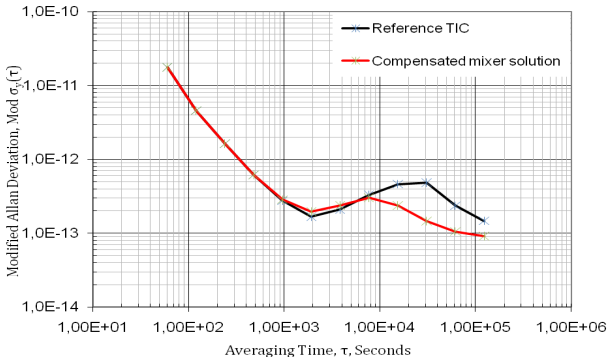


Figure 8. The stability of the reference TIC (black curve) and the scaled compensated mixer solution (red curve)

The stability of the TIC-data and the newly improved scaled compensated mixer solution are plotted as the Modified Allan Variance in figure 8. In the previous experimental setup **[9]** the whole system was affected by white phase noise in the short term and a bump between 500s to almost 4000s was assumed to be noise added from the mixers, amplifiers, long coaxial cables and the voltmeter, but also due to the simple correction algorithm.

A large effort has been made into reducing these noise sources and the result can be seen in figure 8 where the whole bump has disappeared. According to the figure the two systems are comparable up to the vicinity of 5500 seconds. At longer time scales it seems as the mixer systems with compensation is more stable.

VI. CONCLUSION

It is experimentally verified that it is possible to perform a one-way time and frequency transfer of optical single mode fiber with continuous estimate of delay variations. The technique is based on transmission over two wavelengths, separated by 8nm and both residing in the C-band of optical communication. By evaluating the phase difference between these time transfer delay of two signals, a correction can be applied to compensate for the unwanted delay variations along the transmission path.

This experiment has shown proof of concept for long baselines measurement, i.e. distances larger than 100km. Future work includes development of a physical real-time correction component at the end of the link that incorporates the steering signal based on the measured difference between the arriving signals at the two wavelengths. Also there is still some work to be done in reducing the noise in the receiver as well as creating a startup algorithm and a calibration algorithm for the mixer solution. Unlike previous experiments **[7-9]**, many components in the system are replaced e.g, the laser changed to telecom products for DWDM system, and the optical filters and the power splitter at the receiver end are replaced by an 8-channel DWDM DMUX. The transmission fibers are exchanged from 38km to 160km. These changes have resulted in less noise and a more stable system that is possible to incorporate in one band of an active data DWDM-system for evaluation. Nevertheless there are some evaluation of the mixer solution and the TIC that remains.

REFERENCE

[1] S.R. Jefferts, M. Weiss, J. Levine, S. Dilla, and T. E Parker, "*Two-Way Time Transfer through SDH and Sonet Systems*", European Frequency and Time Forum EFTF'96, 5-7 March, 1996.

[2] M. Kihara, A. Imaoka, M. Imae, K. Imamura, "*Two-Way Time Transfer through 2.4 Gb/s Optical SDH Systems*", IEEE Trans. Instr. Meas., vol. 50, pp. 709-715, 2001.

[3] F. Kéfélian, H. Jiang, P. Lemonde and G. Santarelli, "*Ultralow-frequency-noise stabilization of a laser by locking to an optical fiber-delay line*", Optics Letters ,Vol 34, No 7, April 1, 2009.

[4] S.C. Ebenhag, P.O. Hedekvist, C. Rieck, H. Skoogh, P.Jarlemark, and K. Jaldehag, "*A fiber based frequency distribution system with enhanced output phase stability*", 23rd European Time and Frequency Forum, EFTF'09, April 20-24, (2009), Page(s):1061– 1064.

[5] R. Emardson, P.O. Hedekvist, M. Nilsson, S.C. Ebenhag, K. Jaldehag, P. Jarlemark, C. Rieck, J. Johansson, L. Pendrill, P. Löthberg and H. Nilsson, "*Time Transfer by Passive Listening over a 10 Gb/s Optical Fiber*", IEEE Trans. Instr. Meas.,vol. 57, pp. 2495 – 2501, 2008.

[6] S.C. Ebenhag, P. Jarlemark, R. Emardson, P.O. Hedekvist, K. Jaldehag and P. Löthberg, "*Time transfer over a 560 km fiber link*", Proc. of the 22st European Frequency and Time Forum, paper 130 Toulouse France ,April 22-25, 2008.

[7] S.C. Ebenhag, P.O. Hedekvist, and K. Jaldehag,, *"Fiber based frequency distribution based on long haul communication lasers",* Paper 7, 41st Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Santa Ana Pueblo. NM, 2009.

[8] S.C. Ebenhag, P.O. Hedekvist, J. Johansson, "*Fiber based one-way time transfer with enhanced accuracy*", Proc. of the 24th European Frequency and Time Forum, Space Research and Technology Centre of the European Space Agency (ESA/ESTEC) Noordwijk, The Netherlands, April 13-16, 2010

[9] S.C. Ebenhag, P.O. Hedekvist, and K. Jaldehag, *"Single way fiber based time transfer with active detection of time transfer variations",* Paper 32, 42nd Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, Virginia, 2010.

[10] A. Walter, G. Schaefer, "Chromatic Dispersion Variations in Ultra-Long-haul Transmission Systems Arising from Seasonal Soil Temperature Variations", Optical Fiber Communication, OFC02, Paper WU4, March 17-22 Atlanta GA, 2002.

[11] P.J. Vella, P.M. Garrel-Jones, R.S. Lowe, "Measuring Chromatic Dispersion of Fibers", US Patent 4551019, Nov. 5, 1985.

[12] W. H. Hatton, M Nishimura, "Temperature dependence of chromatic dispersion in single mode fiber", Journal of lightwave technology, Vol LT-4, No10, October 1986

[13] L. G.Cohen, J. W Fleming, "*Effect of temperture on transmission in lightguides*", The Bell System Technical Journal, April 1979.

[14] L. A. Bergman, S. T. Eng and A. R. Johnston, "*Temperture stability of transit time delay for a single-mode fibre in a loose tube cable*", Electronics Letters ,Vol 19, No 21, 13th October 1983.

[15] W. H. Hatton, M Nishimura, "*Temperature dependence of chromatic dispersion in single mode fiber*", Journal of lightwave technology, Vol LT-4, No10, October 1986.

[16] K. Cochrane, J. E. Bailey, P. Lake and A Carlson, "*Wavelength-dependent measurements of optical-fiber transit time, material dispersion, and attenuation*" Applied Optics, Vol 40, No 1, January 2001.

[17] K. Cochrane, J. E. Bailey, P. Lake and A Carlson, "Wavelength-dependent measurements of optical-fiber transit time, material dispersion, and attenuation" Applied Optics, Vol 40, No 1, January 2001

[18] W. Sellmeier, "Zur Erklärung der abnormen Farbenfolge im Spectrum einiger Substanzen", Annalen der Physik und Chemie 219, 272-282, 1871.

[19] G. Ghosh, M. Endo and T. Iwasaki, "Temperature-Dependent Sellmeier Coefficients and Chromatic Dispersions for Some Optical Fiber Glasses", J. Lightwave Technol. Vol. 12, pp 1338 – 1342, 1994.

[20] D. Gloge, "Dispersion in weakly guided fibers", Appl. Opt., vol 10, pp 2442 – 2445, Nov. 1971.

[21] G. Keiser, "Optical Fiber Communication, 2nd ed", McGraw-Hill, 1991.

One-Way Temperature Compensated Fiber Link

J. L. Hanssen, S. G. Crane, and C. R. Ekstrom
Time Service Department
US Naval Observatory
Washington, DC USA
james.hanssen@usno.navy.mil

Abstract— **We will discuss our work on a one-way optical fiber frequency transfer system. The link removes thermally generated delays through sensing the difference in temperature dependent changes in the group index of the fiber for two wavelengths. The system monitors the phase difference between two amplitude-modulated lasers at the edges of the C-band. Using this signal, a temperature controlled spool of fiber compensates for thermally induced phase excursions in the link. We will present preliminary results as well as a discussion of performance limitations.**

I. INTRODUCTION

As atomic clock technology has progressed, great improvements in stability and accuracy have been achieved. Atomic fountain clocks have short-term stabilities in the upper 10^{-14}'s at one second and integrate down to the low 10^{-16}'s in a day [1]. Newer optical clocks are reaching short-term stabilities in the low 10^{-15}'s and averaging down into the 10^{-17}'s in an hour [2]. It is necessary to have a time and frequency transfer method capable of supporting this level of performance.

One of the standard ways of transmitting time and frequency is through two way satellite time and frequency transfer (TWSTFT). While it is practical for transmitting time and frequency over long distance, the performance is not acceptable for some of the newer high quality clocks. Currently, TWSTFT can transmit frequencies with instabilities at the 10^{-15} level and time at the 1ns level after a day of integration [3]. It is clear that to transfer stabilities achieved with modern atomic clocks, both microwave and optical, better transfer methods are needed.

One possibility is time and frequency transfer through fiber optic links. There are many promising aspects for using fiber optic links. First, there is an existing infrastructure that provides a large network for transferring time and frequency. The links are low loss over large distances and are well isolated. The access to a common return path allows for measurement and compensation of phase fluctuations. Finally, there is great potential for scalability in fiber optic networks.

Most methods for transferring time and frequency along fiber links involve establishing a two way link. The timing signal is sent down the link and then returned to the transmitting station. At the transmitting station, the phase variation brought about by fiber noise are measured and corrected for. One approach is to use an amplitude modulated signal. Short-term stabilities of 10^{-15} integrating down into the low 10^{-18}'s have been demonstrated [4]. Another approach involves using the phase of the optical carrier itself. Excellent performance has been achieved with short-term stabilities in the 10^{-16}'s averaging down to the 10^{-19}'s in a day [5]. An in depth overview of fiber transfer techniques can be found in [6].

Here we present results on a different type of fiber optic link. This is a one-way temperature compensated fiber link [7]. A one-way, amplitude modulated signal is transmitted along the fiber along with a proxy for measuring the temperature. The proxy signal is then used to compensate for temperature induced phase fluctuations along the link. This approach differs from the previous methods in that it is a one-way transfer. Bidirectional access to the fiber is not necessary as in the case of two-way links. Also, the ability to build simpler broadcast networks exists with improved scalability.

This one-way transfer method relies on measuring temperature indirectly. We use the phase difference between two separate colors propagating along the fiber to determine temperature. This is due to the temperature dependence of dispersion in fiber. All materials have an index of refraction for light and exhibit dispersion, meaning that the index is wavelength dependent. But in addition to being wavelength dependent, the index is also temperature dependent.

We can model the index of refraction for fiber by using the Sellmeier coefficients for fused silica [8]. In particular we can use temperature dependent coefficients to see the expected temperature dependent features. The temperature dependent index of refraction has the following form

$$n^2(\lambda, T) - 1 = \sum_{i=1}^{3} \frac{S_i(T) \cdot \lambda^2}{\lambda^2 - \lambda_i^2(T)} \qquad (1)$$

where

$$S_i(T) = \sum_{j=0}^{4} S_{ij} \cdot T^j \qquad (2)$$

and

$$\lambda_i(T) = \sum_{j=0}^{4} \lambda_{ij} \cdot T^j. \qquad (3)$$

The S_i coefficients are determined by the strength of resonance features that occur within the fiber at wavelengths λ_i. A thorough description of this model as well as the complete set of values for the different coefficients can be found [8].

Based on this model, we can calculate the time delay for a pulse or any amplitude modulation feature traveling down the fiber at a given wavelength and temperature. If we have two pulses at two different colors traveling down the fiber at a given temperature, they will arrive with a fixed time difference. As the temperature changes, that time difference changes based on the changing index of refraction. This is shown in Figure 1.

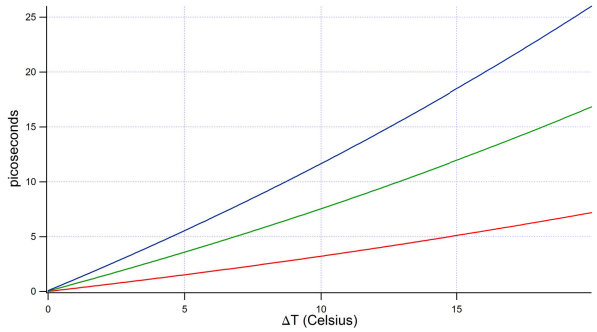


Figure 1. Temperature dependent phase shifts. Calculate results for the changing time delays between two pulses at different colors transmitted on 1km of single mode fiber as the temperature of the fiber changes. Blue corresponds to light at 1530nm and 1565nm. Grey is 1530nm and 1553nm. Red is 1530nm and 1540nm. Increasing wavelength separation gives increasing sensitivity to the temperature dependent phase shift.

From this we see that it is possible to derive a change in temperature in the fiber by measuring the change in time delay between two different color signals. We can use this indirect temperature measurement as a means of compensating for temperature induced phase changes in the fiber. Note that this method only compensates for temperature induced effects. It will not compensate for any mechanically induced phase fluctuations.

II. Experiment

As has been shown, it is possible to use the time delay between pulses at two different colors as a proxy for temperature measurements of a fiber link. Instead of measuring time delays between pulses, the phase change between two amplitude modulated signals were measured. The experimental setup to measure is described in the following and shown in Figure 2.

A master oscillator at 5MHz drives the system on the transmit side. For testing purposes this is just a high quality crystal oscillator. That 5MHz signal is multiplied up to 100MHz. The 100MHz signal feeds a phase locked dielectric resonant oscillator to produce a signal at 6.8GHz. This signal is split and used to modulate two transmit laser modules.

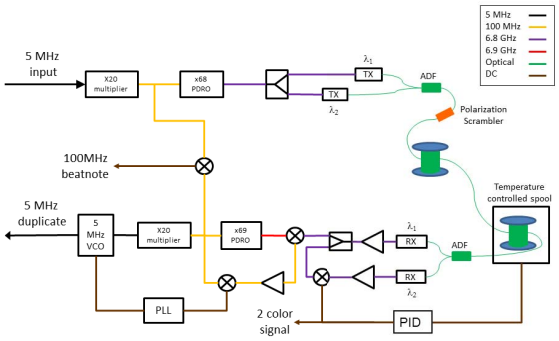


Figure 2. Experimental setup.

The transmit modules are laser diodes with electro-absorptive modulators operating at λ_1 = 1530nm and λ_2 = 1553nm. The two signals are combined onto one fiber using and add/drop filter. The signal then goes through a polarization scrambler. It was experimentally verified that the two color signal has a strong dependence on any birefringence due to residual stress in the fiber. By scrambling the polarization, the signal is averaged over all polarization states and that sensitivity is reduced.

At this point the laser signal travels down the link. For testing purposes, this link is a spool of fiber residing in an environmentally controlled chamber. It was possible to adjust the length of the link from 1m to 30km.

After passing along the link, the signal enters the receive station. The first stage of this station is a temperature controlled spool. The spool consists of 5km of bare single mode fiber wound on an aluminum mandrel. The spool is placed in an insulated box and is temperature controlled using a TEC and commercial temperature controller.

After the temperature controlled spool, the two wavelengths are separated using another add/drop filter and fed to two separate fiber coupled photo receivers. The outputs of each receiver are amplified with a high gain microwave amplifier. The 6.8GHz modulation signal at wavelength λ_1 is split with one half mixed with the entire signal at wavelength λ_2. The change in phase between the two colors' AM signals is the previously described two color signal. The two color signal is amplified and fed to a PI controller. The PI controller determines the set point for the temperature controlled spool to cancel temperature induced fluctuations in the fiber.

The second half of the signal at wavelength λ_1 is used to phase lock a high quality 5MHz quartz crystal to the transmitted signal. The 5MHz signal from the quartz crystal is

multiplied up to 100MHZ and then fed to a phase locked DRO to produce 6.9GHz. The 6.8GHz signal is mixed with the 6.9GHz signal to create a 100MHz signal which is then mixed with the intermediate 100MHz signal in the send station. This error signal is then used to phase lock the crystal to the transmitted signal, completing the clock recovery at the receive station.

In order to test the stability of the transmitted signal, a portion of the 100MHz signal from the transmit station and the send station are mixed together. The resulting beat note is recorded with an analog to digital convert and later analyzed for stability.

III. Results and Conclusions

A first test was to measure the size of the two color signal as a function of temperature change. A 5km spool of fiber was used for the link. The environmental chamber cycled the temperature with 4°C linear ramps every two hours. The PI controller for the compensation spool was turned off, and the temperature set at a fixed value. The two color signal was recorded as was the 100MHz beat note between the transmit and receive stations.

Figure 3 show the results of the measurement. As can be seen, as the temperature changes, there is a change in the relative phase between the transmit and receive stations. For the given link length and temperature difference this is approximately 300ps. Likewise, a phase change in the two color signal is evident, although with a lower magnitude. For the same length and temperature change, the phase change is about 1ps. There is excellent correlation between the 100MHz beat note phase change and the two color signal phase change. The lever arm, or ratio of actual phase change to two color signal phase change, is about a factor of 300. This is an important value in that any noise on the two color signal will get multiplied by the lever arm and written on the actual signal.

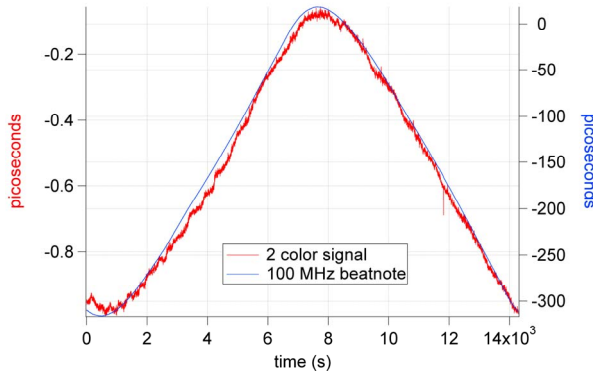


Figure 3. Lever arm measurement. The phase change for the two color signal (red) and the 100MHz beatnote (blue) as a function of changing temperature. The two color signal changes by approximatel 1ps while the 100MHz beatnote changes by over 300ps.

As a note, these results were achieved after ensuring similar fiber throughout the link. The scheme relies on a direct correspondence between temperature and expansion since the actual phase variations are not measured [9]. This means that the coefficient of thermal expansion must be constant for the entire link otherwise different temperature changes give rise to different phase variations. Therefore, bare single mode fiber was used throughout the experiment.

The next test was to use the temperature controlled spool to cancel out the temperature induced phase fluctuations and lock the link. The same 5km spool was used for the link and the same cycling conditions were used in the environmental chamber. The PI controller was turned on and used to adjust the temperature of the compensation spool. The net effect was to drive the two color signal to zero.

The results of locking the link are shown if Figure 4. The link was locked for approximately one day. The two color signal held constant near zero while there are small fluctuations on the 100MHz beat note. For comparison, the link was unlocked, denoted by the black vertical line in Figure 4, and allowed to run for about one day. It is clear that the two color signal as well as the 100MHz signal have much larger phase variations when unlocked.

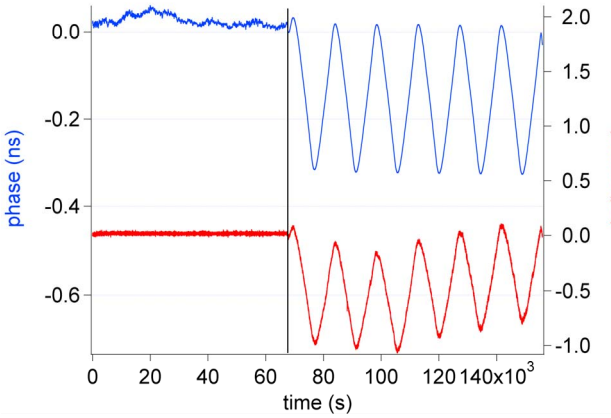


Figure 4. Locked and unlocked link performance. The plot in blue is the 100MHz beatnote signal and the plot in red is the two color signal. The first half of the data set corresponds to the link being locked and compensated. At the black line, the compensation is turned off.

Figure 5 shows the stability of the locked and unlocked links when the environmental chamber is cycling. The Allan deviation is plotted for both situations. First, consider the unlocked case. The stability starts of at 6×10^{-14} at 1 second. It integrates down as $1/\tau$ until about 100 seconds. At that point, the effects of temperature cycling become apparent causing the stability to increase as τ with a peak at half the cycling time of 2 hours.

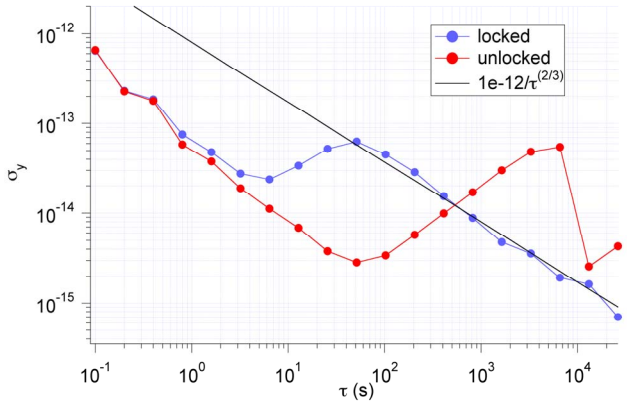


Figure 5. Link Stability. Plotted are the Allan deviations for the locked and unlocked data presented in Figure 4. A reference line of $10^{-12}/\tau^{2/3}$ is plotted as well.

In the case of the locked link, the effects of temperature cycling are removed. Initially, the stability is 6e-14/τ like the unlocked case. At around 10s the stability turns up and eventually latches on to a trend of $1x10^{-12}/\tau^{2/3}$. This comes about because of the temperature control loop and the noise floor of the two color signal. The temperature controlled spool has a time constant on the order of 10 seconds hence the deviation from the unlocked loop around 10 seconds. The stability eventually rolls over to a line defined by the noise floor of the two color signal and the lever arm of the system. For a 10km length of fiber, the noise floor of the two color signal is ~ $3x10^{-15}/\tau^{2/3}$. The measured lever arm is approximately 300; therefore the noise floor is multiplied up by 300 and added to the stability of the loop. That corresponds to ~$9x10^{-13}/\tau^{2/3}$.

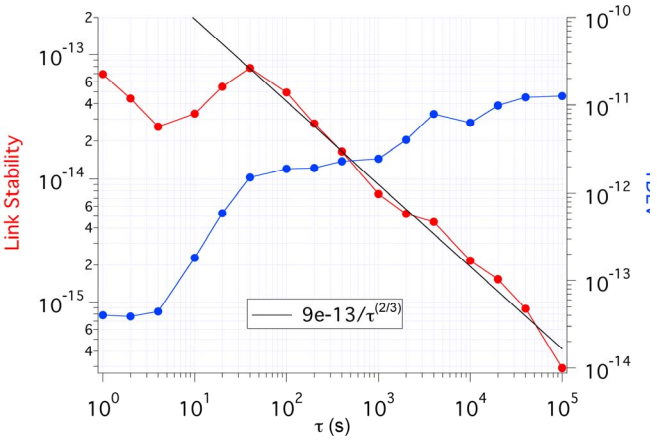


Figure 6. Extended link stability plot. Plotted are the Allan deviation and TDEV for a 5 day continual run of the locked link. . A reference line of $9x10^{-13}/\tau^{2/3}$ is plotted as well.

The results of a longer run with the link locked for 5 days are shown in Figure 6. Throughout, the environmental chamber was changing the temperature as in Figures 3 and 4. From the Allan deviation, we see that the locked link integrates down to the mid 10^{-16}'s in one day with the phase fluctuations suppressed. Also shown is the TDEV for that run. After a day if integration, the TDEV is less than 20ps.

As was pointed out, one of the limiting factors for the performance of the locked link is the noise floor of the two color signal. The limiting stability follows the noise floor times the lever arm. Therefore it is important to know how the noise floor scales with link length. Figure 7 show measurements of the 1 second fractional frequency stability as a function of fiber length. As can be seen, the stability degrades with increasing length and shows that the inherent short term stability increases at $3x10^{-16}$ per kilometer. To ensure that the effect was due to length and not attenuation, the 5km link was remeasured with attenuation equivalent to a 30km link. The results matched the stability of the original 5km link. From this we can infer the performance of longer links. Assuming the same set of transmitting wavelengths, and hence same lever arm, for a 50km link, we should expect the long term stability to behave as $5x10^{-12}$. There is evidence that the polarization scramblers used in the link are not properly optimized. It is possible that improvement of their operation may improve the scale of noise with fiber length.

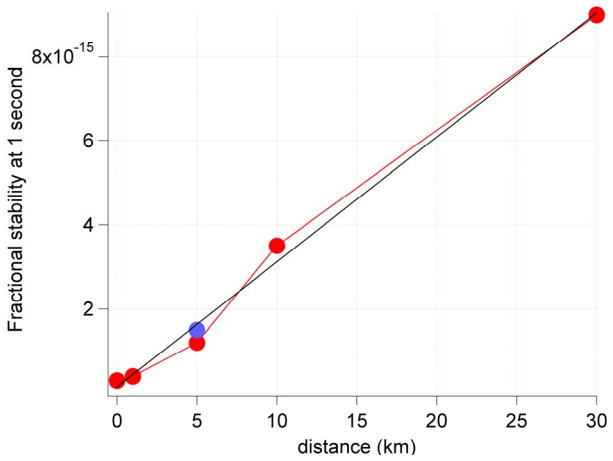


Figure 7. Stability versus length. Plotted is the 1 second stability measurement for the locked link as a function of link length. A fit line corresponding to $1.5x10^{-16} + 3x10^{-16}$/km * distance is plotted as well. The blue data point corresponds to a 5km link with attenuation equivalent to a 30km link.

One means of improving the stability of the link is to decrease the lever arm. This can be done by changing the wavelengths used by the transmit laser modules. A larger wavelength separation gives rise to a larger two color signal and smaller lever arm. Figure 8 shows the possible improvement. By using wavelengths at the edge of the C band (shown in green in the figure), there should be an improvement of roughly a factor of two. By extending even further and using a wavelength in the C band and one at

1300nm (shown in grey in the figure), there is almost an order of magnitude improvement in stability at the cost of no longer being able to use erbium doped amplifiers in the link.

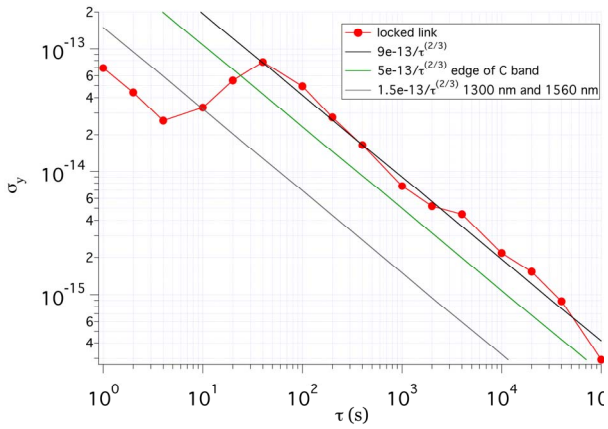


Figure 8. Wavelength improvements. The red curve is the Allan deviation shown in Fig 6. Plotted as well are simulations of changes to the trend line for different wavelength combinations. Black is a fit to the data corresponding to $9\times10^{-13}/\tau^{2/3}$. Green is a simulation for a wavelength combination of 1530nm and 1565nm corresponding to a trend line of $5\times10^{-13}/\tau^{2/3}$. Grey is a simulation for a wavelength combination of 1560nm and 1300nm corresponding to a trend line of $1.5\times10^{-13}/\tau^{2/3}$.

In conclusion, we have demonstrated a one-way temperature compensated fiber optic link. We have shown that by sending two amplitude modulated signals along a fiber optic link, it is possible to recover a record of the thermally induced phase variations and to compensate for them in real time. Some of the length based limitations to performance were discussed as well as possible improvements by appropriate choice of transmit wavelengths. This technology shows promise a viable alternative to two way compensated links when bidirectional access to the fiber is not possible. It also points to easier time transfer through a broadcast protocol as opposed to individual compensated links.

[1] R. Wynands and S. Weyers, "Atomic Fountain Clocks," Metrologia, vol. 42, pp. S64-S79, 2005.

[2] A. D. Ludlow, et al. , "Sr Lattice Clock at 1×10^{-16} Fractional Uncertainty by Remote Optical Evaluation with a Ca Clock," Science, vol. 319, pp. 1805-1808, 2008.

[3] A. Bauch, et al. ,"Comparison Between Frequency Standards oin Europe and the USA at the 10^{-15} Uncertainty Level," Metrologia, vol. 43, pp.109-120, 2006.

[4] O. Lopez, A. Amy-Klein, M. Lours, C. Chardonnet, and G. Santarelli, "High-Resolution Microwave Frequency Dissemination on an 86km Urban Optical Link," App. Phys. B, vol 98, pp. 723-727, 2009.

[5] H. Jiang, et al. "Long-Distance Frequency Transfer Over an Urban Fiber Link Using Optical Phase Stabilization," JOSA B, vol. 25, pp. 2029-2035, 2008.

[6] S. M. Foreman, K. W. Holman, D. D. Hudson, D. J. Jones, and J. Ye, "Remote Transfer of Ultrastable Frequency References via Fiber Networks," Rev. Sci. Instrum., vol. 78, p. 021101, 2007.

[7] S.-C. Ebenhag, P. O. Hedekvist, and K. Jaldehag, "Fiber-Based Frequency Distribution Based on Long-Haul Communicationn Lasers," 41st Annual Presice Time and Time Interval Meeting, 2009.

[8] B. B. Leviton and B. J. Frey," Temperature-Dependent Absolute Refractive Index Measurements of Synthetic Fused Silica," Proc. SPIE 6273, 62732K, 2006.

[9] N. Lagakos, J. A. Bucaro, and J. Jarzynski, "Temperature-Induced Optical Phase Shifts in Fibers," Applied Optics, vol. 20, pp. 2305-2308, 1981.

A compact laser-pumped Rb clock with $5x10^{-13}$ $\tau^{-1/2}$ frequency stability

C. Affolderbach, F. Gruet, R. Matthey, G. Mileti
Laboratoire Temps-Fréquence (LTF)
University of Neuchâtel
Neuchâtel, Switzerland
E-mail: christoph.affolderbach@unine.ch

***Abstract*— We have realized a compact, laser-pumped Rb atomic frequency standard in view of next generation space applications. The clock's Physics Package has a volume of 1.1 liters, and includes a frequency-stabilized diode laser and 2 clock resonator packages. The measured clock short-term stability is $5x10^{-13}$ $\tau^{-1/2}$, reaching < $3x10^{-14}$ at τ > 20'000s. The typical clock drift is $2x10^{-13}$/day, on the same level as the drift observed for a modified GALILEO RAFS operated with laser pumping under vacuum.**

I. INTRODUCTION

Rb atomic frequency standards ("Rb clocks") have found a variety of applications as local frequency references, in fields such as telecommunications, science, and satellite navigation systems, thanks to their compact volume and low mass while providing competitive frequency stabilities. Here we report on our development of a laser-pumped Rb clock [1] with improved clock performance, that meets the performance requirements for on-board clocks in advanced satellite navigation systems ($\sigma_y(\tau) < 1x10^{-12}$ at τ=1s and $< 1x10^{-14}$ at 1 day [2]) while maintaining the compact outline of a high-performance Rb clock (≤3 kg mass, ≤3 liters volume) [3]. To achieve this goal, the approach of a cw laser-pumped Rb clock with optimized key technologies (including the buffer-gas cells, laser stabilization, etc.) is chosen [4].

II. CLOCK REALISATION

In order to keep the clock scheme as simple as possible, a continuous-wave interrogation scheme combined with optimization of the key clock technologies was selected. The block scheme of the clock is given in Figure 1. The light emitted by a DFB laser diode is frequency-stabilized to saturated-absorption lines obtained from a small evacuated Rb reference cell [5]. In order to achieve best possible frequency stability from a simple setup, the cell is temperature-stabilized

This work was supported by the European Space Agency ESA (contract no. 19392/05/NL/CP), the Swiss State Secretariat for Education and Research – Space Affairs Division, the Swiss National Science Foundation, and Fondation en faveur d'un Laboratoire en Recherche Horlogère.

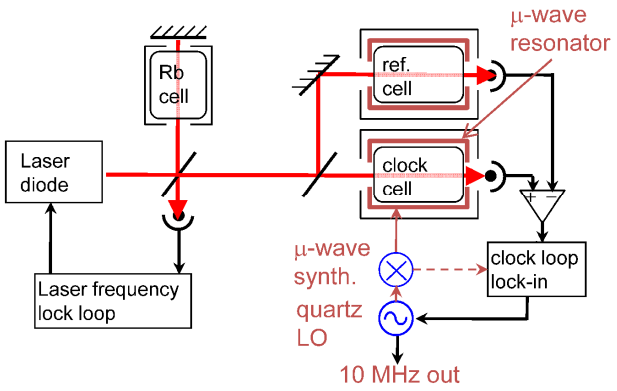


Figure 1. Block scheme of the clock.

and surrounded by two layers of magnetic shields. Laser stabilization is achieved by frequency modulation spectroscopy at a frequency around 50 kHz, resulting in a laser frequency stability of $< 5x10^{-12}$ at all relevant time scales. A more detailed description of the stabilized laser module can be found in [6, 7, 8].

The clock cell contains enriched ^{87}Rb and a buffer-gas mixture and is placed in a small magnetron-type microwave resonator [4]. The resonator and clock cell are temperature-stabilized and surrounded by a 3-layer magnetic shield to form an atomic resonator package, of 4.0cm x 4.4cm x 6.4cm outer dimensions. By injecting low phase-noise microwave radiation into the resonator, a clock signal can be detected on the photodetector placed after the atomic resonator package. The detection noise is dominated by laser-induced noise (including laser AM noise and FM-AM noise conversion), thus a second cell identical to the clock cell is used in a second, identical atomic resonator package into which no microwave radiation is injected, in order to subtract the detection noise from the actual clock signal for improved short-term stability.

The stabilized laser module and the two atomic resonator packages are mounted on a common, temperature-stabilized baseplate and occupy a volume of 1.1 liters. Figure 2 shows

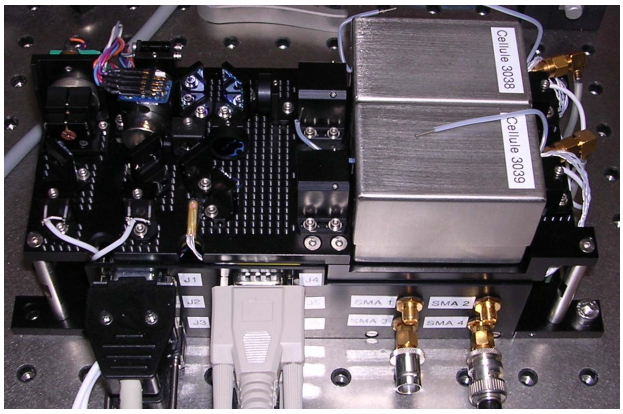


Figure 2. Fully assembled clock package. Overall dimensions are 10cm x 10cm x 24cm, with only 1.1 liter occupied by the core physics package (laser + atomic resonator modules).

the assembled clock, including also proximity electronics in an overall 10cm x 10cm x 24cm volume (2.4 liters) [7].

Under typical operating conditions, the double-resonance clock signal has a linewidth of ≈ 1 kHz and a contrast of 7%, which results in a discriminator slope of 680 pA/Hz (measured at line center in the de-modulated error signal). When using one clock cell only, the shot-noise limit is $3\text{x}10^{-13}\ \tau^{-1/2}$, and the signal-to-noise stability limit is $1.1\text{x}10^{-12}\ \tau^{-1/2}$, mainly due to additional laser noise (including FM-AM conversion in the clock cell).

Because the noise cancellation scheme requires identical clock cells, and in order to investigate on future issues on the production of many clocks, we have investigated on the reproducibility of the clock cell's buffer-gas mixture, which directly affects the cell temperature required to operate at minimized temperature coefficient (inversion temperature, [9]). For this purpose we have produced 4 completely independent batches of cells with nominally identical buffer-gas mixture, distributed over a period of 3 months. Evaluation of the cells' performance in a Rb clock shows that the inversion temperature does not scatter more than ±2°C, which is acceptable for clock operation and production constraints.

III. Clock Performance Results

Figure 3 shows the clock's measured short-term stability.

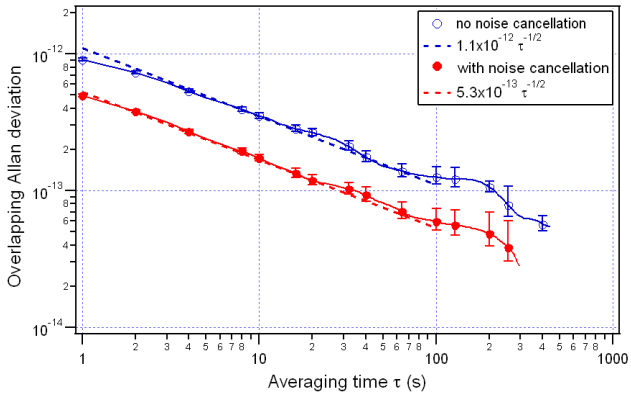


Figure 3. Short-term stability of the clock, measured with and without noise cancellation.

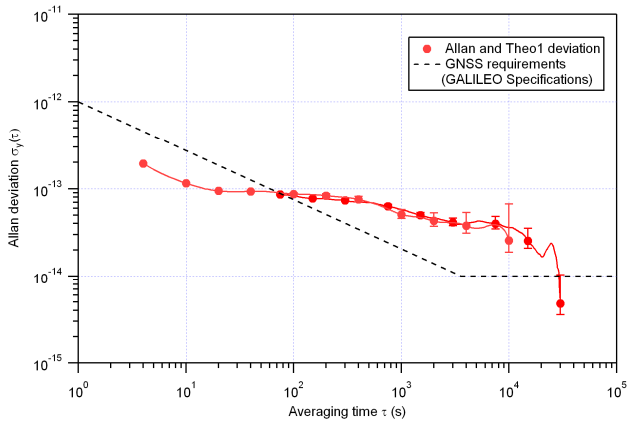


Figure 4. Measured medium-term stability of the clock. Typical frequency drift is around $2\text{x}10^{-13}$/day.

Short-term stabilities of $1.1\text{x}10^{-12}\ \tau^{-1/2}$ and $5.3\text{x}10^{-13}\ \tau^{-1/2}$ are obtained when operating the clock without and with the noise cancellation scheme implemented, respectively, in agreement with the calculated signal-to-noise limits. The latter result is also close to the shot-noise limit of $4.5\text{x}10^{-13}\ \tau^{-1/2}$, slightly increased here due to the implementation of two independent photodetectors when operating with the noise cancellation scheme.

The medium-term stability of the clock is shown in Fig. 4. It reaches $< 3\text{x}10^{-14}$ at $\tau > 2\text{x}10^{4}$s, with a typical drift of $2\text{x}10^{-13}$/day. This stability is mainly limited by temperature variations in our laboratory, acting on the laboratory-style clock electronics and directly on the clock physics (not sufficiently optimized thermal isolation using 2 thermostat stages only operating in standard atmosphere), to be mended in next development stages.

IV. Long-Term Clock Studies

In order to validate and evaluate the operation of a laser-pumped Rb clock under vacuum, we operate a modular Rb clock over extended times in a vacuum chamber held at $< 2\text{x}10^{-6}$ mbar. This laser-pumped Rb clock is composed of a stabilized laser head [5] and a lamp-removed GALILEO RAFS clock [3], as described in [4]. The stability of this clock is $1.6\text{x}10^{-12}\ \tau^{-1/2}$ at $\tau <$ 100s and reaches $2\text{x}10^{-14}$ at 20'000s. Figure 5 shows the clock frequency recorded continuously over 4 months, where the 12-hour mean uncertainty is $2\text{x}10^{-13}$. A stable linear drift of $1.8\text{x}10^{-13}$/day is found, and no frequency jumps linked to unexplained events or changes in the light source intensity (as found in lamp-pumped clocks [10]) are observed over the full measurement time within the 12-hour uncertainty. The only frequency jumps, observed (end of September 2010) are due to maintenance work in the laboratory.

By recording the laser diode's injection current, we can extract the aging of the current on resonance, i.e. of the laser current required for operating at the Rb D2 frequency (at constant laser temperature). A decrease of the current at resonance of –50 µA/month is observed, which corresponds to a + 60 MHz/month change in laser frequency (at constant current and temperature). A similar rate of frequency aging (-80 µA/month) was obtained for laser diodes operating in air

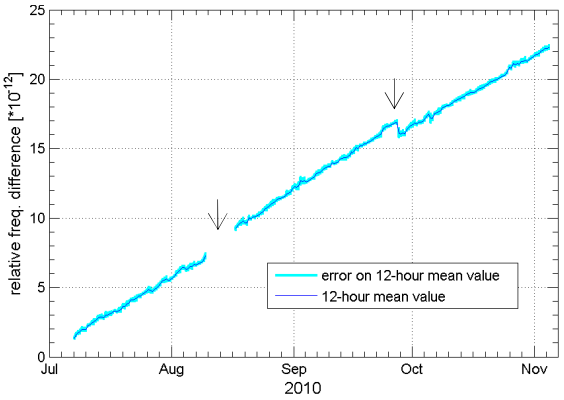


Figure 5. Frequency of the laser-pumped RAFS operated under vacuum, recorded over 4 months continuous operation. The events marked by vertical arrows are due to maintenance and temperature changes made on the laboratory air conditioning system.

in a dedicated test setup over more than 5 months. By linear extrapolation of the data we predict that the laser diodes can be operated at the Rb resonance wavelength for 15 to 20 years, by only slightly adjusting the laser temperature. A detailed study of this laser frequency aging is reported in [11].

V. CONCLUSIONS

We have developed and evaluated a very compact (< 2.5 liters volume) and high-performance laser-pumped Rb clock, that is based on the simplest possible schemes and uses optimized clock technologies. A short-term stability of $5.3 \times 10^{-13}\ \tau^{-1/2}$ is achieved without need for switching of the optical or microwave fields (like used in the pulsed optical pumping approach [8]), which significantly simplifies the clock scheme. Dedicated studies on the long-term behavior of a laser-pumped Rb clock operated under vacuum and on the frequency aging of the implemented laser diodes show that our clock can operate over extended times under vacuum, and that the laser frequency lifetime is compatible with a clock lifetime of at least 15 years. The developed clock is thus an attractive candidate for a number of space applications, including telecommunication, satellite navigation, and science missions.

ACKNOWLEDGMENT

We thank Spectratime SA for support on the laser-RAFS long-term studies, F. Ascarrunz (Spectradynamics Inc.) and C. Calosso (INRIM) for help with the microwave synthesizers, and P. Scherler, M. Durrenberger, D. Miletic, T. Bandi, J. Di Francesco, D. Varidel (all at UniNe-LTF) for their contributions to the work.

REFERENCES

[1] J. Vanier and C. Mandache, "The passive optically pumped Rb frequency standard: the laser approach," Appl. Phys. B vol. 87, pp. 565 – 593, May 2007.

[2] M. Belloni et al., "Space passive hydrogen maser -performances, lifetime data and GIOVE-B related telemetries", Proc. of the 24th European Frequency and Time Forum, Noordwijk, the Netherlands, 13-16 April 2010.

[3] F. Droz et al., "On-Board Galileo RAFS, current status and performances", Proceedings of the 2003 IEEE International Frequency Control Symposium Jointly with the 17th European Frequency and Time Forum, pp. 105-108, 2003.

[4] C. Affolderbach, F. Droz, and G. Mileti, "Experimental demonstration of a compact and high-performance laser-pumped Rubidium gas cell atomic frequency standard," IEEE Trans. Instrum. Meas., vol. 55, pp. 429 – 435, April 2006.

[5] C. Affolderbach and G. Mileti, "A compact laser head with high frequency stability for Rb atomic clocks and optical instrumentation," Rev. Sci. Instrum., vol. 76, 073108, July 2006.

[6] T. Bandi, C. Affolderbach, C. Calosso, and G. Mileti, "Investigations on improved Rb cell standards", Proc. Joint IEEE International Frequency Control Symposium and European Frequency and Time Forum, San Francisco (CA), USA, 2011.

[7] C. Affolderbach, R. Matthey, F. Gruet, T. Bandi, and G. Mileti, "Realisation of a compact laser-pumped Rubidium frequency standard with < 1×10^{-12} stability at 1 second", Proc. of the 24th European Frequency and Time Forum, Noordwijk, the Netherlands, 13-16 April 2010.

[8] S. Micalizio et al., "Pulsed Optically Pumped Rb clock with optical detection: first results", Proc. of the 24th European Frequency and Time Forum, Noordwijk, the Netherlands, 13-16 April 2010, p.116.

[9] J. Vanier, R. Kunski, N. Cyr, J. Y. Savard, and M. Tetu, "On hyperfine frequency shifts caused by buffer gases: Application to the optically pumped passive Rubidium frequency standard", J. Appl. Phys., vol. 53, pp. 5387 – 5391, August 1982.

[10] J. Camparo, "Does the light shift drive frequency aging in the rubidium atomic clock ?", IEEE Trans. Ultrason., Ferroelectr. Freq. Control, vol. 52, pp. 1075 – 1078, July 2005.

[11] R. Matthey, C. Affolderbach, and G. Mileti, "Methods and evaluation of frequency aging in DFB laser diodes for Rubidium atomic clocks", submitted for publication (2011).

Pulsed Optically Pumped Rb Clock with High Frequency Stability Performances

S. Micalizio, C. E. Calosso, A. Godone and F. Levi
Optics Division
Istituto Nazionale di Ricerca Metrologica, INRIM
Torino, Italy
e-mail: c.calosso@inrim.it

C. Affolderbach and F. Gruet
Laboratoire Temps-Fréquence (LTF),
Université de Neuchâtel
Neuchâtel, Switzerland

***Abstract*—In this paper we present a prototype of Rb clock working in pulsed regime with high freq stability performances. In particular, a frequency stability as low as $1.6\times10^{-13}\tau^{-1/2}$ has been measured for integration times up to 10^5 s. In the short term, this represents a record result for a hot-vapor-cell frequency standard.**

I. INTRODUCTION

The pulsed optical pumping (POP) scheme relies on the time separation of the following three phases that generally rule the operation of an atomic clock: laser optical pumping, microwave interrogation with Ramsey scheme, and detection of the clock transition [1]. It is well established that this technique has proven very effective to reduce the transfer of laser noise to the atomic resonance (the so called light-shift), increasing the frequency stability performances with respect to a continuous approach. In the case of the POP maser, in which the clock signal is detected in the microwave domain, we have in fact already reported a frequency stability (Allan deviation) of $1.2\times10^{-12}\,\tau^{-1/2}$ for integration times up to 10^5 s, and the 10^{-15} region has been reached after drift removing [2].

According to the theory we developed [3], the detection of the clock transition in the optical domain can lead to even better results. In fact, because of the higher energy carried by optical photons with respect to microwave photons, a higher signal-to-noise ratio is predicted and so a better short term stability. In addition, there is a secondary advantage: we do not need a high-*Q* cavity because the cavity is used to interrogate the atoms and does not play any role during detection; a lower cavity pulling is expected. The heterodyne detector is not more required and instead we have the photodiode and the transimpedance amplifier. A new physics structure has been then designed and implemented. Moreover, to take full advantage of the POP clock with optical detection, a very low-phase noise synthesis chain has been realized. The pulsed operation is ruled by a pattern generator implemented on a digital board.

Clock measurements have been done using optical pumping on either the Rb D1 or D2 optical lines.

II. PHYSICS PACKAGE

Fig. 1 shows a schematic of the physics package. It has a layer structure, from the innermost component to the outmost the following components can be recognized:

1) Quartz cell;
2) Microwave cavity system (including optical detector: trans-impedance plus photodiode and microwave cable to inject and detect microwave signals, not shown in the figure);
4) C field solenoid;
5) Internal heater;
6) First magnetic shield;
7) External heater;
8) Two external magnetic shields.

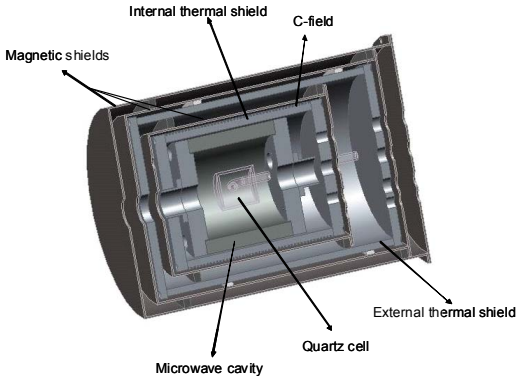


Fig. 1. Schematic of the physics package.

Two different cells have been filled with ^{87}Rb atoms and a mixtures of buffer gases (Ar and N_2 in the pressure ratio 1.6, with a total pressure of 25 Torr) at the LTF facilities. The buffer gas mixture is temperature compensated around 64 °C.

One of the two cells has been placed inside a Mo cavity that has a loaded quality factor lower than 1000. In Fig. 2 it is possible to identify the two cell stems, the microwave cable, the tuning screw and the photodiode.

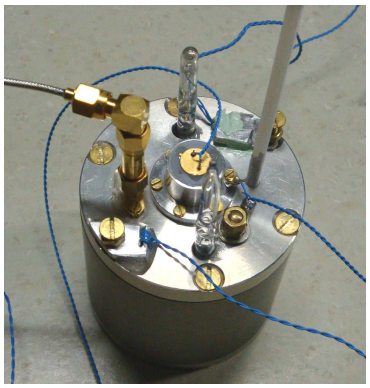


Fig. 2. The cavity system.

The physics structure has been placed in a vacuum chamber in order to isolate it from environmental fluctuations (barometric pressure, humidity, temperature, etc.); the vacuum is provided by rotative and turbo pumps and the vacuum level is below 10^{-6} mbar.

Since the clock frequency is temperature-dependent due to buffer gas collisions, an active control of the physics structure temperature is required. In particular, we implemented two different temperature controls: the innermost controls the cavity temperature and works in AC; the second one works in DC and controls the stem of the cell. Between the stem and the cavity a temperature gradient is maintained to generate a cold point in the structure (the stem) in order to avoid a deposition of Rb inside the walls of the cell that would alter significantly the cavity Q.

As shown in Fig. 3 a temperature stability for the stem and for the rest of the cell inside the cavity below 100 µK up to 100000s is achieved.

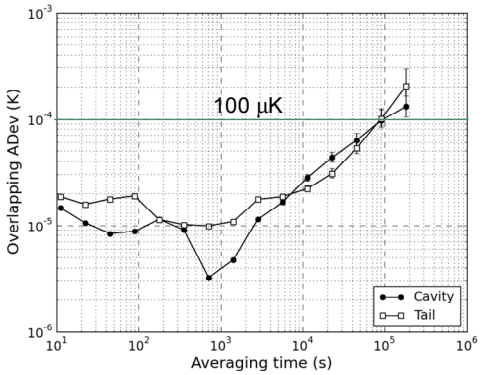


Fig. 3. Allan deviation of the temperature fluctuations of the cell stem (white squared) and of the cell body (black circles).

III. ELECTRONICS

The electronics architecture (see Fig. 4) is very similar to that we developed for the POP maser [3].Here we focus only into the main differences. In the optical detection the laser is not only used to pump the atoms but also to detect the clock resonance; the atomic signal is then processed by a transimpedance amplifier followed by a signal conditioning step to match the input dynamic of the lock-in amplifier. The rest of the electronics has been redesigned in order to support the expected stability.

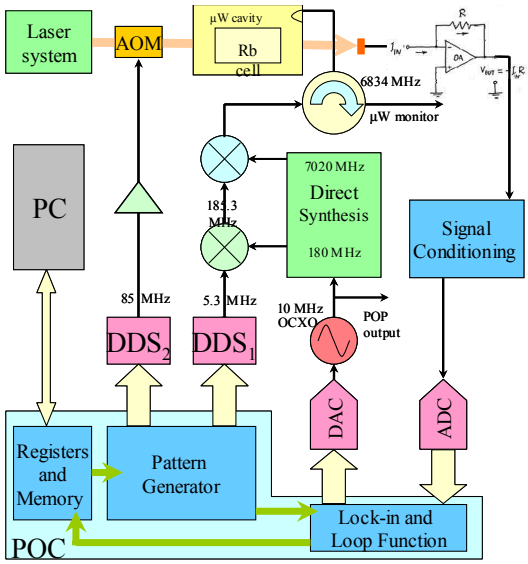


Fig. 4. Scheme of the POP clock from an electronic point of view.

In particular, the synthesis chain (SC) is based on non-linear transmission lines (NLTL) instead of step-recovery diodes (SRD) used in the maser approach. As usual, the SC is a direct multiplication from the 10 MHz of the OCXO to about 7GHz; the first multiplication stages are the most critical one: we have a distribution amplifier to provide the clock output, then a x9 multiplier, followed by a x2 multiplier and finally a PLL that filters the multiplication spurs. The phase noise of the OCXO plus the SC is shown in Fig. 5; it has been obtained by the beat note of two nominally identical devices. The Dick effect turns out $7\text{x}10^{-14}$ at 1s.

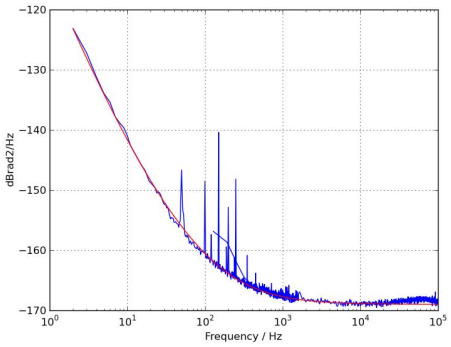


Fig. 5. Phase noise of the SC.

IV. CHARACTERIZATION

The assembled system has been tested using two different optical wavelengths, the Rb D1 line (795nm) and the Rb D2 line (780nm): Optical pumping on the D1 line is achieved with a DFB laser, using a compact and frequency-stabilized laser head build by LTF-UniNe and described in [4].

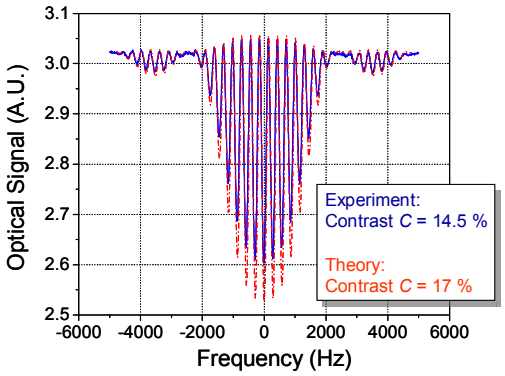


Fig.6. Ramsey fringes with D1 line; pumping time t_p = 3.7 ms; microwave pulse duration t_1= 400 µs; Ramsey time T = 3ms; detection time t_d = 150 µs, laser power during pumping P_L^{pump} = 1.4 mW, P_L^{det} = 300 µW.

Fig. 6 refers to a comparison between experimental (blue line) and theoretical (red line) Ramsey fringes obtained with D1 line. The agreement between the two curves is very good; in particular, the contrast that is one of the key parameter is around 15 %. When the local oscillator is locked on the central fringe of the Ramsey pattern we obtain the frequency stability of Fig. 7. The measured Allan deviation is $4.2 \times 10^{-13} \tau^{-1/2}$ up to average times of 10000 s. the bump around 2000 s is due to the thermal fluctuations not completely understood and solved at the time of the measurement.

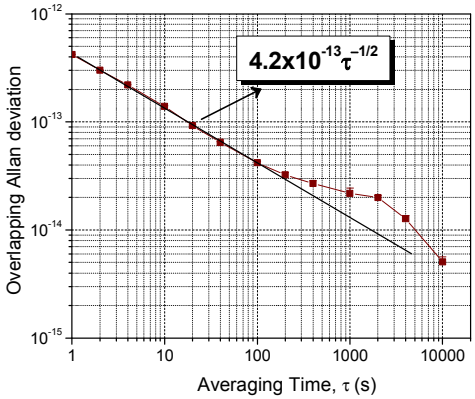
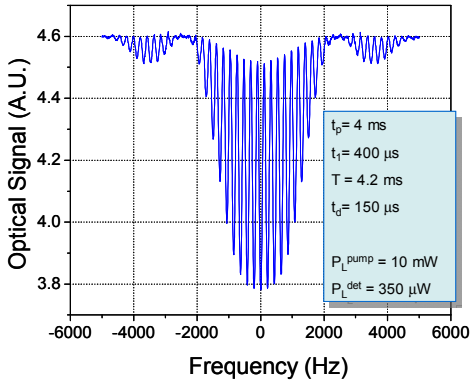
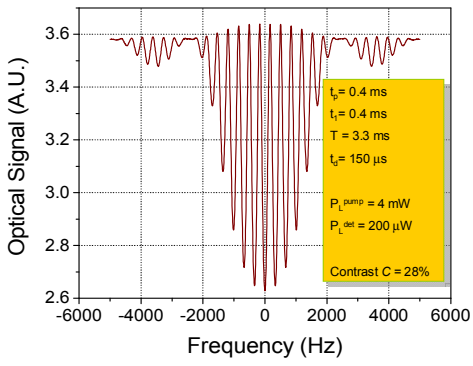


Fig. 7. Allan deviation of the POP clock with optical detection using D1 as pumping light.

The same characterization measurements have been done using a DFB laser locked to the D2 line. Fig. 8 shows the observed Ramsey fringes for two different temperatures.



(a)



(b)

Fig. 8. Ramsey fringes observed with D2 line; (a) temperature T= 58 °C; (b) T= 68 °C. The timing and the adopted laser powers are indicated in the figures.

In particular, a contrast of the 28 % for the central fringe is achieved. The corresponding frequency stability is reported in the following figure 9 and now we have 1.7×10^{-13} at 1s and

the white frequency region noise goes down and reaches $5x10^{-15}$ at 10000s. This is, in our opinion, due to some residual environmental effects.

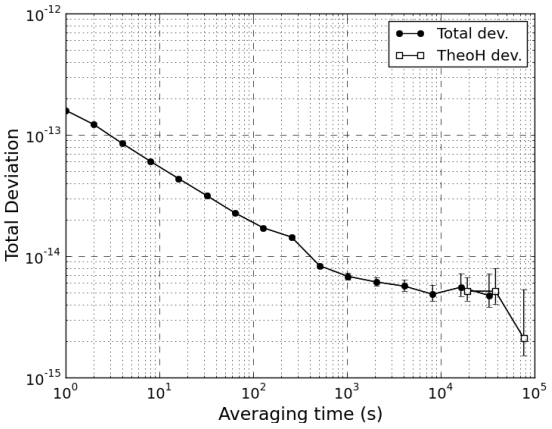


Fig. 9. Allan deviation of the POP clock with optical detection using D2 line as optical pumping.

The drift has not been removed from the data and turns out $-3x10^{-15}$/day.

In conclusion, we have implemented and characterized a Pulsed Optically Pumped frequency standard with optical detection that confirms the significant benefit in the short term stability predicted by the theory. Both the frequency stability and the drift are already fully compliant with the Galileo PHM specifications We might cite the article by Belloni et al. on PHM here, Procs EFTF2010.

ACKNOWLEDGMENT

This work was supported by the European Space Agency ESA (contracts 21504/08/NL/GLC and 19392/05/NL/CP), the Swiss Space Office, and the Swiss National Science Foundation. We thank E. Bertacco (INRIM) for his invaluable help, M. Pellaton for filling the Rb cells and G. Mileti, T. Bandi, P. Scherler, and J. DiFrancesco (all LTF) for their contributions.

REFERENCES

[1] A. Godone, S. Micalizio, and F. Levi, "Pulsed optically pumped frequency standard", Phys. Rev. A vol. 70, p. 023409 (1-11), 2004.

[2] A. Godone, S. Micalizio, F. Levi, and c. Calosso "Physics characterization and frequency stability of the pulsed rubidium maser", Phys. Rev. A vol. 74, p. 043401 (1-12), 2006.

[3] S. Micalizio, A. Godone, F. Levi, and C. Calosso, "Pulsed optically pumped Rb vapor cell frequency standard: a multilevel approach", Phys. Rev. A vol. 79, 013403 (1-19), (2009).

[4] T. Bandi, G. Mileti, F. Gruet, C. Affolderbach, C. E. Calosso, "Investigations on Improved Rb Cell Standards", Proc. joint IEEE International Frequency Control Symposium and European Frequency and Time Forum, San Francisco (CA), USA, May 2011.

The Influence of Laser Polarization Variations on CPT Atomic Clock Signals

Michael Huang and James Camparo

Physical Sciences Laboratories
The Aerospace Corporation
2310 E. El Segundo Blvd., El Segundo, CA
james.c.camparo@aero.org

***Abstract*— The Coherent-Population-Trapping (CPT) vapor-cell atomic clock represents a clear break from previous vapor-cell clock technology. As such it has advantages that may result in significant frequency stability improvements. However, along with the technology's benefits come new effects that may limit the CPT clock's capabilities, and one such effect is laser polarization noise. Here, we present our first results looking at the response of a CPT signal to modulation of a laser's polarization. Our results show that the CPT clock signal undergoes a splitting when the polarization modulation is rapid, suggesting that various Fourier components of a field's stochastic polarization fluctuations will affect the CPT clock signal in qualitatively different ways.**

I. INTRODUCTION

In the CPT clock, the microwave atomic signal is generated in an all optical fashion [1]. Briefly, sidebands are placed on the optical carrier at one-half the frequency of the ground state hyperfine splitting, $\Delta\nu_{hfs}$, (i.e., 6.8 GHz for Rb^{87}), and when the separation between the two sidebands, 2δ, matches the ground state hyperfine splitting both levels are simultaneously coupled to the same excited state. It is a feature of the atom's quantum nature that excitation pathways interfere, and when the two hyperfine states are coupled to the same excited state in this fashion the pathways interfere destructively (i.e., the atom cannot make a real transition to the excited state). If the atoms cannot be excited, there is no absorption of the laser light, and consequently the intensity of light transmitted by the vapor increases. The transmitted light thereby acts as a monitor of the atoms' interaction with a microwave signal (i.e., the laser sideband separation), and this can be used to lock the frequency of a quartz crystal oscillator to the atom's ground state hyperfine splitting.

One obvious advantage to this design is that the microwave cavity of the traditional vapor-cell clock is eliminated. In addition to allowing for "chip-scale" atomic clocks [2], elimination of the microwave cavity removes all microwave power shifts that can give rise to frequency instability [3]. In particular, effects like alkali surface migration on the resonance cell's glass walls [4], which affect the microwave cavity Q, are eliminated. Moreover, fewer components in general translate into greater device reliability. Nevertheless, though the CPT clock has real advantages over the conventional laser-pumped atomic clock [5], its novel nature suggests that it may be subject to new processes that play little, if any, role in the conventional laser-pumped clock. One of these processes concerns the effect of laser polarization fluctuations on the CPT clock signal [6], which the present work was initiated to investigate.

II. THE PROBLEM

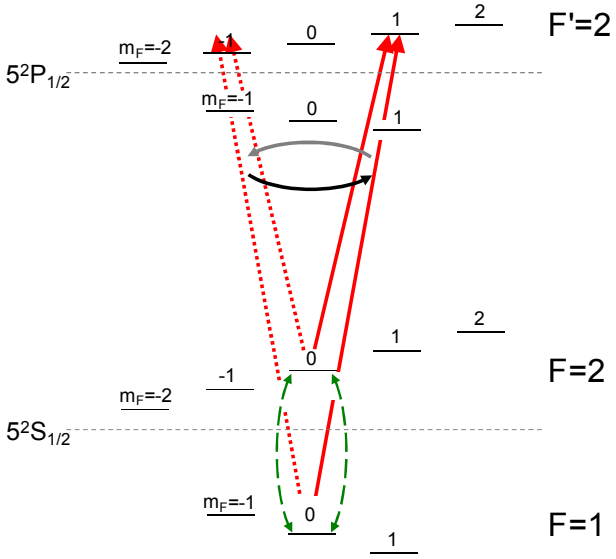


Figure 1: In the typical CPT clock, two modes of a laser couple an atom's two $m_F = 0$ ground-state sublevels to the same excited state; here, the common excited sate is F' = 2, $m_F = +1$. The simultaneous coupling creates a coherence in the ground-state, indicated by the dashed green lines, and it this coherence that is at the heart of the CPT phenomenon. If the laser polarization fluctuates, then the common excited state will momentarily change to F' = 2, $m_F = -1$, and this will affect the ground-state coherence.

In the most typical realization of the rubidium CPT clock, the laser light is circularly polarized and connects the (F = 2, $m_F = 0$) and (F = 1, $m_F = 0$) ground state Zeeman sublevels to the (F' = 2, $m_F = +1$) Zeeman sublevel of the $5^2P_{1/2}$ excited

This work was funded by U.S. Air Force Space and Missile Systems Center under Contract No. FA8802-04-C-0001.

state; this is illustrated in Fig. 1. As is well-known, this "lambda-system" coupling creates a coherence between the two $m_F = 0$ ground state sublevels (i.e., the 0-0 transition), and it is this coherence which is at the heart of the CPT signal. If the circular polarization switches from right-circularly polarized light to left circularly polarized and back again, then a transient will be introduced into the CPT signal since the common excited state changes: $(F' = 2, m_F = +1) \rightarrow (F' = 2, m_F = -1) \rightarrow (F' = 2, m_F = +1)$. If the polarization variations are stochastic [7], then the time series of these CPT transient will appear as noise on the CPT signal.

III. Experiment

Figure 2 shows a block diagram of our experimental arrangement, which is fairly straight-forward. We start with a cleaved-facet Fabry-Perot diode laser, which does not suffer from intrinsic polarization fluctuations [8]. Though not shown in the figure, the laser light passes through an electro-optic modulator (EOM), which places sidebands on the laser at 3.4 GHz. The modulated and linearly polarized field then passes through a ferroelectric liquid crystal (FLC) polarization rotator that has a bandwidth of 10 kHz. This polarization rotator changes the field's polarization by ninety degrees depending on an applied voltage supplied by a function generator. The field then passes through a quarter-wave plate, creating right or left circularly polarized light depending on the sense of linear polarization transmitted by the FLC polarization rotator. With the laser polarization square-wave modulated, the 3.4 GHz sideband spacing is very slowly swept through the CPT resonance (i.e., $2\delta = \Delta\nu_{hfs}$). The CPT resonance is observed as an increase in the transmitted laser light, measured by a silicon photodiode, and we average the CPT resonance lineshape using an averaging oscilloscope. We employ a buffer gas cell, containing 50 torr of N_2, so that the dominant relaxation mechanism in our experiments is electron-spin randomization (also known as S-damping) [9].

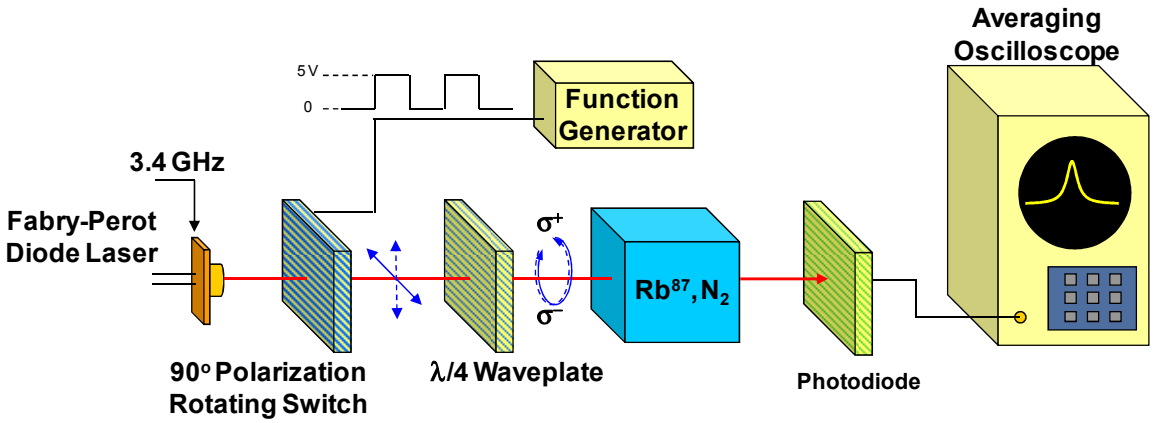


Figure 2: Block diagram of our experimental arrangement. The Rb cell is maintained at 47 °C, corresponding to an alkali density of ~ 10^{11} cm^{-3}, and the Rb atoms are contained with a 50 torr N_2 buffer gas. The laser light passes through an electro-optic modulator (not shown) that places sidebands on the laser at ~ 3.4 GHz. In our experiments the sideband-to-carrier power ratio is 0.17.

IV. Results

Based on our previous work [6] and that of Jau et al. [10], we had certain expectations concerning the results of our experiments. For modulation frequencies, f_m, less than the electronic spin-polarization relaxation rates in our system, we expected the atom to respond quasi-statically to the modulation: the atom would reach steady-state equilibrium between laser polarization changes, and the CPT phenomenon would appear to be unaffected by the polarization modulation. However, as f_m became larger than the rate of spin-polarization relaxation, the atom would not have enough time between laser polarization variations to generate an equilibrium spin-polarization in the vapor. In other words, the laser would not have enough time to put atoms into a "trapping state" [11], and so we expected the CPT resonance to grow in this regime of modulation frequencies. Finally, as f_m became larger than the 0-0 coherence's dephasing rate, we expected the laser polarization variations to begin to have an effect on the CPT lineshape. In particular, we expected the CPT linewidth to grow monotonically (though perhaps not linearly) with f_m in this regime.

Figure 3 shows the main results of our work. There, we show several cases of the CPT resonance for different values of the modulation frequency. At low modulation frequencies, $f_m \sim 3$ Hz, we see the standard CPT resonance. The "noise" on the CPT signal is not noise at all, but a residual of the transients that are induced in the absorption cross section as the laser's polarization changes [6]. At intermediate modulation frequencies, $f_m \sim 30$ Hz, we see an increase in the CPT signal amplitude. As discussed above this is expected, since these modulation frequencies are beginning to exceed the spin-polarization relaxation rate. Surprisingly, however, as f_m exceeds the 0-0 coherence dephasing rate (i.e., $\gamma_2 = 120$ Hz [6]) we don't see the CPT linewidth broaden. Rather, the CPT linesphape splits into a doublet with the separation between the two CPT resonances increasing with the modulation frequency. This is shown more quantitatively in Fig. 4, where we have plotted the center frequency of each doublet component, f_{CPT}, as a function of the modulation frequency. In the figure, the two straight lines correspond to $f_{CPT} = +f_m$ and $f_{CPT} = -f_m$.

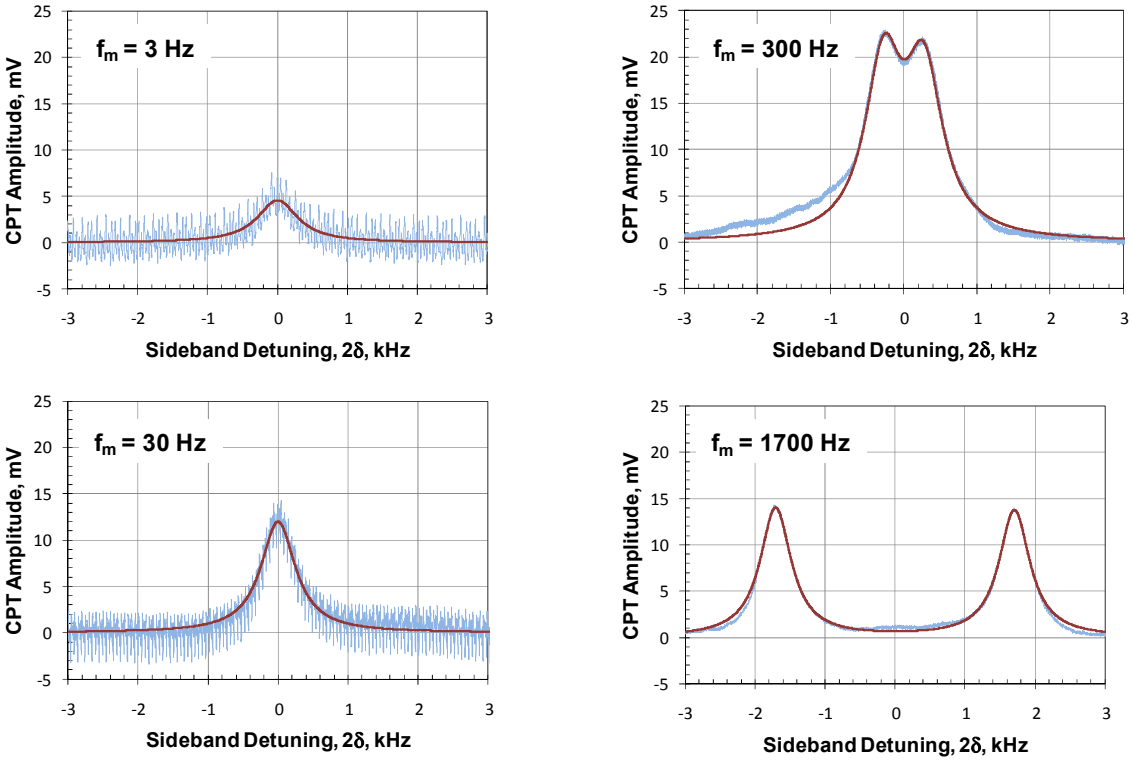


Figure 3: This figure shows our principle experimental results. At low modulation frequencies the standard CPT resonance is obtained. At intermediate frequencies, the CPT amplitude increases since atoms don't have time to be trapped in a non-CPT participating atomic state. At the highest modulation frequencies, our CPT resonance splits into a doublet.

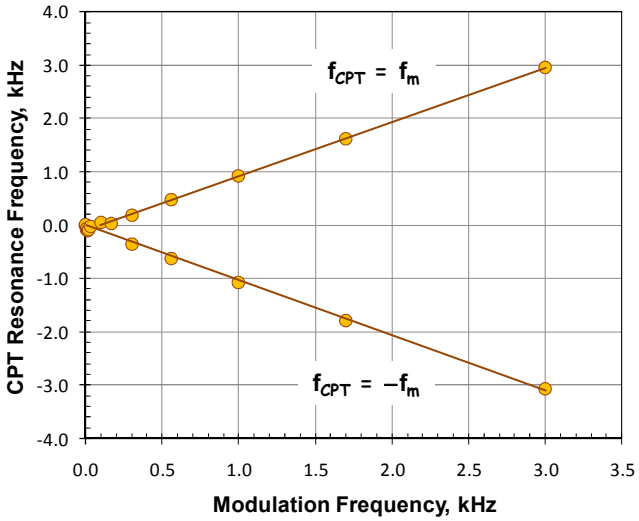


Figure 4: The frequency of each doublet component as a function of modulation frequency. For these results, the laser intensity was four times larger than that yielding the highest contrast CPT resonance in the absence of polarization modulation. Note that the doublet splitting equals $2f_m$.

V. Discussion

As noted in the Introduction, when the laser changes from right-circularly polarized light to left-circularly polarized light the atom is coupled to different Zeeman sublevels of the excited state. As a consequence, the Rabi frequencies change, not in magnitude but in phase. Interestingly, our theoretical analysis (to be discussed in a subsequent publication) shows that only one of the Rabi frequencies changes sign after a laser polarization change, the other is unaffected; and it is this asymmetry in the phase change of the Rabi frequencies that yields the CPT splitting. Briefly, since only one of the Rabi frequencies changes sign, the ground-state 0-0 coherence suffers a phase change following each laser polarization change, which happens twice every modulation period. This modulation of the ground-state coherence at $\pm f_m$ yields a doubling of the CPT resonance frequency: $\Delta\nu_{CPT} = \Delta\nu_{hfs} \pm f_m$. If both Rabi frequencies changed phase, then there would be no net phase modulation of the 0-0 ground-state coherence, and no splitting would occur.

Clearly, the results of the present work have implications for the effect of *random* polarization variations on CPT signals, and these are indicated (very qualitatively) in Fig. 5. For ease of discussion, let f_c be a critical laser polarization modulation frequency: for $f_m < f_c$ no splitting of the CPT resonance occurs; for $f_m > f_c$ the CPT resonance is split into a doublet. As suggested by Noise-Spectrum A, if all Fourier components of the noise spectrum are less than f_c, then laser polarization noise could actually increase the amplitude of CPT resonances. Specifically, those Fourier components less than f_c, but larger than the spin-polarization relaxation rate, would cause the CPT amplitude to increase. In other words, laser polarization noise could enhance a chip-scale atomic clock's signal-to-noise ratio. However, if the high frequency

cutoff of a polarization noise spectrum is above f_c, then as suggested by Noise-Spectrum B the CPT resonance will be reduced in amplitude and it will have an increased linewidth. Thus, for a noise spectrum like B a chip-scale atomic clock's signal-to-noise ratio and its atomic-Q would be reduced. Whether or not these qualitative expectations hold up will be the subject of further research in our laboratory.

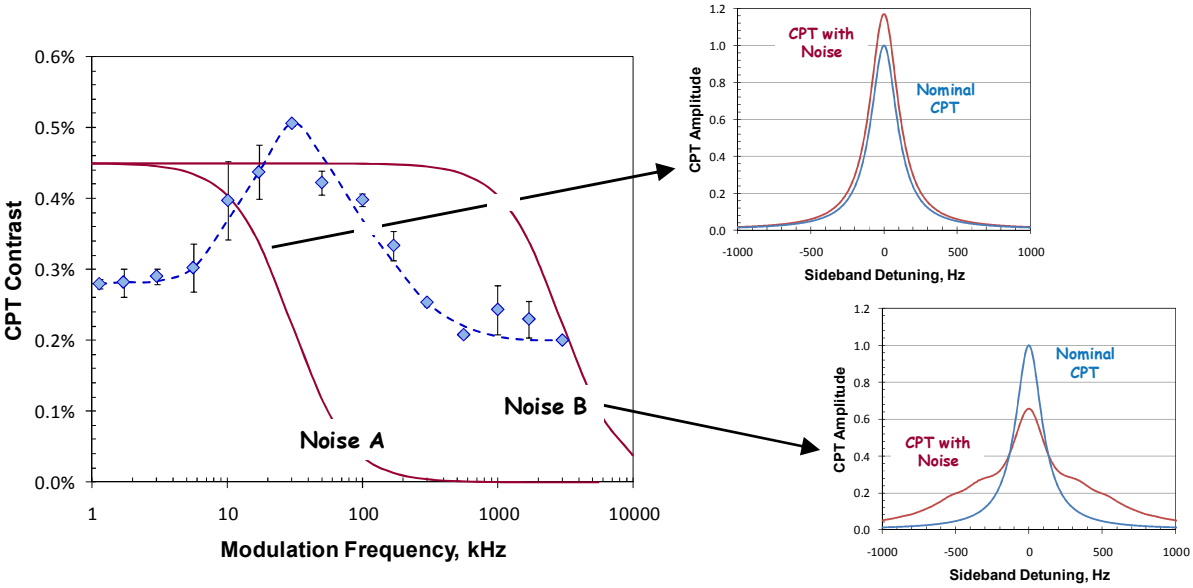


Figure 5: Illustrative considerations of the effect of laser polarization noise on CPT signals. The data points in the figure to the left show the amplitude of the CPT resonance as a function of modulation frequency; for these data the laser intensity maximized the CPT contrast in the absence of polarization modulation. (The dashed line is simply an aid to guide the eye.) In effect, the data points can be thought of as an atomic "transfer function" for laser polarization noise. Assuming that a laser polarization noise spectrum is similar to curve A in the leftmost figure, we would predict the CPT resonance to appear like that shown in the figure to the upper right (nominal implies the absence of laser polarization noise). Alternatively, if the laser polarization noise spectrum was more closely modeled by curve B, we would expect the CPT resonance lineshape to take on the appearance shown in the curve to the lower right.

REFERENCES

[1] J. Vanier, "Atomic clocks based on coherent populatin trapping: A review," Appl. Phys. B, vol. 81(4), pp. 421-442, 2005.

[2] S. Knappe, V. Shah, P. D. D. Schwindt, L. Hollberg, and J. Kitching, "A microfabricated atomic clock," Appl. Phys. Lett., vol. 85(9), 1460-1462, 2004; R. Lutwak, J. Deng, W. Riley, M. Varghese, J. Leblanc, G. Tepolt, M. Mescher, D. K. Serkland, K. M. Geib, and G. M. Peake, "The chip-scale atomic clock – low-power physics package," in Proc 36th Annual Precise Time and Time Interval (PTTI) Meeting. Washington, DC: US Naval Observatory, 2004, pp. 339-354.

[3] A. Risley, S. Jarvis, Jr., J. Vanier, "The dependence of frequency upon microwave power of wall-coated and buffer-gas-filled gas cell Rb^{87} frequency standards," J. Appl. Phys., vol. 51(9), pp. 4571-4576, 1980.

[4] J. G. Coffer, B. Sickmiller, and J. C. Camparo, "Cavity-Q aging observed via an atomic-candle signal," IEEE Trans. Ultrason. Ferroelec. & Freq. Control, vol. 51(1), pp. 139-145, 2004.

[5] J. Vanier, M. W. Levine, D. Janssen, and M. J. Delaney, "On the use of intensity optical pumping and coherent population trapping techniques in the implimentation of atomic frequency standards," IEEE Trans. Instrum. Meas., vol. 52(3), pp. 822-831, 2003.

[6] M. Huang, J. G. Coffer, and J. C. Camparo, "CPT transients induced by rapid changes in laser polarization: Validation of a semi-empirical model," J. Phys. B: At. Mol. Opt. Phys., vol. 43, 135001, 2010.

[7] J. Kaiser, C. Degen, W. Elsässer, "Polarization-switching influence on the intensity noise of vertical-cavity surface-emitting lasers," J. Opt. Soc. Am. B, vol. 19(4), pp. 672-677, 2002.

[8] J. C. Camparo, "The diode laser in atomic physics," Contemp. Phys., vol. 26(5), pp. 443-477, 1985.

[9] F. A. Franz, "Relaxation mechanisms in optical pumping," Phys. Rev., vol. 141(1), pp. 105-112, 1966.

[10] Y.-Y. Jau, E. Miron, A. B. Post, N. N. Kuzma, and W. Happer, "Push-pull optical pumping of pure superposition states," Phys. Rev. Lett., vol. 93(16), 160802, 2004.

[11] J. Vanier, M. W. Levine, D. Janssen, and M. Delaney, "Contrast and linewidth of the coherent population trapping transmission hyperfine resonance line in ^{87}Rb: Effect of optical pumping," Phys. Rev. A, vol. 67, 065801, 2003.

Positioning FBAR Technology in the Frequency and Timing Domain

Rich Ruby (IEEE Fellow)
Wireless Semiconductor Division
Avago Technologies
San Jose, CA, USA
rich.ruby@avagotech.com

Abstract— this paper will focus on the technical differentiation of FBAR from other mechanical resonator technologies for timing applications. The talk will also touch on a recent modification of FBAR (the Zero Drift Resonator, ZDR) that is temperature compensated. A further technology differentiator is the size of the chip-scale packaged resonator. Using the same technology used in manufacturing Avago Technologies' filter duplexer (over 3 billion sold) allows us to leverage qualified processes and thus go to market quickly. Furthermore, the silicon lid is perfectly suitable for placing integrated circuits and this is being done. All of these factors, (wide tuning range, high Q, high frequency, small size, integrated circuitry) are being used to differentiate potential products for the Time and Frequency Markets.

I. INTRODUCTION

In John Vig's tutorial can be found the salient facts and an excellent overview of the contribution of the quartz resonator in the fields of frequency and timing standards [1]. Highlights include

- Applications and Requirements
- Hierarchy of Oscillators (and thus, their applicability)
- Sources of Oscillator Instability

Between 1930 and 1950, the quartz crystal became the 'heart' of commercial time keeping, and by 1972 wristwatches using quartz resonators become widespread.

Today, required stability for USB 2.0 and 3.0 Host Controllers is +/-300 ppm. Stability for Wi-Fi and other wireless applications is +/- 20 ppm. For GPS, the stability must be on the order of 1 to 2 ppm.

Recently, several technologies have begun to target some of those markets traditionally held by Quartz resonators and oscillators. MEMs resonators and oscillators (in the 50 to 100 MHz frequency range) have gotten a lot of press in the past 3 years. Q's on the order of 60,000 to >100,000 [2,3] have been reported. Another technology, Solid State Oscillators (SSO)[4], have also entered the market. This technology is quite different in that it does not use a mechanical resonator of any kind. Rather, SSOs rely on an integrated LC tank circuit (typically running at 2 to 4 GHz) and taking advantage of the property that phase noise improves by -20 log[N] dB as the frequency is divided down by N.

One of the chief drawbacks of the MEMs oscillator is the lack of any real coupling (i.e. the ratio of the motional capacitance to the plate capacitance). Poor coupling reduces the ability to pull the high-Q resonator. So, given that process and temperature variations determine the final frequency, the MEMs oscillator is used as a non-deterministic reference. The output frequency is generated using a follow-on fractional N synthesizer with programmable offsets to compensate for frequency offsets due to process variation and to compensate for the rather large temperature coefficient of frequency (on the order of -30ppm/C) of the silicon MEMs resonator.

The SSO can be pulled quite easily; as the Q of the inductor L is 10 to 20. The resonant frequency is typically chosen to be 3 -4 GHz and then divided down. The divided down frequency is compared to a pilot tone at the factory and offsets are similarly stored into a look-up table to compensate for temperature variations and frequency offsets due to process variations. A varactor as part of the LC oscillator then adjusts the frequency. Influence from external sources of parasitic capacitance is minimized by fabricating a Faraday cage around the die.

The FBAR is a mechanical resonator more akin to quartz resonators than MEMs type devices. However, there are several key differences between all 3 devices:

i) The mode of operation is a longitudinal motion (sound waves normal to the plane of the resonator) as opposed to a flexure mode used by MEMs devices

ii) The thin sputtered piezoelectric and metal electrode films, set the native frequency which is much higher than quartz resonators with their relatively thick plates

iii) The coupling coefficient for FBAR is at least 30X larger than quartz resonators

978-1-61284-110-6/11/$26.00 ©2011 IEEE

iv) The coupling coefficient for FBAR is at >100X larger than MEMs electro-statically driven resonators

However, the f*Q products of these mechanical resonators in commercial applications are all within a range of 10^{12} -10^{13} Hz.

Another technology has recently been the subject of investigation. This would be the Rayleigh-Lamb wave devices. Here, the MEMs silicon bar is replaced by an AlN piezoelectric film. Inter-digitized fingers patterned on top generate an SO symmetric mode (see Fig. 3). The advantage of this device over a MEMs device is the coupling coefficient is much higher than silicon MEMs devices. The coupling coefficient ranges from 1 to 2% for non-temperature compensated devices and 0.56% for temperatures compensated devices (similar to quartz) [5,6].

II. Thin Film Bulk Acoustic Resonator

A. Background

There are two 'flavors' of FBAR, the Free-standing Bulk Acoustic Resonator (or FBAR) and Solidly Mounted Resonators (or SMR-BAW). FBARs were first developed in the '80's [7,8]. In the mid '90's SMR devices came into vogue [9]. At HP, later Agilent, later Avago, we focused on FBAR and in 2001 announced a PCS duplexer [10,11]. Due to the inherent advantages of Bulk acoustic devices over surface acoustic devices (SAWs), FBAR became a major contributor to many of the cell phone refinements that we take for granted today (e.g. small size, internal antennas, more functionality due to lower power dissipation from the duplexer etc…). Part of the staying power of FBAR at Avago was the introduction of a wafer-scale package technique in 2002 that allowed for an all-silicon chip scale package that was hermetic and could be either integrated directly on the customers board or on a 'molded chip on board (MCOB)' at our factory.

Q enhancements, die-shrinking and new topologies have allowed Avago to be a leading supplier of duplexers, multiplexers and Front-End Modules (FEM) that integrate PA and power coupler with duplexer. Some of the more stringent LTE filter requirements have driven Avago to develop a 'zero drift' version of our resonator, or ZDR. So, besides 'screaming' roll-off, some LTE bands require the filter not move over temperature.

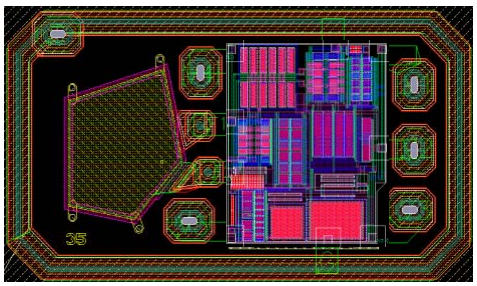


Fig. 1. Diagram of Avago's FMOS Device. Both the resonator and the circuitry are integrated into the same die

Currently, our ZDR resonators can be 'tuned' such that the linear coefficient of frequency with respect to Temperature is zero'd out. This is done with an oxide layer buried inside the stack [12,13]. We target a Turn over Temperature (TOT) in the range of 30 to70°C. The residual quadratic temperature dependance, β, ranges from -15 to -25 ppb/ $^{o}C^{2}$, and depends on stack make up.

More recently, we have begun adding active circuitry in the lid (Fig. 1). The circuitry can either be CMOS (hence, the name 'FMOS') or Bipolar. We have demonstrated both. The circuitry is connected to the resonator by two narrow 'drop-down' posts that form the connection at the same time that the base wafer containing the FBAR is bonded to the lid wafer containing the circuitry. A description of the manufacturing process is given elsewhere [14].

The value of this method of manufacturing (base wafer – FBAR, lid wafer – circuitry) is that all aspects, all process steps, are fully qualified, fully vetted and do not interfere with each other. The latter property is most important, if one has to re-invent each process so as to not interfere with the other technology, then, the whole process must –once again—be fully qualified (an onerous undertaking).

B. FBAR

Experimental data on varying sizes of resonators show that the Q behavior has a $(Area)^{-3/2}$, not $(Area)^{-1}$. The extra -1/2 in the exponent comes from this area-to-perimeter ratio. This power law observation points to where we can find energy losses.

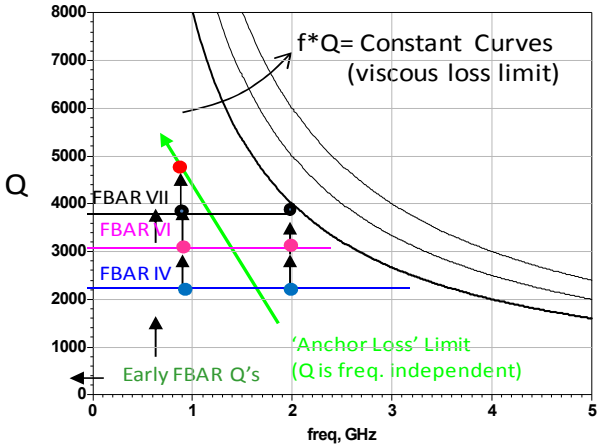


Fig. 2. Measured Q of Avago FBAR Resonators plotted over time and by generation (FBAR IV, V, VI, VII…) suggesting that 'anchor point' losses dominate. At some point resonator Q losses will be dominated by intrinsic or viscous losses.

We have reason to believe that energy leaking out from the perimeter ('Anchor' Losses) is the main loss mechanism for FBAR (at least up to 2 GHz). Figure 2 is a plot of Q vs. Frequency of Avago's standard FBAR resonator (developed over time) and includes a family of f*Q=constant curves (these curves represent the behavior of a mechanical resonator when the losses are dominated by intrinsic, material losses). It is well known that resonator losses due to phonon scattering, friction or heat generation will generate f*Q= constant

behavior. In particular, extracting attenuation, $\alpha(f)$ from the fundamental, the 3rd and the 5th harmonics shows that $\alpha(f) \sim \alpha_1 * f + \alpha_2 * f^2$ [15]. Thus, for a given velocity and resonance;

$$\frac{1}{Q} = \frac{2\alpha(f_o) * v_{ac}}{\omega_o} \sim \alpha_1 + \alpha_2 * f_o$$

(Units of α's assumed to have correct dimensions)

Figure 2 suggests that for fundamental frequencies up to 2 GHz, resonators made in any number of ways (that is the ratio of layers and the materials chosen for those layers – e.g. Mo or W electrodes) have a Q limit that is frequency independent [16].

In an ideal world, an AC voltage applied across the electrodes at or near resonance of an FBAR would generate only a longitudinal motion. However, due to the non-zero cross terms in the piezoelectric stress tensor, lateral modes can and will be generated as well. Specifically, there are a variety of Rayleigh-Lamb (RL) modes that can exist at any given frequency.

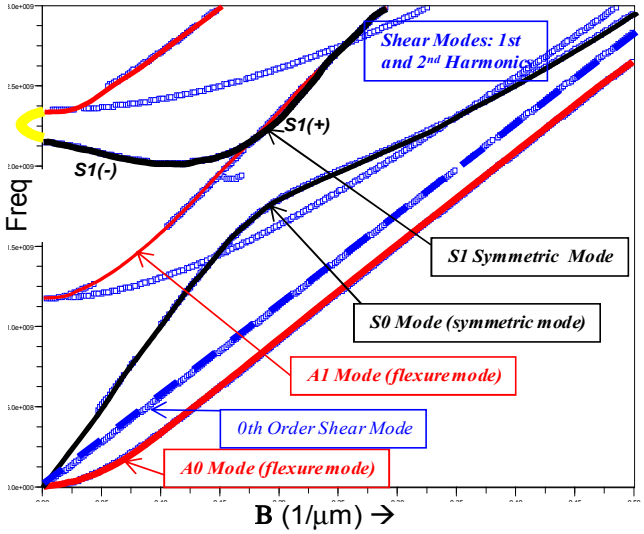


Fig.3. Calculated Dispersion Curves with the Rayleigh-Lamb modes labeled for an AlN Plate 2.5 um thick. Experimental extractions of dispersion curves from spatial Fourier transforms show remarkable agreement between experiment and theory.

Figure 3 shows the modeled dispersion curves of a AlN resonator. All of the allowed mode branches are labeled [17]. The dispersion curves are generated by doing spatial Fourier transforms on measured displacements mapped across a given resonator surface [18]. Longitudinal energy converted into lateral modes can scatter into 'matching' modes in the surrounding silicon substrate (anchor points) [19]. Once energy is deposited into the silicon it is lost and hence, the Q of the FBAR resonator is limited. Various methods exist to improve Q given this reality and some are given in [20]. To date, the maximum Q of an FBAR at 1 GHz is ~5000 (4000 at 2 GHz). The method of measurement is given in [16] and also discusses some of the inherent difficulties when measuring high Q resonators at GHz frequencies.

There are four outstanding attributes of the FBAR resonator that make it ideal for high performance filters; high Q, high kt^2_{eff}, high power handling, and the use of silicon as the substrate (with the attendant advantages of using main stream production tools and the ability to form an all-silicon hermetic package).

Unfortunately, there is a finite temperature coefficient of frequency (TCF) that is on the order of -30 ppm/°C. Thus, to use an FBAR resonator in an oscillator has only limited application, namely modest tuning range VCOs for PLLs. But, that application is still potentially interesting.

Unlike Quartz BAW or SAW devices, FBAR using AlN as the piezoelectric has an effective coupling coefficient on the order of 6.5 to 7.1% (in the crystal world, one uses the variable $r = C_o/C_m$, whereas, FBAR/SMR papers refer to $kt^2_{eff} = 1.2*C_m/C_o$, so a kt^2_{eff} of 6.8% corresponds to an r = 17.7 -in 'crystal' speak). For a Quartz resonator, r typically ranges from 60 to 80 (depending on cut).

The coupling coefficient, kt^2_{eff} can be defined as

$$kt^2_{eff} = \frac{4.8 * (f_p - f_s)}{(f_p + f_s)}$$ where f_s and f_p are the frequencies where the measured Q circle crosses the real axes of the Smith Chart (see Fig. 6a).

With minor modifications, one can 'tune' the kt^2_{eff} down to as low as 2% (this has several advantages: lower pulling range improves at least one source of externally generated phase noise – noise on the varactor, the resonator is less 'influenced' by the variations on the oscillator core (known as PVT, Process, Voltage and Temperature variations), the unloaded Q increases for decreasing kt^2_{eff}, and in simulation, the quadratic term for the temperature dependence goes down.).

The trade-off between Q and kt^2_{eff} can be understood using the 'Bode' equation for Q developed in [21,22]. The unloaded Q is defined by the stored energy divided by the power dissipated in a cycle.

$$Q = \frac{2\pi f \tau_{GD} |\Gamma_{11}|}{1 - |\Gamma_{11}|^2}$$

The stored energy is represented by the group delay τ_{GD} where $\tau_{GD} = -d\phi/d\omega$. The total phase between f_s (the series resonance) and f_p (the parallel resonance) is always fixed at 180 degrees. Therefore, as f_p and f_s get closer in frequency the $-\Delta\phi/\Delta\omega$ term increases. This is partly why low coupling devices have intrinsically higher Q than a device with a large coupling coefficient. [For completeness, the power dissipated in a cycle is $(1 - |\Gamma_{11}|^2)/(2\pi f)$].

C. FBAR Pulling Range

The 'pulling' range of a crystal is a small fraction of its coupling coefficient. Naively, one can think of adding parallel capacitance to the resonator until the parallel resonance matches the series resonance. However, for a variety of reasons that involve the external gain circuitry as well as the $kt^2_{eff} * Q$ product of the resonator plus varactor, one can only 'pull' the frequency a small fraction (<< 30%) of $\Delta f = f_p - f_s$. We have demonstrated a pulling range (while maintaining

high Q) of 8600 ppm using a ZDR resonator with a kt^2_{eff} of ~5%. We believe that with a standard FBAR, we could increase the pulling range to over 10,000 ppm. In the Quartz BAW and SAW world, the pulling range is very small, typically 10 to 50 ppm. The one T&F application where the FBAR looks to very interesting is the use of a high frequency FBAR VCO locked to a low frequency (~10 MHz) quartz crystal [23,24,25,26]. The value of an FBAR PLL locked to an external crystal (where the FBAR resonator is at a much higher frequency) is that the overall phase noise far from carrier can be improved immensely by avoiding frequency multiplication, while close-in, the phase noise is influenced by the quartz reference, as demonstrated in [25,26]. These two properties may find an application where excellent phase noise (jitter) and modestly-wide pulling range are needed.

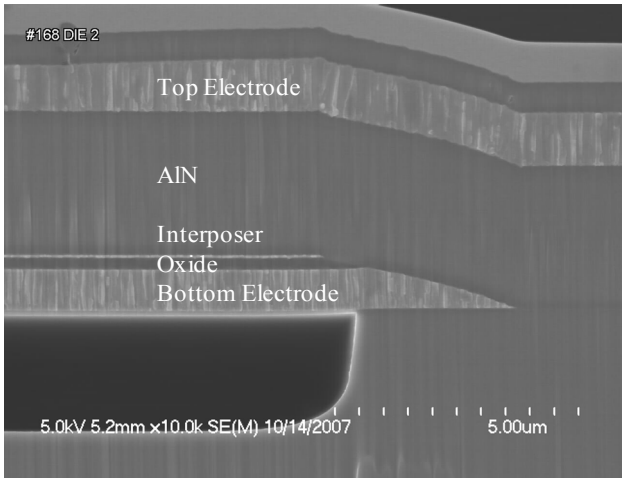


Fig. 4. Cross section of a Zero Drift Resonator (ZDR). The edge of the 'swimming' pool is shown. The layers are labeled.

Many clocking applications need a 'dither' applied to the clocking signal. The 'dither' moves the frequency up and down by a predetermined amount (usually in a saw-tooth fashion) 'spreading' the energy found in each of the harmonics over a finite frequency space. These kinds of products are known as SSXOs or Spread Spectrum Crystal Oscillators. They are used in electronics where some energy of the clock harmonic lies in a sensitive band (e.g. GPS) designated by the FCC or other governing agency as being 'off-limits'. The alternative, a Faraday cage around the electronics, is not always workable.

Typical SSXOs use the Quartz based oscillator as a reference. A Frac-N synthesizer consisting of a LC oscillator in a PLL generates an output frequency that is tied to the Quartz reference. Slowly changing divide ratios move the output frequency up and down. The fractional-n synthesizer can burn as much as much as 20 to 100 mW of power. The phase noise is determined by the LC oscillator and not the Quartz reference. An FBAR PLL dispenses with the LC oscillator and can directly be pulled up to +/- 5000ppm.

D. Zero Drift Resonators (ZDR)

References [12,13] describe a modified FBAR where a layer of oxide is buried inside the bottom electrode. This thin layer has a positive Temperature Coefficient and can be used to 'tune out' the linear TCF term of the FBAR resonator. Figure 4 shows a cross section of such a device (patent pending). To date, the best Q achieved with ZDR devices is on the order of 3000 (768 MHz resonators). Work is in progress to improve this. It is not clear if the device Q is limited by the intrinsic losses of the added oxide layer, or due to non-optimized 'anchor' losses.

Figure 5 shows the temperature dependence of two ZDRs. The TOT (or Turn Over Temperature) can be varied from 0 °C and 100 °C depending on the thickness of the oxide where all other layer thicknesses are fixed.

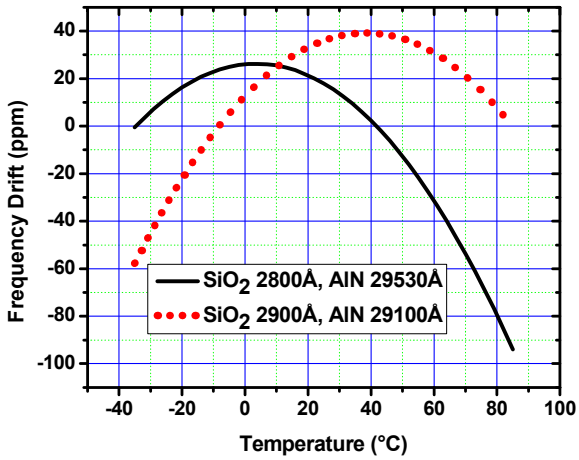


Fig. 5. Frequency Drift (in ppm) vs. Temperature for two ZDR devices with different oxide thicknesses under the metal interposer. Thicker oxide moves the TOT up in temperature. The β for these two devices is ~ -18 ppb/°C

III. ZDR/FBAR Resonators

Both FBAR and ZDR can be fabricated with the fundamental resonance as low as ~400 MHz and higher than 10 GHz. Furthermore, Q and the power handling abilities of FBAR (or ZDR) allow excellent phase noise when used in an oscillator. In an earlier section there was a discussion of FBAR oscillators used as VCOs in PLL type devices. For ZDR oscillators, we aspire to be the frequency source. Thus, the target is accuracy, stability, Q and reliability. Beyond these four 'musts'; the need for product differentiation is also very important. Product differentiation can come from power handling, turn on time and chip-scale integration.

A. Accuracy

Today's quartz crystals and FBAR filters are trimmed to frequency in much the same way. Raw quartz blanks are ion milled with a high energy ion beam of argon atoms – knocking off quartz atoms and thus slowly increasing the native frequency of the quartz blank. A similar process is done on FBAR resonators on wafer; the only difference is the magnitude of the Δfrequency/Δthickness. In the case of a quartz resonator, the quartz is a fraction of a millimeter thick; the thickness of the FBAR resonator is on the order of microns. The native frequency of a quartz plate might be 5 to

50 MHz, the native frequency of an FBAR resonator might be 400 MHz to 10 GHz.

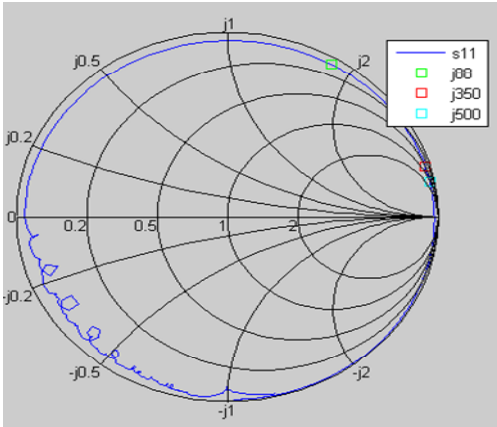


Fig. 6a, Measured Q-circle of a 768 MHz ZDR resonator. Markers show the reactance points where the oscillator core might conjugate match the impedance

A simplistic view of relative accuracy can be described as the ratio of the amount of frequency change, Δf (Hz), per atomic layer removed by argon etching. For a millimeter thick substrate, the Δf (Hz) per atomic layer removed is very small. For a 1 micron thick substrate, the amount of Δf will be three orders of magnitude larger. A quartz resonator can be purchased with a native accuracy of 10 to 30 ppm. An FBAR, after ion milling the resonators can be 'tuned' to 100 ppm 1-σ at 1.5 GHz. If the accuracy required is 1 to 2 ppm (GPS applications), then neither quartz nor FBAR resonators can be 'tuned' accurately enough. External circuitry driving a varactor or switched capacitors must be used to pull the resonator/oscillator. If an FBAR had the same coupling coefficient as quartz, then at most, the pulling range would be 10 to 100 ppm. This is much less than necessary to achieve the needed accuracy due to process variation. However, FBAR coupling is 3 to 6% and allows for 5000 to 10,000 ppm pulling range, compensating for any process variations.

B. Stability

Because the linear term of the TCF in a ZDR resonator can be 'tuned out' and a TOT set to temperature in the center of the operating temperature range, one is left with predictable quadratic temperature dependence. In theory, this can be taken out with electrical feedback, or in some cases, the variation in temperature is allowable for certain applications. External circuitry can be used to counteract the parabolic temperature dependence. This was demonstrated for a ZDR oscillator using a PTAT and a look up table that controls a DAC attached to a varactor [27].

The problem with this technique is that the short term stability is poor due to constant shifting in and out different discrete voltages on the varactor. Future enhancements to this technique might be to add a high frequency 'dither' to the LSB of the DAC. Another circuit approach might be to use an on-chip diode whose voltage (for a fixed current) varies smoothly with temperature and whose output is fed into an Op Amp with programmable gains [28].

One important lesson of ZDR, is that with mechanical compensation only, a small varactor across the resonator is needed to electrically compensate for the residual quadratic error. The tuning range of the varactor need only be +/- 150ppm. Compensation for process errors can be 'tuned' out (for example) by using CMOS switched capacitors in a Pierce oscillator topology (for C_1 and C_2) where the relative losses of the switching caps is relatively small and one can cover a larger frequency range in a linear fashion. This avoids the use of a large and lossy varactor to tune over large frequencies.

C. Quality Factor Q

The Q's of ZDR have not (yet) achieved the same numbers as FBARs made in our traditional line of filters (when comparing comparable frequencies). There are two schools of thought as to why this is so. 1) The oxide is acoustically lossy, bringing the f*Q curves in Fig. 2 lower. 2) Mainstream FBAR filters for duplexers bring in millions of dollars in revenue and due to the competitive landscape; much effort and resources continue to be spent on improving Q. This has not yet been the case for ZDRs. Although many if not all techniques for improving Q for mainstream FBAR are available to ZDR, it still takes time and resources to catch up.

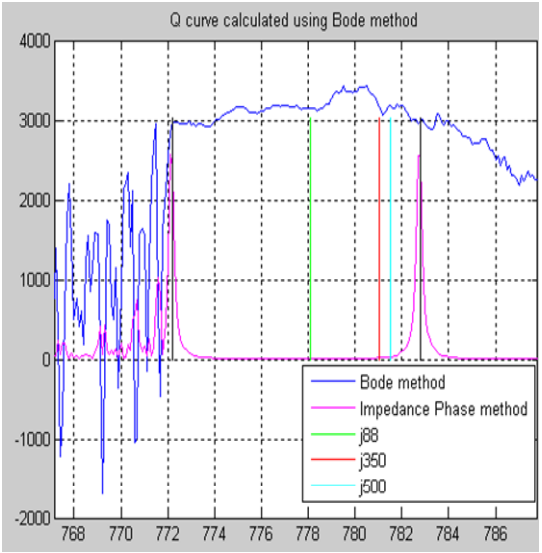


Fig. 6b, The extracted Q vs. freq using the Bode equation [21,22]. The magenta markers show Q at fs and fp using Dicke's equation for Q. The R,B,G markers match the reactance markers shown in Fig. 6a and mark where the frequencies are located for a given reactance.

Figure 6a shows the Q circle measured on our 768 MHz, 54K μm^2 ZDR resonators. The markers highlight points where the oscillator designer may wish to target the conjugate reactance of his oscillator circuit. Simple trade-offs can be made (or made for the oscillator designer). For example, more parasitic capacitance will find the conjugate match of the resonator closer to f_s (even lower than j88 Ω). A 'tightly' designed oscillator with little parasitic capacitance will operate closer to f_p. This in turn, will reduce the power used in the oscillator and allow for more pulling range.

Figure 6b is the Q vs frequency using the Bode equation. Also, the standard Q measurement for Q at f_s and f_p via Dicke's impedance phase technique is presented. Both methods are in agreement at at f_s and f_p [29]. The markers correlate to the markers in Figure 6a.

D. Reliability

The value of having 30,000 to 100,000 die on a wafer is that we can get a very good statistical sense of what the medians, spreads and outliers look like for die subjected to various strife tests. Typical strife is 85 ^{o}C /85% R.H. for hundreds to thousands of hours. Another strife subjects die to temperatures ranging from 125 ^{o}C to 150 ^{o}C in 85% R.H. (but for shorter time than 85 ^{o}C /85R.H.). Vibration tests, autoclave (121C/100%RH for 96 hrs), high temperature storage (125 ^{o}C for 1000 hrs), thermal shock (1000 cycles -65 ^{o}C /150 ^{o}C) are also included in our strife [30]. Strife is a means to force parts to failure to better understand the failure mechanisms. This is done in anticipation of the qualification runs.

Figure 7a shows the combined measurements of 6 control wafers where approximately 2000 resonators per wafer are tested. All wafers are tested twice (wafers measured, removed from the testers, the tester is then re-calibrated and the wafers are place on again and re-tested. This simple algorithm eliminates much grief during strife testing. If the initial cal was bad, then conclusions moving forward become suspect. Test points 1 and 2 are the initial test and re-test. After the rest of the non-Control wafers are strife'd for a period of time (e.g. 120 hours at 125 ^{o}C in 85% R.H. -- HAST), all the wafers are re-tested. The control wafers give an indication if –at the time of testing—there is some issue with the tester. The control wafers are tested at each strife milestone.

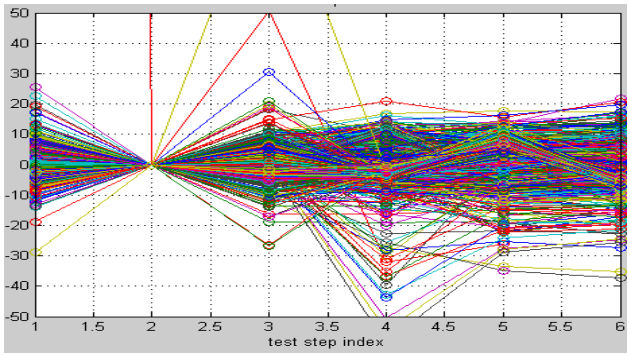


Fig. 7a, The measured frequency differences (in ppm) of 6607 randomly picked die off of 6 control wafers. The x-axes define the 6 test points when the wafers under test are taken out of the environmental chambers and tested. The median does not change (expected), but the spread degrades (presumably from probe damage). We see a 1σ spread going from 2.5ppm to 4.7 ppm. Note that several 'errant' dies go out of spec and then come back to the fold.

Figure 7a shows just how much the spread increases as the wafers are repeatedly tested. This suggests probe damage. In fact, in Fig.7b, the strife'd wafers (non-control) were 'tortured' in HAST environment and although the medians did not move after a 'burn-in' phase the 'spread' in individual resonators grew. The upshot is that guaranteeing an aging number for a raw resonator is limited by the spread. Interestingly, one can trace many individual resonators moving 10's if not 100's of ppm up or down from one test point to the next, but, then 'coming back to the fold'. This suggests another source of error, electro-mechanical interference (EMI) of some kind occurring during test. Going to an oscillator, we expect much better data with less spread and less sensitivity to probe damage and EMI.

One large drawback to qualifying resonators is that the Network Analyzers (NWAs) used to measure the frequencies do not have the necessary guaranteed stability and accuracy of their phase measurement. Comparing resonators on different NWAs can introduce relatively large offsets. Typically, a NWA only offer phase accuracy to a degree and this can translate to large ppm swings. The data showed in Fig's 7 are that measured on the same NWA.

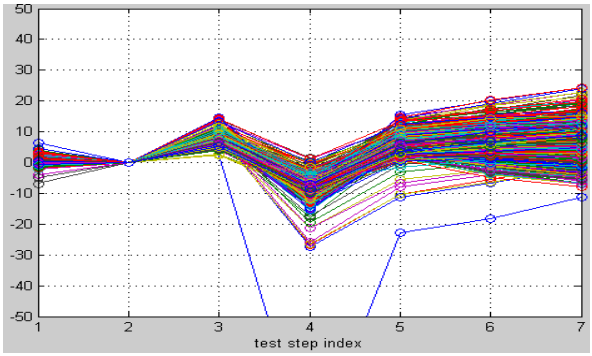


Fig. 7b, The measured frequency differences (in ppm) of 1839 die measured on a wafer undergoing HAST (130C/85%RH). The parts are taken out at stated intervals and measured. Die are screened at testpoint 3 and we see the affect of accrued testing at 48, 96, 192 and 288 hours of HAST. After 96 hours, we see the median stabilize, but, the 1σ spread continues to increase. Hard to explain is the errant die that goes way out of spec, but then returns to the fold.

IV. PROPERTIES OF FBAR, ZDR, AND OSCILLATORS

A. Phase Noise and Jitter

The phase noise measurement of a 600 MHz temperature compensated oscillator built in 2007-2008 gave some indication of the very good performance of ZDR oscillators [12]. The "knee" for the flicker noise onset is between 10 and 20 kHz. The core oscillator is an in-house, CMOS 0.4 um technology node. This oscillator draws 2 mA from a 3.3V supply voltage. Phase noise at 1 kHz offset is -102 dBc/Hz and -156 dBc/Hz at 10 MHz offset. The topology used here was a differential Colpitts design. An Agilent 5052B was used to measure the phase noise and to calculate the integrated rms jitter.

Figure 8a is a plot of SSB Phase Noise for a recently built 1 GHz discrete bipolar oscillator using a ZDR resonator in a Pierce topology. The phase noise was -103 dBc/Hz @ 1 kHz offset and -172 dBc/Hz at 10 MHz. The power was 23.5 mW (4.7V, 5mA). The resonator impedance was ~150 Ω (area =10K um^2). The Q was somewhere ~ 3000. A NEC 851 NPN silicon bipolar transistor was used as the gain element. The integrated jitter (10 kHz to 20 MHz) was measured to 9.4 fs. The better 'far-from-carrier phase noise (an additional 8 dB) comes from using more power.

Figure 8b is a SSB Phase Noise plot of a recently built 3.4 GHz discrete bipolar oscillator using a ZDR resonator with a Q of ~1000. The phase noise was -80 dBc/Hz @ 1 kHz offset. Compared to the 1 GHz oscillator (for the same offset), the difference in phase noise can be mostly accounted by -80 dBc - 20 log[(3.4/1)* (Q @ 1 GHz/Q @ 3.4 GHz)]. Thus, 'normalizing' the phase noise of the 3.4GHz oscillator w.r.t. 1 GHz, phase noise one obtains approximately -100 dBc or only ~3 dB unaccounted for. The point of this exercise is that Leeson's model predicts the behavior of two ZDR oscillators operating at different frequencies and no unexpected surprises.

At a 10 MHz offset, the phase noise was measured to be -156 dBc/Hz. The power was ~20 mW (Vdd=3V, I=6mA). The resonator impedance was ~20 Ω (area=10Kum^2). The RMS jitter was 15 fs. Careful analysis of the Q vs. frequency for our resonators [31] suggests that by further optimization of our circuit, we can take advantage of the region where our resonators' Q is a maximum.

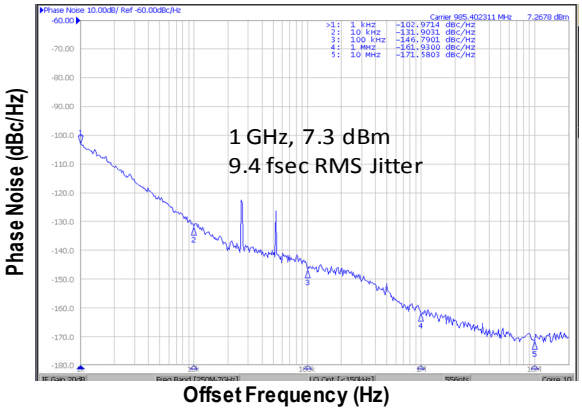


Fig. 8a. Phase Noise measured of 1 Ghz Pierce Oscillator

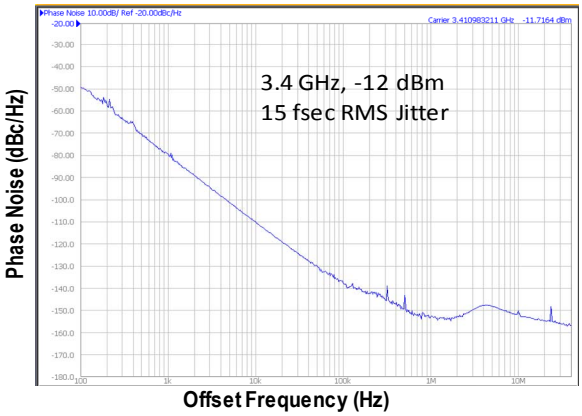


Fig. 8b. Phase Noise measured of 3.4 Ghz Colpitts Oscillator

B. Q ,Power, Area and Frequency Considerations

Given the range of frequencies FBAR/ZDR can cover, it is tempting to target high frequency oscillators and then divide down to the desired frequency; similar to the strategy chosen by the solid state oscillators (we assume that dividers are noise, power and cost free – more on this later). Starting at a 2X higher frequency than our target frequency, f_o, and then dividing by 2X would seem to give -20 log(2) = -6 dB phase noise improvement over a resonator resonating at f_o, for a fixed frequency offset f_m. If Q is dominated by 'anchor point' losses (i.e. Q is constant w.r.t. frequency) then, from Leeson's model, we know that phase noise is proportional to +20 log[f_o /Q] for a given offset frequency, f_m. Therefore, starting at a 2X frequency and then dividing down by 2 does not gain anything. If the resonator losses are intrinsic (1/f dependence), then, starting at 2X higher frequency means that the Q will be 2X worse. Dividing back down to f_o there will be a net 6dB degradation in phase noise for a fixed offset.

Again, comparing the two oscillators in Fig. 8; if one uses the higher frequency oscillator and a 'perfect' divide by 3.4, the improvement in phase noise at 1 kHz offset would be 20 log[3.4] or ~11 dB improvement. Thus the 3.4GHz ZDR oscillator with a divide by 3.4 would have an output frequency of 1 GHz, with -91dBc @ 1kHz offset compared to the ZDR oscillator running at its fundamental 1 GHz and a phase noise offset at 1 kHz of -103 dBc.

It is the additional loss due to Q in the Leeson model (-20 log[3.4] = -10.6 dB) that teaches us to avoid higher frequencies if the loss mechanism is frequency dependent. Other sources of phase noise probably come from additional losses in the oscillator core having to run at 3.4 GHz, and the choice of transistors. The moral of this story is to go for lower frequency ZDR reesonators.

However, the trade-offs in frequency then become the trade–off in area (area goes as f^2 for constant impedance resonators), native accuracy, and impedance matching to the external oscillator circuit. Area is cost sensitive and smaller resonators at higher impedances have their own issues (see earlier discussion). Dividers (if necessary) are not free and cost both power and added phase noise (especially if not designed correctly). High frequency resonators will not be frequency centered as accurately as low frequency resonators, thus, more pulling range is necessary to offset process variations.

A CMOS divide by 2 circuit in 0.4 um technology roughly dissipates 0.5 mW dividing a 800 MHz frequency in half [32]. With each node, the power dissipated decreases linearly, thus a 0.13 um CMOS technology node, the divider might dissipate one third the power, or 0.14 mW etc.. . There is a power/Area trade-off for any oscillator where dividers are used to obtain the final output frequency. To be clear, going to lower frequencies lowers the divider power consumption, but, for the same impedance, the area of the resonator grows as f^{-2}.

Other trade-offs with high frequency oscillators is the power needed to drive the output buffers. A typical load assumed by Quartz oscillator is 10 pF. This will account for both pad capacitance of the CMOS oscillator, any trace capacitance associated with the PCB and the pad capacitance of those ICs requiring ESD. Power is proportional to frequency and capacitance. Lower frequencies and/or means to minimize the pad and trace capacitance will lower the power budget associated with the output driver.

The core oscillator power of an FBAR or ZDR oscillator is proportional to R_p the real resistance seen by the oscillator

circuit at the parallel resonance at f_p (note, f'_p for a resonator in a circuit is 'pulled' down from f_p, of the unloaded parallel resonance due to the capacitance of the circuit and the varactor). It can be shown from the Vittoz equations that the condition for oscillation g_m*R_p ~ 4 means that higher R_p's allow for a lower g_m [33]. The lower the gain needed to sustain oscillation, the lower the power dissipated by the core oscillator. We have demonstrated FBAR oscillators (Q near 5000 and Rp near 7000Ω) dissipating ~ 20 µW at 2 GHz [34].

C. Turn On Time

The turn-on time of an oscillator is proportional to Q/f_{osc}. Thus, for a constant f*Q product, a higher oscillation frequency is desirable from a start-up time perspective. As expected, As expected FBAR/ZDR exhibit a start-up time orders of magnitude faster than a quartz oscillator. Figure 9 shows (in the time domain) the turn on time of an FBAR oscillator. With Q is the 1000 to 5000 range and frequency at or near 2 GHz, the turn on time should be on the order of microseconds (µsec) compared to milliseconds for a quartz crystal resonator with Q's on the order of 60,000 and operating frequency on the order of 20 MHz.

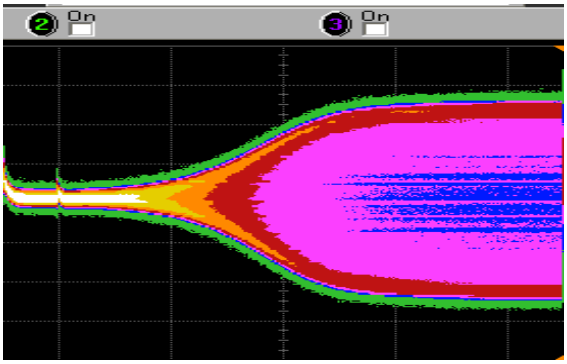


Fig. 9. Turn-on of a 2 GHz Oscillator (measurement courtesy of B. Otis). Each division is 2 us and turn on is ~ 8 us.

This is useful for those applications where the clock can be turned off for power savings, but quickly turned back on when needed.

D. Phase Noise in the Presence of Vibration

Three 1.5 GHz ZDR Oscillators were built up and tested on a shaker table to evaluate the sensitivity to acceleration. Measurements were made by Mike Driscoll and Noah Hughes at Northrup-Grummen. Sensitivity ranged from 1 to 80 ppb/g), and are comparable to SMR-BAWs made by TFR [35] and better than whispering gallery resonators [36]. There were some unexplained 'jumps' in frequency shift measured below 80 Hz and above 800 Hz. It was speculated that this was due to noise from the cables.

To convince ourselves that we were not seeing microphonic-generated noise from the resonator, we used the Acoustic Imaging Microscope [18] to map out each of the flexure modes from 10 KHz to 5 MHz (for the given boundary conditions in x and y) of the A0 mode shown in Fig. 3. As predicted in Fig.3, we expect that the resonant frequency of the fundamental to be inversely proportional to the length squared. We looked at the fundamental 'drum' mode (as it appears in the AIM) for two different areas (35K um^2 and 54K um^2) and saw that these resonances (899 KHz and 542 KHz respectively) did scale inversely with area (length squared).

The area of the resonator/oscillator measured by Hughes and Driscall was 30K um^2 and the first 'drum' mode resonance should occur around 1 MHz. Hence, we believe that the noise seen by their measurement at 1 to 2 KHz and below 80 Hz is not due to microphonics of our device.

V. FMOS

A. Description

In [14], we report a 1.5 GHz bipolar FMOS oscillator with a phase noise of -80 dBc/Hz at 1 kHz offset. The oscillator core uses 0.3 mA of current at 1.8V (0.54 mW). Due to a correctable process error, the Q was quite low and varied from 500 to 800 across the wafer (not the more typical ~1600 for on-chip ZDRs). Thus, we would have expected about -86 to -88 dBc/Hz at 1 kHz offset. Still, this is very good phase noise for a 0.54 mW DC power draw and an oscillator operating at 1.5 GHz. In discrete ZDR oscillators operating at 1.5 GHz, we have obtained ~ -103 dBc/Hz at 1 kHz offset. But, here the power used was ~44X higher; thus, in line with Leeson's model.

FMOS integrates much of the necessary oscillator core into the lid wafer. We have demonstrated working dividers, varactors, temperature sensing diodes, resistors and capacitors integrated into our lid – along with the oscillator core and buffer amplifiers [14]. The area allotted for these circuits is on the order of 300 by 300 μm^2 (as shown in Fig. 1).

Having a temperature sensing diode to monitor temperature directly from the die gives the promise of temperature compensation circuitry placed inside the lid. Thus, the oscillator can be made to both precise and stable over process variations and temperature.

Intuitively, it would seem that the overall yield of FMOS would be the product of the various yielding processes and would therefore be less than if the devices were separately built (note bonding occurs regardless of whether there is circuitry in the lid or not). However, this logic may not follow. An example of how adding circuitry to the lid will increase overall yield is the case where a working FBAR resonator is off-frequency. Parametrically, being off frequency is equivalent to being non-functional. With FMOS, a varactor (or switched capacitors) in the lid circuitry along with some non-volatile memory and a DAC (and of course, the oscillator core) could allow one to measure the FMOS oscillator frequency (on-wafer measurement), load in a correction term and thus have a working die – now on frequency.

At Avago, we are blessed with having both a qualified bipolar process (with an ft ~ 25GHz) and a 0.4 um CMOS process. Using the bipolar process, we can target FMOS toward high frequency and high performance VCOs with excellent jitter and phase noise. Using our CMOS process, we can target the fixed frequency market using NVM, switched capacitor techniques and low power dividers to generate Quartz oscillator frequencies.

B. Aging Data

Avago is in the process of evaluating long term aging properties of our ZDR. One of the inherent difficulties in determining aging is that multiple measurements are needed. Given that the accuracy and repeatability of measurement needs to be in the few ppm range, we are stymied by artificial effects such as accruing pad damage due to multiple contacts.

One elegant way around this problem is to integrate the

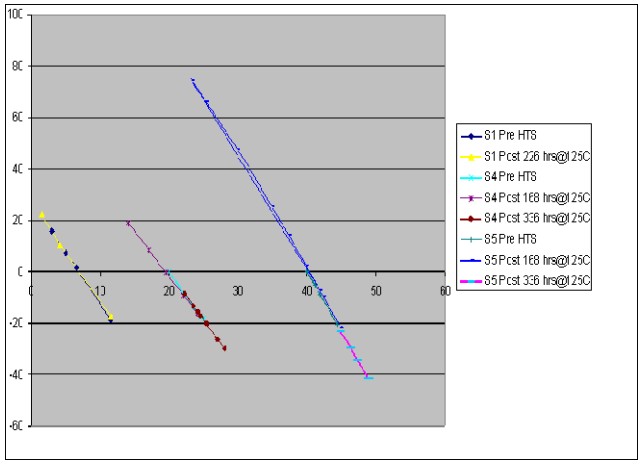


Fig. 10. Three FMOS oscillators measured, re-measured after 168 hours in 125C storage and re-measured again after 336 hours. The x-axes is relative temperature and the 3 devices are offset to better show the individual performances. The y-axes is in Δf/f ppm.

oscillator inside the ZDR microcap lid (i.e. FMOS as described above). Testing FMOS wafers (oscillators); loading correction factors, inking out dead die prior to singulation has inherent advantages over testing raw resonator die on wafer. First, the oscillator frequency (properly biased and buffered) measured on wafer will be the frequency seen by the customer. Testing for aging and/or quality issues (e.g. non-hermetic package) is much simplified since frequency measurement via a counter is inherently more stable and accurate (at the ppm level) than measuring complex impedances of the resonator via a Network Analyzer). Unlike the case of a resonator, probe/pad contact variation will not 'pull' the oscillator frequency. Strife and formal Qualification Runs hopefully can yield more deterministic data than raw microcap'd resonators.

Figure 10 shows three measured FMOS oscillators (individually mounted on pc boards) over a small temperature range before and after 300 hours at 125 °C. By 'dithering' the temperature, we eliminate the possible ppm excursions due to temperature variations that occur over time in the test lab. Temperature was measured using an integrated diode inside the FMOS circuit. Frequency was measured using a spectrum analyzer. As can be seen in Figure 10, the absolute frequency and temperature dependence are nearly indistinguishable between initial test and after 336 hours at 125 °C. We are now in the process of automating the frequency shift vs. strife (i.e. High Temperature Storage, HAST, vibration, etc….) on the wafer level.

C. Comparison to CMOS Solid State Oscillators

In a recent talk at ISSCC, 2011 [37], a 1.2 GHz oscillator using 65 nm CMOS was presented demonstrating 570 fs integrated jitter (Phase noise, here integrated from 1 kHz to 10 MHz) and burning 52 mW of power, the area was 0.12 mm^2. In their paper, they use a Figure of Merit [38]

$$FOM = 10Log[(\sigma/1s)^2(P/mW)]$$

and compare their work along with 7 other results showing FOMs ranging from -210dB to -227 dB.

In contrast, the discrete oscillator shown in Fig. 8b has an FOM of -253 dB, → 26 dB better. This is consistent with the 30dB improvement predicted the side-by-side LC- and FBAR-oscillator comparison presented in [39]. [One caveat; the device in Fig. 8a oscillates at 1 GHz not 1.2 GHz] When the discrete version is integrated in FMOS, we expect an area to be only 3 to 5X larger (including the hermetic package with pads). The conclusion here is that the performance of mechanical-resonator based oscillators still dominates over solid state devices.

VI. Conclusions

Currently, the mobile market has developed a huge thirst for low insertion loss filters with excellent isolation. As more bands open up around the world and the requirements on the filter response is 'tightened', the demand for better quality filters is a unremitting. Since 1993, Q's for FBAR have steadily improved. In the last two years, Qs have gone from 3000 to 5000 in our production lines at Avago.

Riding on the FBAR filter production coattails are our Zero Drift Resonators and in conjunction with in-house CMOS and Bipolar processes, FMOS devices. The goal is to find a technology niche that provides reasonable volume and ASP (average sale price). The value proposition is the small size and relatively high performance of our ZDR and FMOS oscillators. A low end application would be a USB3.0/e-SATA FMOS oscillator. The target area is 0.3 mm^2 with a thickness < 0.16 mm leading to a total volume of 0.04 mm^3. In comparison a commercially available miniature quartz resonator operating at the same frequency is 0.672 mm^3 or 17X larger. This size would allow the ASIC manufacturer to insert this oscillator inside their QFN package or do a 'die-on-die' stacking of the part, which reduces the net BOM of their customers as well as guaranteeing performance and size.

However, longer term, FMOS can be used as critical elements in ultra-small low power radios, GPS receivers, and other electronic devices where small size and low power are key differentiators.

Acknowledgment

The Author wishes to thank his colleagues Martha Small, Duane Fazen, John Larson, Marty Shipley, Jyrki Kaitila, Don Lee, Frank Bi, Qiang Zou, Reed Parker, Lori Callighan, Steve Ortiz along with the management team of Bryan Ingram, Mark Unkrich, Mark Anderson, Stan Strathman, John Choy, and Tina Lamers. Also, the Author wishes to thank the students under Prof. Brian Otis, of the University of Washington, Helen Zhang, Jian-Lei Shi, Andrew Nelson, Kannan

Sankaragomathi, Julie Hu, Ethan Shih. Shailesh Rai. Lastly, the Author thanks Mike Driscoll and Noah Hughes for their kind interest and measurements of acceleration sensitivity.

REFERENCES

[1] J. Vig, "Quartz crystal resonators and oscillators for frequency control and timing applications - A Tutorial", Rev. 8.5.3.2, 2006, http://www.ieee-uffc.org/frequency_control/teaching/vig/vig3_files/frame.htm

[2] Y-W Lin, S. lee, S-S Li, Y, Xie, C. Nguyen,"Series-resonant VHF micromechanical resonator reference owcillators", *IEEE Journal of Solid-State Circuits,* vol. 39, no. 12, December 2004, pp. 2477-2491

[3] YW. Lin, SS. Li, Z. Ren, and C. T.-C. Nguyen, "Low phase noise array-composite micromechanical Wwine-glass disk oscillator", *IEDM Tech. Digest, 2006*

[4] M. McCorquodale, "Self-referenced, trimmed and compensated RF CMOS harmonic oscillators as monolithic frequency generators," in IEEE *Int. Frequency Control Symp., 2008,* pp. 408-413

[5] J. Kuypers, C. Lin, G. Vigevani and A. Pisano, "Intrinsic temperature compensation nitride lamb wave resonators for multiple-frequency references", *International Frequency Control Symposium 2008*, pp. 240-249

[6] C. Lin, T. Yen, Y. Lai, V. Felmetsger, M. Hopcroft, J. Kuypers, A. Pisano, "Temperature compensated AlN lamb resonators", *IEEE Trans. on Ultrasoncis, Ferroelectrics and Frequency Control*, vol. 57, no. 3, march 2010,pp.524-532

[7] T. Grudkowski, J. Black, T. Reeder, DE Cullen, RA, "Fundamental-mode VHF/UHF minature acoustic resonators and filters on silicon", Applied Physics Lett., v. 37 1980 pp. 993-995

[8] k. Nakamura, H. Sasaki, H. Shimizu, "A Piezoelectric composite resonator consisting of a ZnO film on an anisotropically etched silicon substrate", Jap. Jour. Of Appl. Phys., vol 20 (1981), pp. 111-114 – From proceedings of 1rst Symposium on Ultrasonic Electronics, Tokyo, 1980

[9] K.M. Lakin, K.T. McCarron, and R.E. Rose, "Solidly mounted resonators and filters", IEEE Ultrasonics Symposium 1995, pp. 905-908

[10] R. Ruby et al., "PCS 1900 MHz duplexer using thin film bulk acoustic resonators (FBARs)", Electronic Letters, vol. 35, no. 10, pp.794-5, May 1999", Trans. Elect. Lettr., 1999

[11] R. Ruby P. Bradley, A. Chien, J. Larson III, "Thin Film Bulk Acoustic Resonators for Wireless Applications", Ultrasonics Symposia 2001, Atlanta, pp. 813-21

[12] W. Pang, R. Ruby, R. Parker, P. Fisher, M. Unkrich, and J. Larson, "A temperature-stable film bulk acoustic wave oscillator," Electron Device Letters, IEEE, vol. 29, no. 4, pp. 315–318, April 2008

[13] Qiang Zou et al., "High Coupling Coefficient Temperature Compensated FBAR Resonator for Oscillator Application with Wide Pulling Range ", International Frequency Control Symposia, June 2010, pp. 646-51

[14] M. Small, R. Ruby, R. Parker, S. Ortiz, "Wafer-scale packaged FBAR oscillators for commercial applications", IFCS, held San Francisco, May 2011

[15] *Private Communications with John Larson*

[16] R. Ruby, "Method of fitting Q-circles of measured mechanical resonators", Ultrasonics Symposia, Held in San Diego, October, 2010

[17] R. Ruby, J.Larson, C.Feng, S. Fazzio, "The Effect of Perimeter Geometry on FBAR Resonator Electrical Performance" International Microwave Symposium Digest 2005

[18] K. Telschow, V. Deason, D. Cottle, J. Larson, "UHF Acoustic microscopic imaging of resonator motion", Ultrasonics Symp., 2000, pp. 631-634

[19] F. Thalmayr, K. Hashimoto, M. Ueda, T. Omori, M. Yamaguchi, "Quantitative Analysis of Power Leakage in an FBAR Device at the Anti-Resonance Frequency", Japanese Journal of Applied Physics (2010), 49 (2010) 07HD11

[20] R. Ruby, "FBAR Resonators and Filters," Ch. 5, RF Bulk Acoustic Wave Filters for Communications, K. Hashimoto ed., MA: Artech House, 2009, pp. 117-158.

[21] D. Feld, et al.,, "After 60 years:a new formula for computing Q is warrented" , Ultrasonics Symposium 2008 (Beijing)

[22] Ruby, R, et al. "Method of extracting unloaded Q applied across different resonator technologies" IEEE Ultrasonics Symposium 2008 (Beijing), pp. 1815-1818

[23] J. Hu, W. Pang, R. Ruby, B. Otis, "A 750 uW 1.575 GHz Temperature-stable FBAR-based PLL", Radio Frequency Integrated Circuits Symp., 2009 RFIC , RM)4C-1

[24] H. Ito, H. Lakdawala, A. Ravi, S. Pellerano, R. Ruby, K. Soumyanath, K. Masu, "A 1.7-GHz 1.5-mW Digitally-controlled FBAR oscillator with 0.03-ppb resolution", Solid-State Circuits Conference, 2008, ESSCIRC 2008, pp. 98-101

[25] J. Hu, R. Parker, R. Ruby, B. Otis, "A Wide-tuning digitally Controlled FBAR-based oscillator for frequency synthesis", International , pp. 608-612Frequency Control Symposium, 2010

[26] J. Hu, R. Ruby, B. Otis, "A 1.56GHz Wide-tuning all digital FBAR-based PLL in 0.13m CMOS", Custum Integrated Circuits Conference (CICC), September 2010

[27] S. Rai, Y. Su, W. Pang, R. Ruby, B. Otis., "A Digitally compensated 1.5 GHz CMOS/FBAR Frequency Reference", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, no. 3, March 2010

[28] T. Adachi, "A Study on MOS Temperature Compensated Crystal Oscillator", Japan-Taiwan Workshop on Future Frequency Control Devices, March 2007, pp.19-24

[29] Principles of Microwave circuits, Montgomery, Dicke, and Purcell, IEEE, Electromagnetic wave series 25, 1948, chapters 5 & 7

[30] Electromagnetic wave series 25, 1948, chapters 5 & R. Ruby, L. Kekoa, "Wafer Level Packaging (WLP) of FBAR Rilters", 3rd International Symposium on Acoustic Wave Devices for Future Mobile Communication Systems, March 2007, pp. 101-103

[31] R.Ruby," Review and comparison of bulk acoustic wave FBAR, SMR technologies", Invited Paper, IEEE Ultrasonic Symposium, 2007,pp.1029-1040

[32] *Private Communication with B. Otis, A. Nelson, K. Sankaragomathi of Univ. of Washington*

[33] E. Vittoz, M. Degrauwe, S. Bitz,"High-performance crystal oscillator curcuits: theory and application ", *J. Solid State Cir.* v. 23, June 1988, pp. 774-83

[34] A. Nelson, J. Kaitila, R. Ruby, B. Otis, "A 22uW, 2.0 GHz FBAR Oscillator", *In Press*, submitted to the RFIC Conference to be held in July 2011 in Baltimore.

[35] S. Birdsall, P. Dever, J. Donovan, M. Driscoll, K. Lakin, T. Pham,"Measurement of Static and Vibration-Induced Phase Noise in UHF Thin-Film Resonator (TFR) Filters", IEEE Trans. Ultrasonics, Ferroelectrics, Freq. Control, v. 49, no. 5, May 2002 pg. 643-648

[36] *Private Communications with Mike Driscoll*

[37] A. Sai, T. Yamaji, T. Itakura, "A 570 fs Integrated-Jitter Ring VCO Based 1.21 GHz PLL with Hybrid Loop", Int. Solid-State Circ Conf., 2011, pg 98-99

[38] X. Gao, E. Klumperink, P. Geraedts, B. Nauta, "Jitter Analysis and a Benchmarking Figure of Merit for Phase Locked Loops", IEEE Trans. On Circ. And Systems, v. 56, no. 2, Feb. 2009, pp 117-121

[39] S. Rai, B. Otis, "A 600 uW BAW-Tuned Quadrature VCO Using Source Degenerated Coupling", IEEE Journal of Solid State Circuits, Vol 43, No. 1, pp.300-305, January 2008

Tunable RF SAW/BAW Filters: Dream or Reality?

Ken-ya Hashimoto[1], Shuji Tanaka[2], and Masayoshi Esashi[2]
[1]Department of Electrical and Electronics Engineering, Graduate School of Engineering, Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba-shi 263-8522 Japan
[2]Department of Nanomechanics, Graduate School of Engineering, Tohoku University
6-6-01 Aza-Aoba, Aramaki, Aoba-ku, Sendai, 980-8579 Japan
[1]k.hashimoto@ieee.org

Abstract—**This paper discusses possibilities of tunable radio frequency (RF) surface and bulk acoustic wave (SAW/BAW) filters. First, configuration of RF front-end in current wireless communication systems is surveyed, and research trends of the flexible RF front-end and software defined radio (SDR) are discussed. Second, various RF tunable filters are introduced, and we discuss how high performances are expected by the use of tunable RF SAW/BAW filters provided that above mentioned key technologies are developed. Finally, we explain what kinds of research subjects are going on in our projects for realization of the flexible RF front-end for the SDR.**

I. INTRODUCTION

Adding to ultimate down sizing, supporting multi-band and multi-standard operation is required in current mobile phones. Nowadays high-end phones support eight frequency bands for two communication standards[1]. This makes the radio-frequency (RF) front-end section of the transceiver very complex[2]. It seems very difficult to support further standards and/or frequency bands, although current RF-ICs are powerful enough to support them.

As a possible solution to relax this problem, the software defined radio (SDR)[3-7] in combination with the flexible (tunable) RF front-end[8-10] has been paid much attention. Majority of transceiver functionalities is realizable by skillful use of digital signal processing and digitally-controlled analog technologies. Nevertheless, development of high performance tunable RF filters is mandatory for the realization of really flexible RF front-end. They should be able to change the passband width and/or center frequency, preserving performances achieved by current RF surface and/or bulk acoustic wave (SAW/BAW) devices[11-15]. Although certain number of attempts was reported[16-21], achieved performances were pessimistically low for practical use.

Recently, the author's group proposed a variable RF filter composed of wideband RF SAW/BAW resonators and variable capacitors (VCs)[22], and showed that not only the filter bandwidth but also the center frequency can be tuned for a wide range provided that following three key technologies are developed; (a) interconnection of SAW/BAW resonators with multiple VCs and controlling circuitry, (b) extremely wideband SAW/BAW resonators, and (c) high quality factor (Q) VCs with wide tuning range. Recent advancement of related technologies is going to resolve these hard obstacles. For example, (a) the micro-electromechanical system (MEMS) technologies enables chip-on-chip integration among heterogeneous technologies, (b) new piezoelectric materials with extremely large electromechanical coupling factor were proposed, and (c) high performance ferroelectric VCs were developed.

In collaboration with leading-edge research labs in Japan, namely, SAW R&D division in Murata MFG, Co. Ltd, the wireless communication research center in the National Institute of Information and Communications Technologies (NICT), and other outstanding technical supporters, we started a project to develop a really flexible RF front-end including tunable RF SAW/BAW filters applicable to present and/or future communication systems such as the cognitive radio[23].

This paper reviews current status and future prospects of this project.

First, configuration of RF front-end in current wireless communication systems is surveyed, and research trends of the flexible RF front-end and SDR are discussed. Second, Various tunable RF filters are introduced and compared, and we show how high performances are expected by the use of tunable RF SAW/BAW filters provided that above mentioned key technologies are developed. Finally, we explain what kinds of research subjects are running in our projects for realization of the flexible RF front-end for the SDR.

II. WHY SAW/BAW DUPLEXERS ARE NECESSARY?

Fig. 1 shows a basic configuration of radio receivers. Receive (Rx) signals detected by an antenna (ANT) are down-converted to the base-band (BB) by the frequency mixing with the local oscillation (LO) after the band selection by the RF band-pass filter (BPF) and amplification by the low noise amplifier (LNA). Then target signals are selectively detected by a detector circuit after the channel selection by a BB low-pass filter (LPF).

Most important task of the RF-BPF is to suppress incidence of jammer signals to the LNA and mixer without badly affecting the noise figure (NF) of the receiver. At the frontend, the insertion loss causes NF deterioration of the same amount.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

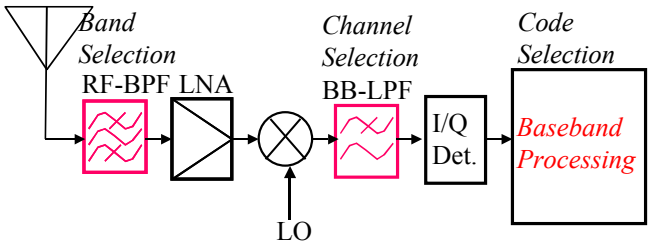


Fig. 1 Simplified receiver architecture

Thus ultimate reduction of the insertion loss is requested to the RF-BPF.

When a strong jammer signal is incident to the LNA and mixer, they may loose their gains (desensitization) and/or may generate in-band signals through the non-linear mixing. Thus the RF-BPF is also requested to offer good out-of-band rejection. However, the insertion loss and the out-of-band rejection are trade-off for the RF-BPF design in general, and we can not optimize both of them simultaneously.

Some wireless communication systems employ the frequency division duplex (FDD), which separate uplink (terminal to base-station) and downlink (base-station to terminal) signals in frequency domain. In the case, the strongest jammer for the receiver is transmit (Tx) signals generated by a transmitter circuit in the same terminal.

Fig. 2 shows the RF front-end of the FDD-mode transceiver. The ANT is also used for emitting Tx signals generated by a power amplifier (PA). An RF BPF is sandwiched between the ANT and LNA so as to reject unwanted signals detected by the ANT and/or those emitted by the PA. Another RF-BPF is sandwiched between the ANT and PA so as to suppress unnecessary signals generated by the PA. The insertion loss of this filter in the Tx band must be also minimized because it results in increased battery power consumption and heat up of the PA and the Tx filter.

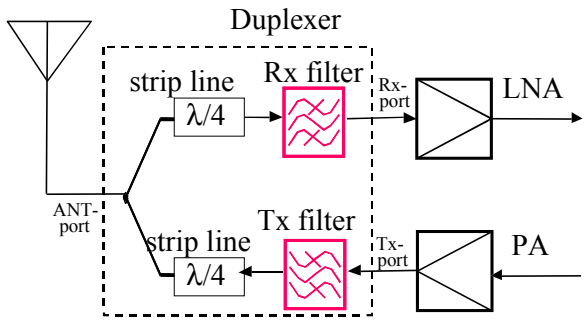


Fig. 2 RF front-end of FDD-mode transceiver.

The circuit block composed of two filters and two λ/4 strip lines is the antenna duplexer (DPX), which is one of the most important applications for RF SAW/BAW devices. Due to the frequency gap between the TX and Rx bands, we can suppress power leakage of the TX signal and noise output from the PA to the LNA to some extent. The strip lines are used to reduce electrical interferences between two filters, and are usually embedded into the DPX package.

Rx signals detected by the ANT can be extremely weak, say in –100 dBm order, which is slightly higher than the thermal noise level [2]. On the other hand, Tx signal emitted from PA can be strong, say in +25 dBm order[2]. From the W-CDMA specification, the receiver function should withstand jammer intensity up to –15 dBm. This indicates that nonlinearity of the DPX should be negligible.

Currently only the RF SAW/BAW DPXs fulfil all of these stringent requirements for the use in mobile phones, and thus are able to occupy the market share. DPXs employing dielectric resonators (DRs) exhibit excellent performances, and are widely used in base-stations. However, they are bulky and do not fit into compact mobile phones.

Until very recent year, interstage SAW filters has been inserted between the LNA and mixer to relax the nonlinear problem. Currently, many attempts so called "SAW-less" have been reported to remove them[2]. Then extremely high linearity is requested to the mixer which is exposed to jammer and leaked Tx signals without attenuation by the interstage SAW filter. This causes self-mixing of the modulated Tx leakage and generates unwanted signals at the baseband in addition to the DC component. This is called the cross modulation. For example, when the Tx output power is +25 dBm and the Tx-Rx isolation in the Tx band is 50 dB, the receiver circuit is requested to have extremely high input referred second-order intercept point (IIP2) close to +46.3 (≈110+2*(25-50-3)+α) dBm to suppress the non-linear output below -110 dBm[24]. In the equation, α (=-7.7 dB) is a correction factor for the modulated UMTS up-link channel[24]. This fact clearly indicates how deeply current mobile phones rely on the isolation given by the DPX.

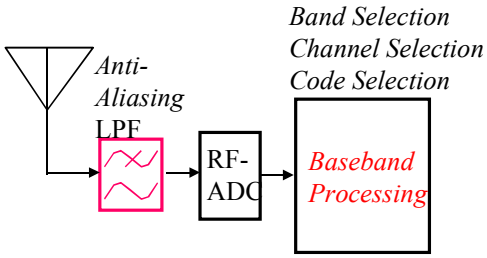
III. Software Defined Radio

An ideal solution to support multi-band and multi-mode operation is introduction of the SDR[3-7]. In an ideal SDR receiver shown in Fig. 3(a)[3], incoming RF signals are directly analog-to-digital converted, and all necessary signal processing including the band and channel selection is carried out in digital domain. Only a broadband low-pass filter (LPF) is necessary for anti-aliasing in the RF section. The operation mode can be changed and upgraded by uploading specific software. In an ideal SDR transmitter, outgoing RF signals are directly generated by a digital-to-analog converter (DAC) from date stream.

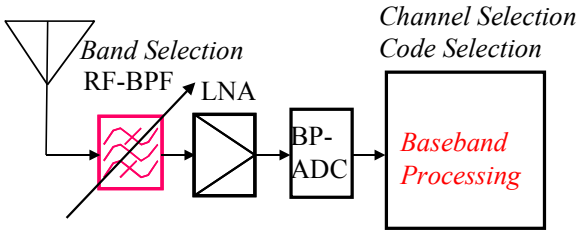
The biggest obstacle for the realization is implementation of high-speed, low-noise and high-resolution analog-to-digital and digital-to-analog converters (ADC/DAC). In high frequency ADC/DAC, the RMS signal-to-noise ratio (SNR) is limited by the clock jitter J_{RMS}[25] rather than the quantization noise, i.e.,

$$\mathrm{SNR} < -20\log_{10}\left[2\pi f_{\max} J_{\mathrm{RMS}}\right], \qquad (1)$$

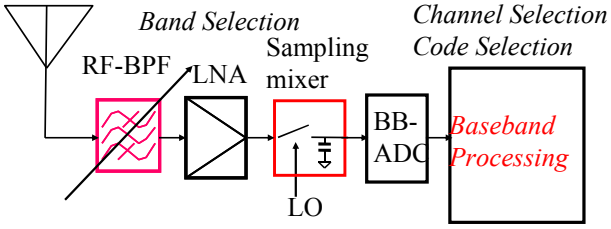
where $f_{\max}$ is the maximum signal frequency. Eq. (1) suggests that even when J_{RMS} is reduced to 0.1 ps, resulting SNR is limited to 58 dB for $f_{\max}$=2 GHz. This value is insufficient for mobile communications. This fact clearly indicates that the ideal SDR is not realistic in GHz range.



(a) Ideal scheme



(b) Under-sampling scheme



(c) Down-conversion scheme

Fig. 3 Software defined receiver architecture.

A little realistic scheme is combination of the tunable RF frontend with the bandpass ADC shown in Fig. 3(b)[8,26]. In this configuration, the band selection is done by RF analog filtering, and remaining channel selection is carried out digitally. Due to under-sampling operation of the ADC, f_{max} in Eq. (1) is given by the bandwidth of the RF filter. In this case, extremely high out-of-band rejection is requested to the RF BPF because the under-sampling does not offer the frequency selectivity and jammer signals must be suppressed to the noise level by the RF BPF alone.

More realistic scheme is combination of the down-conversion with the sampling mixer[27] shown in Fig. 3(c) [5,6,9,10]. Output current of the transconductance LNA is sampled by a CMOS switch and accumulated to a capacitor. This action is equivalent to the boxcar integrator or the infinite impulse response (IIR) filter, and thus the down-conversion, sampling and LPF functions are simultaneously achieved. High-speed (~80 MHz sampling rate) and high-resolution (~14bits) are requested to the ADCs to avoid disturbances caused by strong in-band interferers[6]. Such devices are commercially available, but still expensive and power hungry.

Another scheme is the same as the one shown in Fig. 1 but tunablity is given to the BB-LPF for the channel selection [5,6].

Fig. 4 shows an extension of the sampling mixer called the charge domain filter[5,6,28,29]. Current output of a trans-conductance amplifier is sampled by following N capacitors C_r successively, and summed charges are re-sampled (decimation) and detected by the ADC. This action is equivalent to the N-tap finite impulse response (FIR) filter.

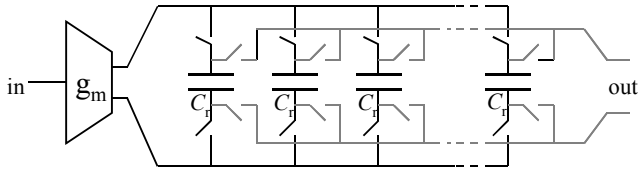


Fig. 4 Charge domain filter

Theoretically, sharp cut-off characteristics can be realized by cascading multiple sections[5,6]. Furthermore, in addition to the CMOS compatibility, the circuit offers adjustability of the filter bandwidth using the sampling clock frequencies. From these features, it seems that RF-BPF can be eliminated provided this topology offers sufficient out-of-band rejection. This is the perfect solution as the analog front-end of the SDR. However, extremely high linearity is required to the LNA in this case because it will be directly exposed to jammer signals[5,6]. In FDD transceivers, lack of DPXs will request the IIP2 higher than +120 dBm.

Thus majority of reported SDRs still equip series of high-performance fixed frequency RF-BPFs based on SAW/BAW technologies, and choose appropriate one(s) for the band selection[4,5,7,30,31].

In current mobile phones, most of all functionalities except RF-BPFs including DPXs and PAs have already been integrated into a single CMOS chip. In addition, use of the fixed RF filters is not flexible enough to introduce new communication standards. Thus such SDRs do not offer practical merit for the use in mobile phones.

IV. Tunable RF Filters

It is clear that tunable RF frontend, especially tunable RF-BPFs, can be an invaluable tool for realization of fully flexible SDRs.

Combination of RF BAW filters with RF micro-electromechanical system (MEMS) switches gives a possibility to integrate them with CMOS circuitry on a single Si chip[32,33]. It is possible, but not practical. This is because the AlN thickness control is too tough to produce multiple RF BAW filters with different design simultaneously.

Use of RF MEMS filters using in-plane vibration can relax the requirement for the AlN thickness control[29,32]. However, achievable filter bandwidth is narrow due to small electromechanical coupling for the vibration mode. It should be noted that small dielectric constant of AlN forces the device area large in order to obtain small input impedance for impedance matching with peripherals such as the antenna.

Various types of continuously tunable filters have been also proposed and developed. Magnetically tunable filters using the ferromagnetic resonance, such as yttrium iron garnet (YIG) spheres and films[34], offer low insertion loss and

multi-octave tuning range, and are widely used in tele-communication systems[35]. However size, weight, price and current consumption do not fit to the use in handy terminals. Intrinsic non-linearity is also a problem.

Mechanically tunable filters are also commercially available. But they are too large, too heavy and too slow for the current purpose[36].

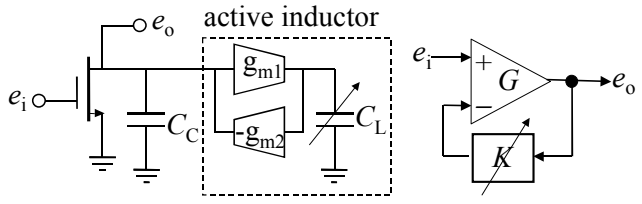
Certain number of attempts was reported on electrically tunable filters and some of them have already been mass produced[36]. Basic idea is to combine inductive circuit elements with VCs[36-38]. When lumped-element inductors are used, achievable Q is limited. On the other hand, when micro-strip lines are used, achievable Q is better, but the device size is large in a few GHz range.

As VCs, various types have been used. The GaAs VCs has been used because of its large tuning range[36]. However its Q is not so high and nonlinearity is large. VCs based on MEMS technologies[39,40] offer high Q while their physical size is relatively large and tuning range is relatively narrow. Currently, VCs employing ferroelectric ceramics such as bismuth strontium titanate (BST) are paid attention for their wide tuning range and relatively high Q[41].

Combination of CMOS (or MEMS) switches with fixed capacitor arrays is an interesting approach[42,43]. This offers large tuning range and high Q simultaneously. In addition, the total capacitance can be set digitally, and the device can be integrated with driving and control circuitry.

DRs offers extremely high Q but bulky in a few GHz range. A widely-tunable BST-based DR filter was reported, but it operates in 30 GHz range[44].

RF active filters have been also investigated extensively. There are two types. First one composes a resonator circuit using the gyrator-based active inductor[45] (see Fig. 5(a)). It causes the parallel resonance with C_C, and the voltage gain e_o/e_i becomes large. The resonance frequency and Q can be tuned by C_L and the trans-conductance g_m, respectively. Another one employs the positive feedback loop and acts as the IIR filter[46,47] (see Fig. 5(b)). The resonance frequency and Q can be tuned by adjusting phase shift in K and the gain G of the amplifier, respectively. These devices can be implemented into RF-ICs. In addition, they offer wide tuning range and high Q. In fact, Q can be infinite, and the situation corresponds to the self-oscillation. Thus we must apply fine gain setting to achieve high Q. Another problem is nonlinearity. Since the total gain becomes large when Q is high, the input-referred intercept points will be small due to output saturation.



(a) Active inductor-based filter (b) Recursive filter

Fig. 5 RF active filters embedded in LNA

Several types of tunable RF SAW/BAW filters were also proposed. Use of ferromagnetic, ferroelectric, and intrinsic and induced piezoelectricity has been discussed extensively to adjust acoustic wave velocities[48,49]. However, reported tuning ranges are limited to less than a few percents. Fancy one is to place an acoustic load above the RF BAW resonator using a movable hinge of the MEMS switch[17]. Although change of the center frequency more than 10% is expected, experimental results were not reported until now.

The author's group proposed another approach, which employ combination of VCs with extremely wideband SAW resonators[22]. Fig. 6 shows its topology. Basic design is exactly the same with the conventional ladder-type filter, and VCs connected in parallel with SAW resonators shift their anti-resonance frequencies while VCs connected in series shift their resonance frequencies. Thus proper VC setting enables us to adjust location of upper and lower edges of the filter passband independently.

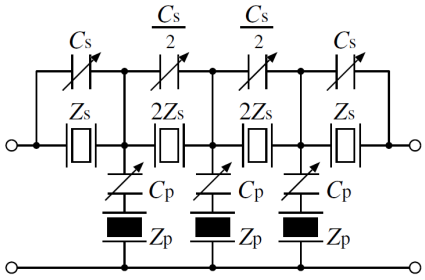


Fig. 6 Tunable filter configuration employing acoustic wave resonators and variable capacitors

We fabricated two SAW filters with identical resonator design but different capacitance values, and action of the tunable filter was simulated experimentally[22]. For achieving wide tunable range, SAW resonators on a Cu-grating/15°YX-$LiNbO_3$ (15-LN) substrate structure[50] were employed. Shear-horizontal (SH) SAWs on this structure possess extremely large k^2 (~30%).

Figure 7 shows the measured frequency responses of these two filters. It is seen that the passband can be translated. An

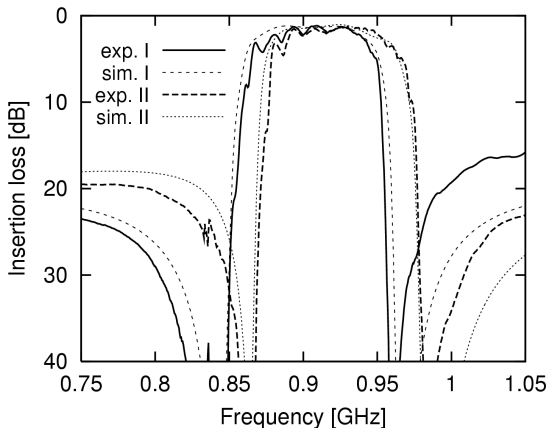


Fig. 7 Frequency response of a variable filter composed of wideband SAW resonators.

interesting feature is the fact that the passband characteristic is insensitive to the Q factor of VCs. A drawback of this design is limited tunability due to existence of a certain frequency range which is always included in the filter passband.

We also showed that passband width and location can be adjusted more flexibly by using the topology shown Fig. 8 [22]. One drawback of this configuration is sensitivity of filter performances to the Q factor of VCs. Namely that of C_s^s and C_p^p in Fig. 8 influences the filter insertion loss and steepness of the passband edges significantly.

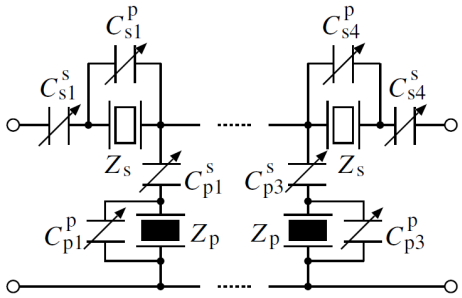


Fig. 8 Another tunable filter configuration employing acoustic wave resonators and VCs.

It is clear that wide tunability is realizable when we adopt variable inductors and/or fixed inductor with switch arrays in this configuration[51]. However, this approach seems not feasible because the Q factor of conventional inductors is pessimistically low in GHz range.

Tunable frequency range of these filters is inherently limited by the capacitance ratio γ of employed RF SAW/BAW resonators, and thus use of piezoelectric materials with extremely large k^2 is essential for its realization.

An interesting attempt for high k^2 realization is use of the piezoelectric layer prepared by polishing a ferroelectric single crystal plate[52,53]. This approach allows us to use arbitrary materials with arbitrary crystal orientation. Presently, the plate thickness can be reduced close to 1 μm. When the technology is applied for fabrication of RF BAW and Lamb wave resonators, small γ and moderate Q factor are achievable simultaneously in GHz range using commercially available high-quality and large-size $LiNbO_3$ wafers.

Use of highly piezoelectric materials such as PMN-PT[54], PIM-PMN-PT[55] and $KNbO_3$[56] is also a possible approach. For example, Matsushima, et al reported development of an FBAR with high k^2 (~0.5) and relatively high anti-resonance Q (~150) using a sputter-deposited $Pb(Mn,Nb)O_3$-$Pb(Zr,Ti)O_3$ film[57].

V. Integration of MEMS and SAW Filter

For realization of practical tunable SAW/BAW filters, multiple VCs and control electronic circuitry must be interconnected in high density with RF SAW/BAW resonators, which are very sensitive to parasitic impedance. Because use of highly piezoelectric ferroelectric materials is mandatory, the chip-on-chip integration seems to be most feasible and pragmatic.

There are two possible configurations: (a) VCs are fabricated at first on a piezoelectric substrate with SAW electrode patterns, and then the flip-chip interconnection is done with a Si chip where RF and control circuits are made [58], and (b) VCs and RF BAW resonators are fabricated on a bare Si substrate at first, and then it is flip-chip bonded with another Si chip with RF and control circuits.

As a first trial for the configuration (a), integration of MEMS VCs on $LiNbO_3$ was investigated[59]. Fig. 9 shows top and side views of the device structure, which consists of 5 SAW resonators, 2 VCs and one fixed capacitor. Fig. 10 shows the fabrication process. First, SAW resonators and fixed capacitors are fabricated using Au electrodes (110 nm) on a 15°YX-$LiNbO_3$ wafer by the lift-off process (Steps 1 and 2). The lift-off process is again applied for the fabrication of Au stopper posts, on which the capacitor bridge lands when pulled in (Step 3). Next, a sacrificial layer is formed using

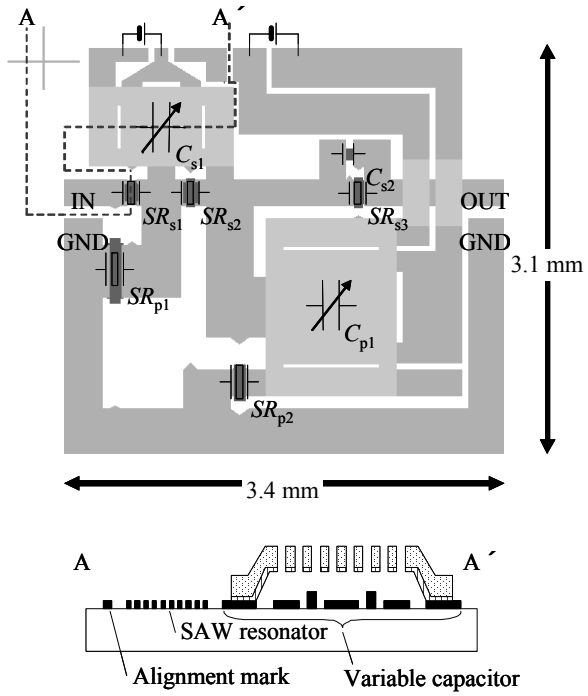


Fig. 9 Device layout and cross sectional views of tunable filter.

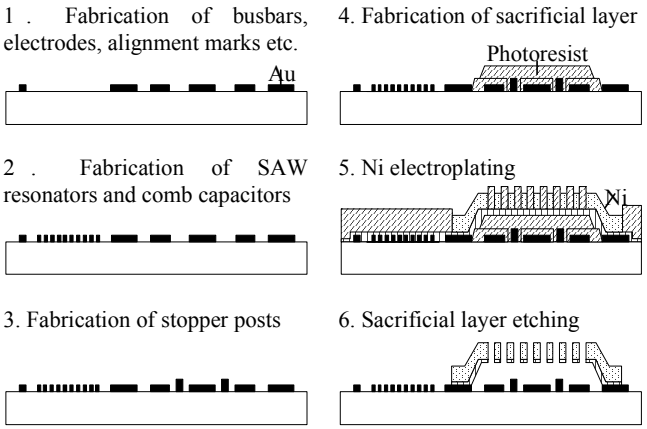


Fig. 10 Fabrication process for the tunable SAW filter.

negative photoresist, and a Cu seed layer is sputter-deposited (Step 4). To make the capacitor bridge, a positive photoresist is patterned for a mold, and 2 μm thick Ni is electroplated in the mold (Step 5). The photoresist mold is removed by organic stripper, and the Cu seed layer is removed by ammonium persulfate. Finally, the double sacrificial photoresist layer is removed by O_2 ashing to release the capacitor bridge (Step 6).

Figure 11 shows the measured and theoretical transmission characteristics of the filter. Although relatively wideband width of about 13% was achieved both theoretically and experimentally, experimental insertion loss (12 dB) was much larger than the theory. This deterioration is mainly due to surface contamination by the photoresist residue, which could not be completely removed by O_2 ashing at the Step 6.

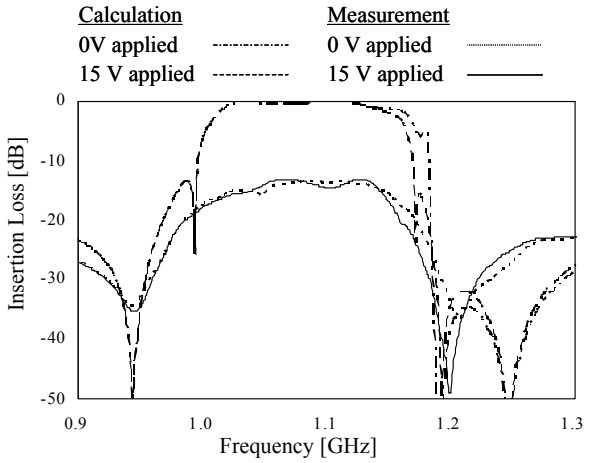


Fig. 11 Measurement and Simulation

For the passband tuning, DC voltage was applied to both C_{s1} and C_{s2} (see Fig. 9a). The application of 15 V made C_{s1} pulled in, and -3 dB bandwidth was reduced from 146 MHz to 130 MHz, as shown in the figure. On the other hand, C_{p1} could not be actuated even when 100 V was applied. This may be due to the buckling of the capacitor bridge, which might be caused by the stress of electroplated Ni.

It is seen from Fig. 9 that the MEMS VCs are much larger than the SAW resonators. This is because vacuum is used as the dielectrics for the MEMS VCs. For the size reduction, we also investigate on-chip integration of BST VCs on the SAW substrate. The result will be presented in future.

We also need to develop driving and control circuitry for the VCs on a Si chip, which will be interconnected with the SAW substrate using the flip-chip technology. These research topics are also involved in this project.

VI. CONCLUSIONS

This paper reviews current status and future prospects of the project aimed at developing tunable RF SAW/BAW filters.

Although we only discussed use of RF SAW devices in this paper, we also investigate integration of RF BAW resonators and VCs on a Si chip.

Applicability of the on-chip technologies is not limited to tunable filters. For example, we are also interested in RF oscillators and amplifiers.

Although the tunable filter under concern is expected to offer low loss and wide tunability, its performances are still not enough for the replacement with current RF SAW/BAW DPXs. Especially, high isolation between their TX and RX ports is hard to realize.

Currently, various approaches are investigated extensively to relax the isolation requirement given to DPXs. Integration of an interstage notch-filter for the TX signal leakage[24] and its active cancellation[60] are their examples. Use of RF active filters described in Section IV seems also effective for the purpose. IMD cancellation using base-band signal processing is also investigated[61]. These technologies may make tunable DPXs more realistic.

Our goal is far beyond the horizon. But the strong demand is the enough reason to challenge this difficult task.

ACKNOWLEDGMENT

The authors thank members of the project and anonymous students in Chiba and Tohoku Universities for their supports and discussions. This research is supported by the Japan Society for the Promotion of Science (JSPS) through its Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program).

REFERENCES

[1] http://galaxys2.samsungmobile.com/html/specification.html

[2] D.Kaczman, M.Shah, M.Alam, M.Rachedine, D.Cashen, L.Han and A.Raghavan, "A Single–Chip 10-Band WCDMA/HSDPA 4-Band GSM/EDGE SAW-less CMOS Receiver With DigRF 3G Interface and +90 dBm IIP2," IEEE J. of Solid State Circuits, **44**, 3 (2009) pp. 718-739

[3] J.Mitola III, "Software Radios - Survey, Critical Evaluation and Future Directions", IEEE Telesystems Conf. (1992) pp. 13/15-23.

[4] H.Harada, "Software Defined Radio Prototype toward Cognitive Radio Communication Systems", Proc. First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (2005) pp. 539-547.

[5] R.Bagheri, A.Mirzaei, S.Chehrazi, M.Heidari, M.Lee, M.Mikhemear, M.Tang and A.Abidi, "An 800MHz-to-5GHz Software-Defined Radio Receiver in 90nm CMOS", Dig. Tech. Papers, International Solid State Conf. (2006) pp. 1932-1941.

[6] A.A.Abidi, "The Path to the Software-Defined Radio Receiver," IEEE J. Solid-State Circuits, **42**, 5 (2007) pp. 954-966.

[7] H.Harada, "A Feasibility Study on Software Defined Cognitive Radio Equipment," Proc. Third IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (2008) pp. 1-12.

[8] J.F.Lu, T.Mueller, T.Mack and A.Terzis, "Configurable RF Receiver Architecture", IEEE Microwave Magazine, **4**, 2 (2004) pp. 75-82

[9] L.Maurer, T.Burger, T.Dellsperger, R.Stuhlberger, G.Hueber, M.Schmidt, and R.Weigel, "On the Architectural Design of Frequency-Agile Multi-Standard Wireless Receivers," Frequenz, Proc. IST Mobile and Wireless Summit (2005)

[10] L.Maurer, R.Stuhlberger, C.Wicpalek, G.Haberpeuntner and G.Hueberet, "Be Flexible -- Highly flexible digital front–end

enhanced CMOS–based RF transceivers", IEEE Microwave Magazine, **9**, 2 (2008) pp. 83-96.

[11] R.Weigel, D.P.Morgan, J.M.Owens, A.Ballato, K.M.Lakin, K.Hashimoto and C.C.W.Ruppel, "Microwave Acoustic Materials, Devices and Applications", IEEE Trans. Microwave Theory & Tech., **50**, 3 (2002) pp. 738-749.

[12] K.Hashimoto, M.Yamaguchi and Y.Satoh, "RF Filters Based on Surface-Acoustic-Wave and Film-Bulk-Acoustic-Resonator Technologies", Proc. of the European Microwave Association, **1**, 1 (2005) pp. 38-44.

[13] R.C.Ruby, P.Bradley, Y.Oshmyansky, A.Chien, and J.D.Larson III: "Thin Film Bulk Acoustic Resonators (FBAR) for Wireless Applications", Proc. IEEE Ultrason. Symp. (2001) pp. 813-821.

[14] T.Nishihara, M.Iwaki, G.Endo, X.Mi, S.Taniguchi, M.Ueda, and Y.Satoh, "BAW/SAW/IPD Hybrid Type Duplexer with Rx Balanced Output for WCDMA Band I", Technical Digest, IEEE Microwave Symp. (2008) pp. 831-834

[15] M.Kadota, T.Nakao, K.Nishiyama, S.Kido, M.Kato, R.Omote, H.Yonekura, N.Takada and R.Kita, "Small Surface Acoustic Wave Duplexer for Wide-Band Code-Division Multiple Access Full-Band System Having Good Temperature Characteristics", Jpn. J. Appl. Phys., **46**, 5 (2007) pp. 4714-4717

[16] C.Zinck, E.Defay, A.Volatier, G.Caruyer, D.P.Tanon, and L.Figuiere, "Design, Integration and Characterization of PZT tunable FBAR," Proc. International Symp. on Applications of Ferroelectrics (2004) pp. 29-32.

[17] M.El Hassan, E.Kerherve, Y.Deval, A.A.Shirakawa, P.Jarry, and A.Cathelin, "A Study of FBAR Filters Reconfiguration", Proc. IEEE International Conf. on Electronics, Circuits and Systems (2005) pp. 1-4.

[18] S.Gevorgiana and A.Vorobiev, "DC Field and Temperature Dependent Acoustic Resonances in Parallel-Plate Capacitors Based on $SrTiO_3$ and $Ba_{0.25}Sr_{0.75}TiO_3$ Films: Experiment and Modeling," J. Appl. Phys., **99** (2006) 124112-1~11

[19] X.Zhu, J.D.Phillips, and A.Mortazawi, "A DC Voltage Dependant Switchable Thin Film Bulk Wave Acoustic Resonator Using Ferroelectric Thin Film," Tech. Digest of IEEE Microwave Conf. (2007) pp. 671-674.

[20] A.Vorobiev, J.Berge, M.Norling and S.Gevorgian, "Silicon Substrate Integration of BST Based Tunable TFBARs Using All-dielectric SiO_2/AlN Bragg Reflectors," Proc. Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (2010) pp. 41-44.

[21] A.Vorobiev and S.Gevorgian, "Tunable $Ba_xSr_{1-x}TiO_3$ FBARs Based on $Si0_2$/W Bragg Reflectors," Tech. Digest of IEEE Microwave Conf. (2010) pp. 1444-1447.

[22] T.Komatsu, K.Hashimoto, T.Omori and M.Yamaguchi, "Tunable Radio Frequency Filters Using Acoustic Wave Resonators and Variable Capacitors," Jpn. J. Appl. Phys., **49**, 7 (2010) 07HD24-1~4.

[23] S.Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications", IEEE J. Selected Areas in Communications, **23**, 2 (2005) pp. 201-220.

[24] R.Vazny, W.Schelmbauer, H.Pretl, S.Herzinger, and R.Weigel, "An Interstage Filter-Free Mobile Radio Receiver with Integrated TX Leakage Filtering," Technical Digest, IEEE Radio Frequency Integrated Circuit Symp. (2010) pp. 21-24

[25] W.Kester, "Aperture Time, Aperture Jitter, Aperture Delay Time— Removing the Confusion," MT-007 Tutorial, Analog Devices, Inc. (2008)

[26] K.Koli, S.Kallioinen, J.Jussila, P.Sivonen, and A.Parssinen, "A 900-MHz Direct Delta-Sigma Receiver in 65-nm CMOS," IEEE J. Solid-State Circuits, **45**, 10 (2010) pp. 2807-2818.

[27] J.Yuan, "A Charge Sampling Mixer with Embedded Filter Function for Wireless Applications, " Proc. 2nd Int. Conf. Microwave and Milimeter Wave Tech. (2000) pp. 315-318.

[28] K.Muhammad and R.B.Staszewski, "Direct RF Sampling Mixer with Recursive Filtering in Charge Domain," Proc. of ISCAS, Vol. I (2004) pp. 577-580

[29] R.B.Staszewski, K.Muhammad, D.Leipold, C.M.Hung, Y.C.Ho, J.L.Wallberg, C.Fernando, K.Maggio, R.Staszewski, T.Jung, J.Koh, S.John, I.Y.Deng, V.Sarda, O.Moreira-Tamayo, V.Mayega, R.Katz, O.Friedman, O.E.Eliezer, E.de-Obaldia, and P.T.Balsara, "All-Digital TX Frequency Synthesizer and Discrete-Time Receiver for Bluetooth Radio in 130-nm CMOS," IEEE J. Solid-State Circuits, **39**, 12 (2004) pp.2278-2291

[30] C.T.Y.Nguyen, "Integrated Micromechanical Circuits Fueled By Vibrating RF MEMS Technology", Proc. IEEE Ultrason. Symp. (2006) pp.953-962

[31] T.Matsumura, M.Esashi, H.Harada, F.Thalmayr, K.Hashimoto, and S.Tanaka, "Selective Mode Excitation of Piezoelectric Disk-type Resonator by Electrode Pattern Definition", Proc. IEEE Ultrason. Symp. (2010) [to be published]

[32] L.Elbrecht, R.Aigner, C.I.Lin and H.J.Timme, "Integration of Bulk Acoustic Wave Filters: Concepts and Trends", Technical Digest, IEEE Microwave Symp., Vol. 2 (2004) pp. 927 – 929

[33] M.A.Dubois, J.F.Carpertier, P.Vincent, C.Billard, G.Parat, C.Muller, P.Ancey and P.Corti, "Monolithic Above-IC Resonator Technology for Integrated Architectures in Mobile and Wireless Communication", IEEE J. Solid State Circuits, **41**, 1 (2006) pp. 7-16.

[34] A.Isobe and K.Asai, "Contour-Mode AlN Resonator with High Coupling Factor," Jpn. J.Appl. Phys. **49**, 7 (2010) 07GG07-1~3.

[35] J.D.Adam, L.E.Davis, G.F.Dionne, E.F.Shloemann and S.N.Stitzer, "Ferrite Devices and Materials", IEEE Trans. Microwave Theory and Tech., **50**, 3 (2002) pp. 721 – 737.

[36] J.Uher and W.J.R.Hoefer, "Tunable Microwave and Millimeter-Wave Band-Pass Filters," IEEE Trans. Microwave Theory and Tech., **39**, 4 (1991) pp. 643-653.

[37] G.M.Kraus, C.L.Goldsmith, C.D.Nordquist, C.W. Dyck, P.S. Finnegan, F.Austin IV, A.Muyshondt, and C.T.Sullivan, "A Widely Tunable RF MEMS End-Coupled Filter," Technical Digest, IEEE Microwave Conf., Vol.2 (2004) pp. 429-432.

[38] J.C.Estes, "Tunable RF Band Pass Filter with Variable Resonator Coupling", Technical Digest, International Microwave Symp. (2008) pp. 1035-1038.

[39] H.A.C.Tilmans, W.de Raedt and E.Beyne, "MEMS for Wireless Communications: from RF-MEMS Components to RF-MEMS-SiP', J. Micromechanics and Microengineering, **13** (2003) pp.139-163.

[40] H.J.deL.Santos, G.Fischer, H.A.C.Tilmans and J.T.M.van Beek, "RF MEMS for Ubiquitous Wireless Connectivity, Part II Applications", IEEE Microwave Magazine, **7**, 6 (2006) pp. 50-65

[41] B.York, "BST Filters and Voltage-Activated BAWs", Workshop on Recent Advances in Reconfigurable Filters, IEEE Microwave Symp. (2010)

[42] A.v.Bezooijen, M.A. de Jongh, C.Chanlo, L.C.H.Ruijs, F.v.Straten, R.Mahmoudi, and A.H.M.v.Roermund, "A GSM/EDGE/WCDMA Adaptive Series-LC Matching Network Using RF-MEMS Switches," IEEE J. Solid State Circuits, **43**, 10 (2008) pp. 2259-2268.

[43] A.Chamseddine, J.W.Haslett, and M.Okoniewski, "CMOS Silicon-on-Sapphire RF Tunable Matching Networks,"

EURASIP J. on Wireless Communications and Networking, 2006 (2006) pp. 1-11

[44] S.Courrèges, Z.Zhao, K.Choi, A.Hunt and J.Papapolymerou, "Ferroelectric Tunable Bandpass Filters for X and Ka-Band Applications," Workshop on Recent Advances in Reconfigurable Filters, IEEE Microwave Symp. (2010)

[45] H.Xaio, R.Schaumann, W.R.Daash, P.K.Wong, and B.Pejcinovic, "A Radio-Frequency CMOS Active Inductor and Its Application in Designing High-Q Filters", IEEE Symp. on Circuit and systems (2004) pp. 197-200.

[46] S.Darfeuille, Z.Sassi, B.Barelaud, L.Billonnet, B.Jarry, H.Marie, N.T.L.Le and P.Gamandet, "A Fully Differential 2 GHz Tunable Recursive Bandpass Filter on Silicon", Germany Microwave Conf. (2005) pp.102-105

[47] S.Andersson and C.Svensson, "An Active Recursive RF filter in 0.35um BiCMOS", J. Analog Integrated Circuits and Signal Processing, **44**, 3 (2005) pp. 213-218.

[48] P.Smole, W.Ruile, C.Korden, A.Ludwig, E.Quandt, S.Krassnitzer, and P.Pongratz, "Magnetically Tunable SAW Resonator," Proc. IEEE International Freq. Contr. Symp. (2003) pp. 903-906

[49] J.H.Kuypers, M.Schmidt, S.Tanaka, and M.Esashi, "Phase Velocity Control of Surface Acoustic Waves Based on Surface Shorting and Electrical Field Application using MEMS Switches," Proc. IEEE Ultrasonics Symp. (2007) pp. 1233–1238.

[50] T.Omori, K.Matsuda, N.Yokoyama, K.Hashimoto and M.Yamaguchi, "Suppression of Transverse Mode Responses in Ultra-Wideband SAW Resonators Fabricated on a Cu-grating/15^{o}YX-$LiNbO_3$ Structure", IEEE Trans. Ultrason., Ferroelec., and Freq. Contr., **54**, 10 (2007) pp. 1943-1948.

[51] N.O.Fenzi, P.J.Turner, B.A.Willemsen, J.R.Costa, E.R.Soares, and S.Jimenez, "Multimode bandpass SAW filter using Reconfigurable Resonance Technology," Proc. IEEE Ultrason. Symp. (2010) [to be published]

[52] Y.Osugi, T.Yoshino, K.Suzuki, and T.Hirai, "Single Crystal FBAR with $LiNbO_3$ and $LiTaO_3$ Piezoelectric Substance Layers," Technical Digest, IEEE International Microwave Symposium (2007) pp.873-876.

[53] M.Kadota and T.Ogami, "5.4 GHz Lamb Wave Resonator on $LiNbO_3$ Thin Crystal Plate and its Application," Jpn. J. Appl. Phys., **50**, 7 (2011) [to be published]

[54] J.Chen, H.Wang, K.Hashimoto, T.Omori, and C.J.Ahn, "Suppression of Rayleigh Wave Spurious Signal in Shear Horizontal Surface Acoustic Wave Devices Employing (1-*x*)$Pb(Mg_{1/3}Nb_{2/3})O_3$-*x*$PbTiO_3$ Crystals," [under review]

[55] J.Chen, H.Wang, K.Hashimoto, T.Omori, C.J.Ahn, X.Y.Zhao, and H.S.Luo, "Theoretical Analysis of Ultrahigh Electromechanical Coupling Surface Acoustic Wave Propagation in PIN-PMN-PT Crystals," J. Appl. Phys., **109** (2011) 054104-1~3.

[56] T.B. Pollard, T.D.Kenny, J.F. Vetelino, and M.P. da Cunha, "Pure SH-SAW Propagation, Transduction and Measurements on $KNbO_3$," IEEE Trans. on Ultrason., and Freq. Contr. **53**, 1 (2006) pp. 199-208.

[57] T.Matsushima, N.Yamaguchi, T.Shirai, T.Yoshihara, Y.Hayashi, T.Ueda, I.Kanno, K.Wasa and H.Kotera, "High Performance 4 GHz FBAR Prepared by $Pb(Mn,Nb)O_3$-$Pb(Zr,Ti)O_3$ Sputtered Thin Films," Proc. IEEE Freq. Contr. Symp. (2010) pp. 248-251.

[58] K.D.Park, M.Esashi and S.Tanaka, "Wafer-level Hetero-integration Process of SAW Devices and LSI," Proc. IEEE Ultrason, Symp. (2010) [to be published]

[59] T.Yasue, T.Komatsu, N.Nakamura, K.Hashimoto, M.Esashi and S.Tanaka, "Wideband Tunable Love Wave Filter Using Electrostatically-Actuated MEMS Variable Capacitors Integrated on Lithium Niobate," Proc. Transducers 2011 (2011) [to be published].

[60] V.Aparin, G.J.Ballantyne, C.J.Persico, and A.Cicalini, "An Integrated LMS Adaptive Filter of TX Leakage for CDMA Receiver Front Ends," IEEE J. Solid-State Circuits, **41**, 5 (2006) pp. 1171-1182.

[61] M.Valkama, A.S.H.shahed, L.Anttila, and M.Renfors, "Advanced Digital Signal Processing Techniques for Compensation of Nonlinear Distortion in Wideband Multicarrier Radio Receivers," IEEE Trans., Microwave Theory and Tech., **42**, 6 (2006) pp. 2356-2366.

Cascaded Channel-Select Filter Array Architecture Using High-K Transducers for Spectrum Analysis

Eugene Hwang, Tanay A. Gosavi, Sunil A. Bhave
School of Electrical and Computer Engineering
Cornell University
Ithaca, NY, USA

Ronald G. Polcawich, Jeffrey S. Pulskamp, Sarah Bedair
US Army Research Laboratory
Adelphi, MD, USA

***Abstract*—This work presents a novel cascaded filter array architecture for large-scale MEMS filter arrays that overcomes the limitations associated with the large shunt capacitance to efficiently scale the number of filters in the array. The proposed architecture effectively reduces the capacitive loading at the input by utilizing a series connection of the filters, resulting in improvement in the return loss and stopband rejection for large arrays. This work demonstrates the feasibility of large-scale filter arrays without the need for additional costly isolation strategies for real-time spectrum monitoring applications.**

I. Introduction

On-going research and development of high quality factor (Q) RF MEMS resonators and filters has allowed their use in a number of RF applications, such as timing references and front end band-select filters [1-4]. Many of these devices have been designed to fit the specifications of current communication system architectures as direct replacement components for quartz or SAW devices with notable success. Currently, most radio architectures use the concept of band-select filtering followed by heterodyning to a lower intermediate frequency, where further signal processing and channel-select filtering is performed. This architecture has been established as the standard for most radio front ends since its invention by Armstrong [5]. It, however, involves a number of tradeoffs between various performance metrics, including linearity, power consumption, and noise, which have been unavoidable largely due to the lack of extremely narrowband filtering capability at RF frequencies using conventional resonator technologies. RF MEMS resonators may be able to relax some of these tradeoffs due to their ability to realize extremely high Q exceeding 10,000 on chip, providing the potential for extremely narrowband filters even at VHF and UHF frequencies. Furthermore, this capability may even enable new types of communication standards and radio architectures for specific applications. One such application is cognitive radio, systems that are able to scan the spectrum and setup a communication link on relatively unused channels or identify and filter blocker signals before they saturate other sensitive electronics [6,7]. Large scale RF MEMS filter arrays have emerged as a key technology which may enable cognitive radio for mobile applications due to the unique ability of this technology to provide both channel-select filtering and lithographical definition of resonant frequency on an integrated platform. However, the number of filters necessary for such applications is on the order of hundreds to thousands.

Recent work has demonstrated arrays each with two and ten separate filters [8,9]. In both cases, the filters were connected in the conventional parallel configuration that is normally envisioned for such an array. While these works demonstrated arrays with multiple filters, they left unanswered questions about how the performance of such arrays would scale as the number of filters is increased. This is a problem due to the shunt capacitance present at the input of any RF MEMS resonator or filter, regardless of the transducer used. As the size of the array increases, the input loading from these shunt capacitors will increase in proportion to the number of filters in the array, significantly degrading performance. The first example [8] shown above was able to circumvent these negative effects by using a differential architecture and a

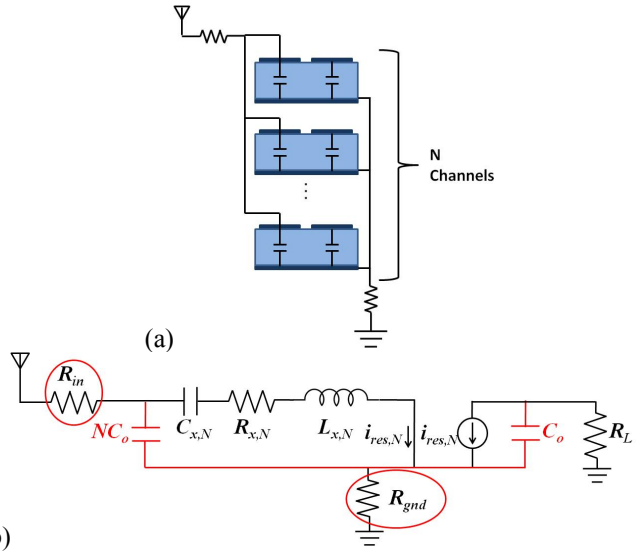


Figure 1. (a) Simplified schematic of the conventional parallel filter array architecture. (b) Equivalent circuit model used for analysis with feedthrough path highlighted in red and important parasitic resistances circled.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

parallel inductor to resonate with the shunt capacitance. The utility of such methods, however, are severely limited when it comes to scaling to hundreds of filters. On the other hand, the second example [9] used an AlN transducer, which is known to be a very efficient transducer, and thus is relatively unaffected by the shunt capacitance at such low frequencies and small array sizes. However, as the filter array increases in size and is designed to operate at higher frequency, it will also suffer from the input loading presented by the shunt capacitance. Others have suggested alternative methods of alleviating this problem by added isolation using integrated switches [10] but mechanical switches have yet to be successfully integrated into radio systems or cellular phones, making their use difficult in large scale arrays.

While the problem of shunt capacitance has been raised as an issue potentially limiting the use of MEMS filter arrays, to the authors' knowledge, a rigorous analysis of its effects on filter array performance for arbitrarily large array sizes has not yet been reported. In this paper, we present an analysis based on Butterworth-Van Dyke (BVD) equivalent circuit models with relevant parasitic elements and find that these parasitics combined with the large shunt capacitance severely limit filter performance when using a simple parallel array architecture. We then propose a novel series cascaded filter array architecture that provides a means of alleviating the increasingly negative effects of the shunt capacitance as the array increases in size. Experimental results will then be presented to verify our analysis and we will conclude with a summary and implications of this work.

II. Conventional Parallel Architecture and Effects of Shunt Capacitance

While the MEMS community has been focusing much of its energy on trying to interface directly to a 50 Ω environment (i.e., obtaining motional resistance of 50 Ω), in many cases this has been achieved naively by increasing the transduction area, which also increases the shunt capacitance. In general though, unless the filter array interfaces directly to a 50 Ω distributed system (e.g., antenna), this is not a necessary requirement. In fact, for a given motional resistance, it is always possible in the ideal case (i.e., no parasitic shunt capacitance) to terminate the filter with an optimal impedance which will result in nearly ideal filter characteristics, as long as the Q of the resonators is sufficiently high (i.e., $Q_{res} >> Q_{filter}$). Systems terminated with impedances on the order of 1 kΩ or lower have been demonstrated with low insertion loss and suitable shape factor, indicating that this is not the fundamental limitation [9]. The main problem in the context of large scale filter arrays then is not the motional impedance but the loading of the shunt capacitance. Even the work in [9] achieves its well-behaved filter response only after using the methods previously discussed to mitigate the negative effects of the shunt capacitance. Some of the well-known effects include large insertion loss and passband return loss due to a large impedance mismatch with the source that increases with array size. There are, however, other seldom considered effects – particularly in the stopband response – that arise due to parasitic routing resistances that become increasingly important as array size increases. As a starting point for analyzing these effects, we use a model based on the illustration in Fig. 1(a). An equivalent circuit model is shown in Fig. 1(b). Here, N represents the number of channels in the filter array (i.e., array size), R_{in} models any parasitic interconnect losses, and R_{gnd} models imperfect grounding that results from interconnect losses in the ground plane. The input is loaded by the parallel combination of the shunt capacitances of N filters. We will discuss in this section specifically how a number of important filter metrics – return loss, stopband rejection, and insertion loss – are affected by the shunt capacitance loading and parasitic resistances.

Before discussing the simulation results, note that the following simulations are based on data collected from individual 20 MHz PZT-on-Si filters fabricated at the U.S. Army Research Laboratory. The transduction area of these devices is scaled to achieve 50 Ω motional impedance. These devices are used specifically because the extremely high dielectric constant of PZT makes the loading effects significant at relatively low frequencies and small values of N. This significantly reduces the burden on fabrication and measurement setup for experimental demonstration. This does not, however, restrict the generality of this analysis. While these loading effects in other strongly coupled transducers may become significant at higher frequencies or larger N, this problem cannot be avoided by simply choosing a different transducer, especially when the goal is to design a large scale

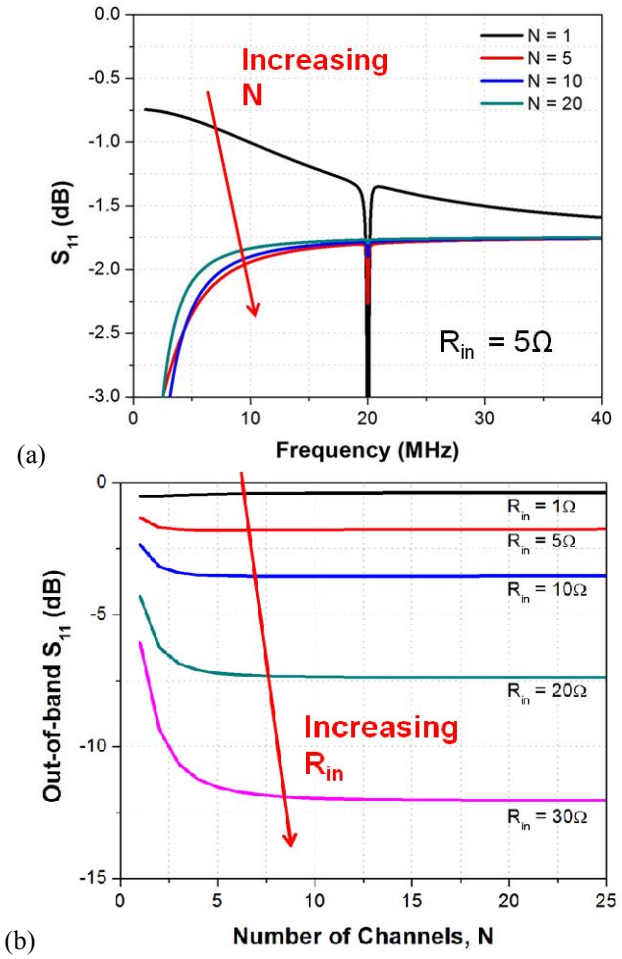


Figure 3. Return loss simulation results (a) as a function of frequency for R_{in} = 5Ω and varying N and (b) stopband return loss as a function of N at 20 MHz.

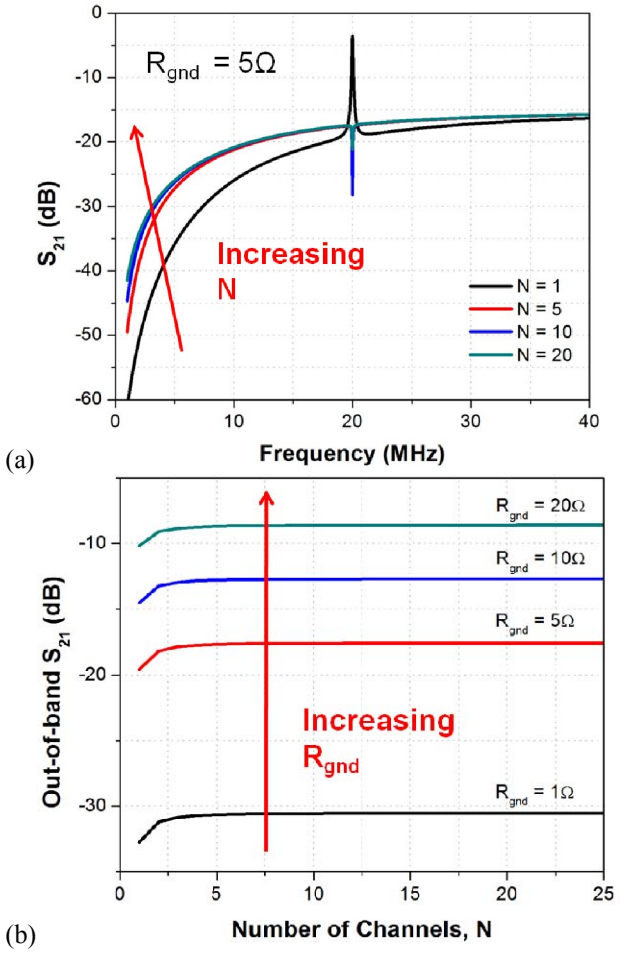


Figure 4. Transmission simulation results (a) versus frequency for R_{gnd} = 5Ω and varying N and (b) feedthrough signal as a function of N at 20 MHz.

filter array with hundreds or thousands of filters.

A. Return Loss

For filters, a large value of return loss in the stopband and a small value in the passband are typically desired. While the need for small passband return loss is well known, it is also important that the stopband return loss be large, a fact often overlooked by designers of MEMS filters. This requirement ensures that energy at these frequencies is not intercepted and wasted in the filter array – even -1 dB of return loss in the stopband means 20% of the incident power is absorbed. Also, nonlinearities may cause interference to appear in the filter passband if strong signals interact with one another. To see how this parameter is affected as the array size increases, we simulate the frequency response of the return loss for a fixed value of R_{in} = 5 Ω and various values of N (Fig. 3(a)). We see that as N increases, the stopband return loss converges to a certain value. This is due to the decreasing impedance of the effective shunt capacitance with increasing N, which virtually terminates the parasitic resistance R_{in} to ground. This provides a path for stopband energy to be absorbed by the filter array and dissipated in R_{in}, in contrast to the case when R_{in} = 0 where this energy was completely reflected. This is also illustrated in Fig. 3(b), which plots the stopband return loss as a function of N for different values of R_{in} at 20 MHz and shows that the stopband return loss converges to a value for large N that is highly dependent on R_{in}. Interestingly, Fig. 3(a) also indicates that the notch depth in the passband decreases as N increases. This indicates that as N increases, less passband energy is absorbed in the filter and any energy that is absorbed is dissipated in the parasitic resistances.

B. Stopband Rejection

Another common problem for RF MEMS filters is the feedthrough signal which limits the stopband rejection. This is often the result of imperfect grounding of the device. The path though which this signal may flow is highlighted in red in Fig. 1(b) and includes the shunt capacitance. Thus, we expect that this feedthrough signal will also be affected as the filter array increases in size. The simulated transmission results for R_{gnd} = 5 Ω and different values of N shown in Fig. 4(a) confirm that the feedthrough signal increases with N, and for certain values of R_{gnd} and N, the feedthrough may even exceed the insertion loss, eliminating the filter response. As in the case for the stopband return loss, the feedthrough converges to a value which is highly dependent on the parasitic resistance R_{gnd}, as shown in Fig. 4(b). This, again, is due to the decreasing impedance of the effective shunt capacitance as N increases, making the feedthrough signal a very strong function of R_{gnd}.

C. Insertion Loss

As previously mentioned, the main reason why insertion loss increases with N is because of impedance mismatch – energy is reflected away rather than flowing into the filters in the array. When these parasitic resistances are now added to the analysis, it is evident that energy dissipation through them – which was also found to impact stopband return loss – is an additional mechanism that prevents energy from reaching the filters. As the previous simulation results show, this dissipation increases with N for given values of the parasitic resistances R_{in} and R_{gnd}, and as already discussed, beyond a threshold value for N depending on R_{in} and R_{gnd} the filter response is practically lost. Again, this behavior is highly sensitive to these parasitic components for large N, setting a limit to the achievable array size. Of course, it is always desirable to minimize these parasitics through good fabrication. However, there may be certain limits to how low they can be made in practical systems. For increased robustness and insensitivity to process variations, it would therefore be desirable to also employ architectural methods of reducing the sensitivity of the filter array to these parasitic components. We see, however, that because of the large capacitive loading at the input, the parallel filter architecture does not provide any robustness to these parasitic resistances, especially as the size of the array increases.

III. Series Cascaded Filter Array Architecture

The analysis in the previous section indicates that the main issue limiting the performance of large scale MEMS filter arrays using the parallel architecture is the large effective shunt capacitance at the input. This combination of N shunt capacitances results in an impedance that decreases with

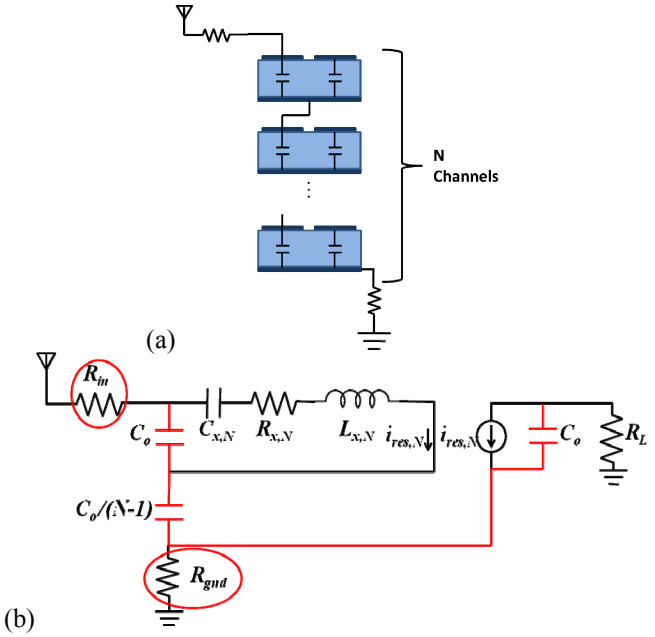


Figure 5. (a) Simplified schematic of proposed series cascaded filter array architecture. (b) Equivalent circuit model used for analysis with feedthrough path highlighted in red and important parasitic resistances circled.

increasing N, leading to a large feedthrough signal and allows the wasteful dissipation of energy through parasitic resistances. Furthermore, the filter response for large N becomes very sensitive to these parasitic resistances, which is undesirable. To avoid this problem, we propose an array architecture where the filters in the array are connected such that a series connection of their shunt capacitance is realized. This will be referred to as the series cascaded filter array architecture and Fig. 5 shows a simplified illustration of this architecture with its equivalent circuit model. For this architecture, since the filters are now connected in series, we expect the effective shunt capacitance to be inversely proportional to N, which results in an impedance that actually increases with array size. It is then expected that both the stopband return loss and feedthrough signal should improve as the array size increases.

We verify the improvement in stopband return loss, by simulating the return loss as a function of frequency for a constant value of R_{in} = 5 Ω and for different values of N (Fig. 6(a)). The results show that the stopband return loss approaches 0 dB as N increases, as expected. More significantly, Fig. 6(b) shows that this is true regardless of the value of R_{in}, as opposed to the parallel filter array architecture in which this parasitic resistance determines the value of return loss for large N. This indicates that the high impedance of the series cascaded array architecture effectively prevents stopband energy from being dissipated in R_{in} and allows it to be reflected back to the source.

Next, the transmission behavior of the series cascaded architecture is simulated as shown in Fig. 7 to verify the improvement in the stopband rejection and feedthrough signal. Fig. 7(a) shows the filter transmission as a function of frequency for R_{in} = 5 Ω and different values of N. This graph shows that as N increases, the feedthrough signal decreases as expected. Fig. 7(b) also confirms that the relationship is

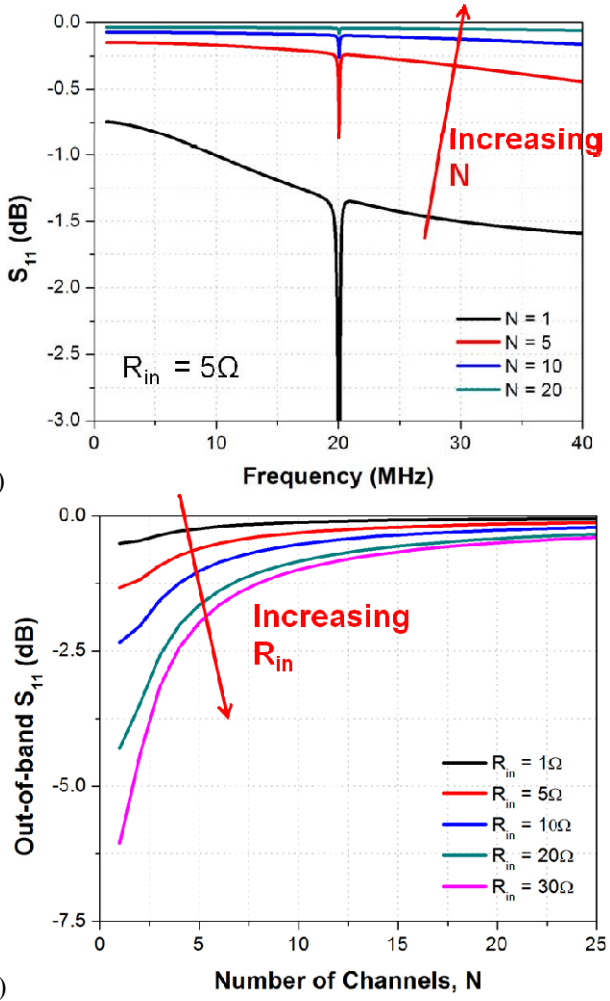


Figure 6. Return loss simulation results (a) as a function of frequency for R_{in} = 5Ω and varying N and (b) stopband return loss as a function of N at 20 MHz.

inversely linear with N, in contrast to the case of the parallel array architecture, where the feedthrough signal increases to some constant level for large N. Since the insertion loss also increases with N due to the voltage division occurring from the series connection of the N filters in the array, the result is a stopband rejection that is relatively insensitive to the array size and, more importantly, a transmission response that does not lose its filter behavior beyond a certain value of N. These simulation results all indicate that the series cascaded array architecture significantly improves the stopband return loss and filter transmission characteristics over its parallel counterpart for large arrays.

IV. Experimental Setup and Results

In order to demonstrate the performance enhancement that results from the use of the series cascaded filter array architecture, two 9-channel arrays of PZT-on-Si filters – one using the parallel architecture, the other using the series cascaded architecture – designed for center frequencies around 20 MHz with channel spacing of 30 kHz were fabricated in the process described in [11]. As previously mentioned, this technology was chosen due to the high dielectric constant of

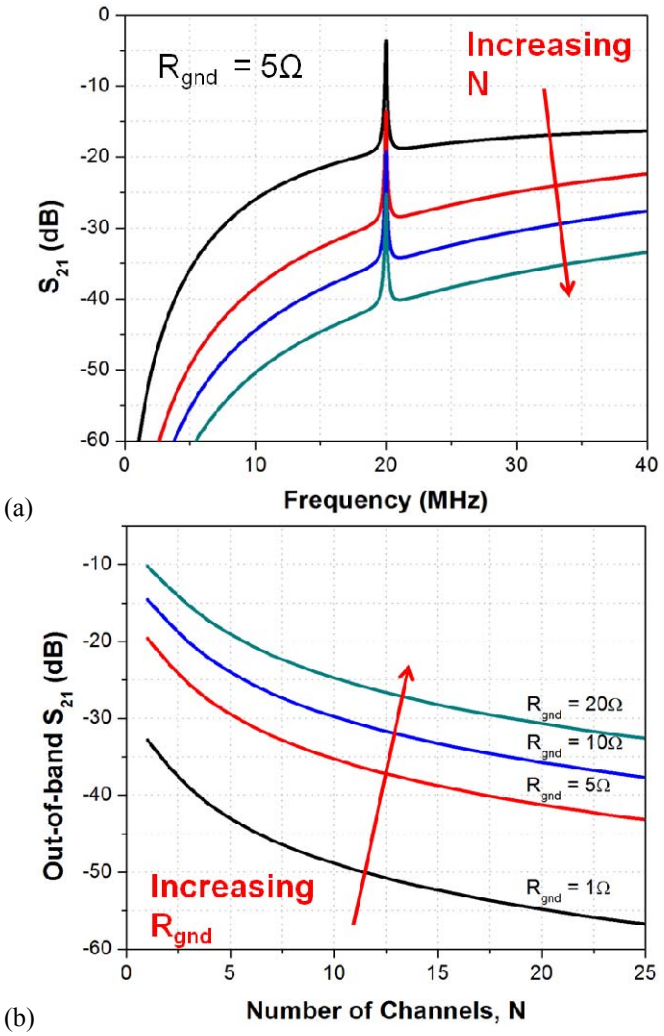


Figure 7. Transmission simulation results (a) versus frequency for R_{gnd} = 5Ω and varying N and (b) feedthrough signal as a function of N at 20 MHz.

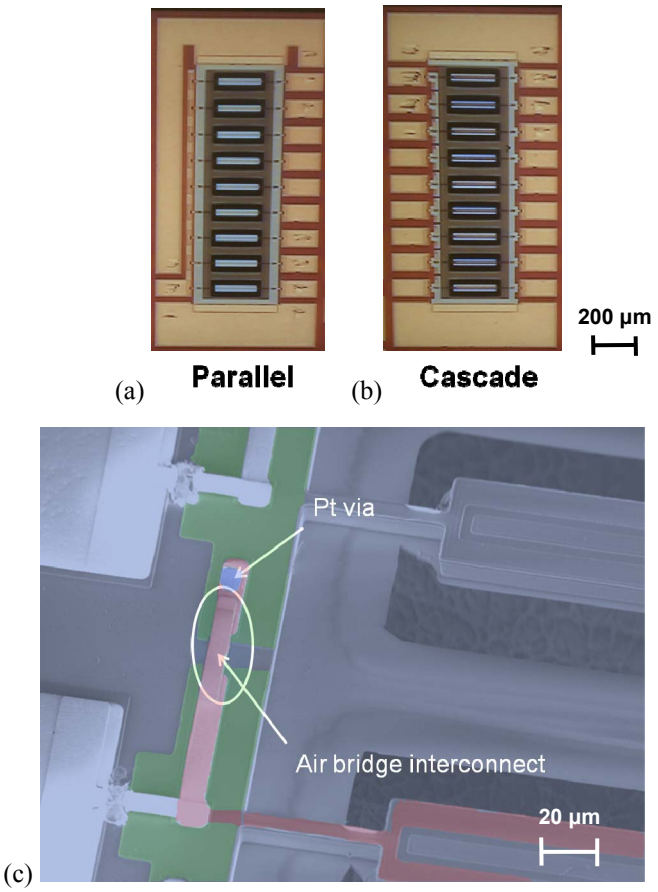


Figure 8. Chip micrograph of 9-channel filter array centered around 20 MHz implemented with (a) conventional parallel architecture and (b) series cascaded architecture using PZT-on-Si filter technology presented in [11]. (c) SEM image showing the routing between filters in the series cascaded architecture.

PZT enabling this work to demonstrate the capacitive loading effects at lower frequency using fewer filters, reducing the burden on fabrication and testing. Micrographs of these arrays are shown in Fig. 8(a) and (b). The SEM image in Fig. 8(c) shows how the series connections are implemented. The area shaded in green indicates the bottom electrode layer of the top filter underneath the PZT layer. A via, shaded blue, is opened to route to the gold air-bridge, which is then connected to the top electrode of the next filter, all of which is shaded in red. This connection is repeated for all filters except the last one in the chain, whose bottom electrode is connected to ground.

Fig. 9 shows the details of the two-port testing setup used to characterize these filter arrays. Port 1 of the vector network analyzer is connected to the GSG probe on the left and provides the input stimulus to the array. The outputs are probed using a 9-point probe structure whose connections are routed to a switching matrix. This matrix selects one filter output to be connected to port 2 of the network analyzer while all others are terminated with 50 Ω impedances. Typical two-port measurements are then repeated to characterize the response of each filter in the array.

Experimental results for the return loss and transmission are shown in Fig. 10(a) and (b), respectively. First, in Fig. 10(a), we see that the parallel filter array has a stopband return loss of roughly -3.5 dB while the notch depth is only about 0.1 dB, indicating that a majority of the absorbed energy is dissipated in parasitic resistances rather than passing through the filters. On the other hand, the series cascaded array shows a larger stopband return loss of roughly -0.7 dB with a notch depth of 0.4 dB, indicating a larger portion of passband energy is reaching the filters. In total, the stopband return loss is improved by roughly 2.9 dB.

Fig. 10(b) shows the results of transmission measurements again confirming the benefits of the series cascaded architecture. The parallel filter array has filters with only 7 dB stopband rejection compared with a roughly 23 dB rejection and 20 dB lower transmission floor for the series cascaded array. The insertion losses are also similar between the two arrays, which is expected. The series cascaded filter array has somewhat higher insertion loss, but this may be due to lower quality factors and non-zero routing resistance between filters. These have been noted and will be improved in future designs. In general, while these filter characteristics have room for improvement, they do indicate that the series cascaded filter array is favorable in terms of both return loss and transmission response of large scale filter arrays.

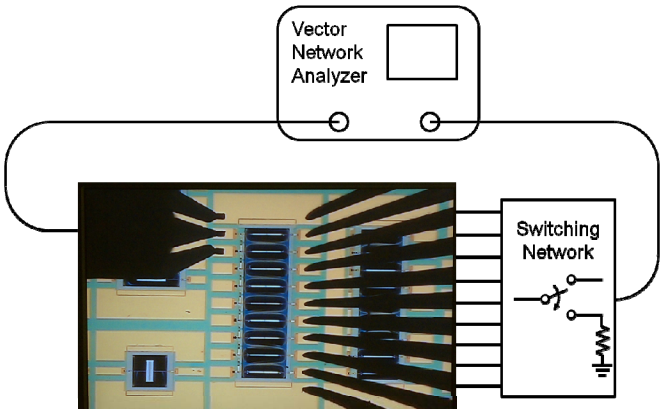


Fig. 9. Two-port measurement setup with switching network used for characterization of the two filter arrays.

V. Conclusions

In summary, this paper presents a detailed analysis of how the performance of the conventional parallel filter array architecture scales with array size using BVD equivalent circuit models. Through this analysis, the large effective shunt capacitance which increases with array size for the parallel array architecture was found to be the main factor limiting the performance of the array, specifically the stopband return loss, stopband rejection, and insertion loss. In addition, it was found that the filter performance was also extremely sensitive to parasitic routing resistances. While it is always desired to minimize these parasitic losses, some might be unavoidable, especially when targeting the design of very large scale filter arrays. Thus, any architectural changes that may reduce this sensitivity would be very beneficial. In response, the series cascaded filter array architecture was proposed to improve the scaling of filter performance with array size by making the effective shunt capacitance now inversely proportional to array size. Simulated results also indicate that this architecture reduces the sensitivity of the filter array to parasitics. This performance improvement was experimentally verified using 9-channel filter arrays with PZT-on-Si filters around a frequency of 20 MHz. While this technology was chosen to demonstrate this concept without the excessive burden of fabricating and testing much larger arrays, the validity of this analysis and architecture is not limited by the type of transducer. This problem of capacitive loading for the parallel array will become an issue even for more efficient transducers when frequency and/or array size is increased. The proposed architecture presents a system-level solution to one of the inherent device limitations of RF MEMS filters, thus opening up the potential for realizing chip-scale spectrum analysis for cognitive radio applications.

Acknowledgment

The authors wish to thank Rockwell Collins and the DARPA MTO ASP and ART programs for their support and the U.S. Army Research Laboratory, specifically Joel Martin and Brian Power, for fabrication. The authors also would like to thank the Cornell Nanofabrication Facility for help with post-processing and measurements.

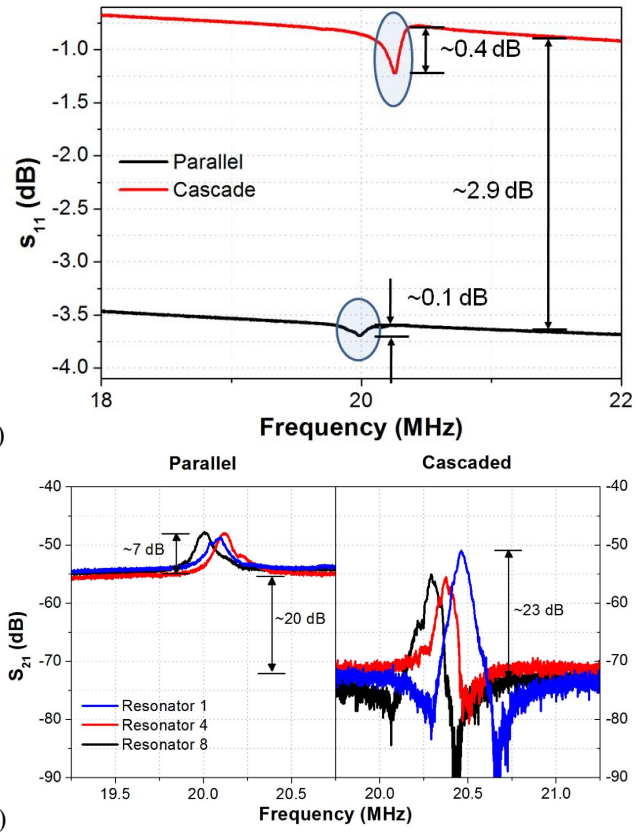


Fig. 10. Experimental results showing the (a) return loss and (b) transmission response for both parallel and series cascaded 9-channel filter arrays.

References

[1] W.-T. Hsu and K. Cioffi, "Low phase noise 70MHz Micromechanical Reference Oscillators," IMS 2004, pp. 1927-1930.

[2] H. M. Lavasani *et al.*, "A 145MHz low phase-noise capacitive silicon micromechanical oscillator," IEDM 2008, pp. 675-678.

[3] B. P. Otis and J. M. Rabaey, "A 300-µW 1.9-GHz CMOS oscillator utilizing micromachined resonators," IEEE J. Solid State Circuits, vol. 38, no. 7, July 2003, pp. 1271-1274.

[4] R. Ruby *et al.*, "Ultra-minature high-Q filters and duplexers using FBAR technology," ISSCC 2001, pp. 120-121, 438.

[5] E. H. Armstrong, "A study of heterodyne amplification by the electron relay," Proc. IRE, vol. 5, no. 5, pp. 145-159, 1917.

[6] C. Stevenson *et al.*, "IEEE 802.22: The first cognitive radio wireless regional area network standard," IEEE Communications Magazine, vol 47, no. 1, pp. 130-138, January 2009.

[7] A. B. MacKenzie *et al.*, "Cognitive radio and networking research at Virginia Tech," Proc. IEEE, vol. 97, no. 4, pp. 660-688, April 2009.

[8] S.-S. Li *et al.*, " A micromechanical parallel-class disk-array filter," IFCS 2007, pp. 1356-1361.

[9] R. H. Olsson *et al.*, "Multi-frequency aluminum nitride micro-filters for advanced RF communications," GOMAC 2010.

[10] N. Sinha *et al.*, "Integration of AlN micromechanical contour-mode technology filters with three-finger dual beam AlN MEMS switches," IFCS-EFTF 2009, pp. 1-4.

[11] J. Pulskamp *et al.*, "Monolithically integrated piezoMEMS SP2T switch and contour-mode filters," MEMS 2009, pp. 900-903.

Laterally coupled narrow-band high overtone bulk wave filters using thinned single crystal lithium niobate layers

Dorian Gachon
Laboratoire PROMES-CNRS, UPR 8521
Université de Perpignan Via Domitia,
Perpignan, France
dorian.gachon@univ-perp.fr

Thomas Baron, Gilles Martin, Eric Lebrasseur, Emilie Courjon, Florent Bassignot, Sylvain Ballandras
FEMTO-ST, UMR 6174 CNRS-UFC-ENSMM-UTBM
Time and Frequency Department
Besançon, France
sylvain.ballandras@femto-st.fr

***Abstract*—This work is devoted to the development of laterally coupled filters built on compound single crystal substrates. High overtone bulk acoustic resonators are built on $LiNbO_3$ thinned films bonded on thick $LiNbO_3$ or Quartz substrates. Two resonators are coupled via a narrow gap between their electrodes, yielding the possibility for their modes to interact and to produce coupled mode resonance conditions. The coupling efficiency is shown to be dependent on the frequency (wavelength), yielding well defined coupled-mode filters at intermediate frequencies (300 – 800 MHz) and single mode transfer functions above 1 GHz. The implementation of an oscillator stabilized by such a filter at 1.7 GHz is demonstrated.**

I. Introduction

Radio-Frequency (RF) filters generally are built using surface acoustic wave resonating devices or bulk acoustic resonators exhibiting single resonance behavior well-suited for the definition of high spectral purity filter capable to answer the drastic system specification generally imposed for telecommunication applications [1]. Hence, a rather small interest was devoted to RF filters based on high overtone bulk acoustic resonators (HBAR) apart from pioneer work by Lakin in early nineties [2]. However, these devices were mainly based on impedance element and required a very large number of resonators in series and parallel (6 "L" stages) for achieving a rejection of about 40 dB. Although developing filters based on HBAR yields the replication of the corresponding transfer function more or less for each resonator harmonics (at least those exhibiting similar coupling than the mode considered for the filter design), their capability to accurately select given frequencies finds some interest for oscillator or sensor purposes. Therefore, the possibility for manufacturing compact filter structures with reduced number of cells compared to the above-mentioned solution appears relevant. Particularly, structures comparable to those used in monolithic filters would be particularly interesting if transferable to HBAR because of their capabilities to produce very large rejection with a limited number of poles. Until now, only few attempts for developing such filters on thin film based material stacks have been successfully achieved [3].

In this work, we investigate the possibility for manufacturing laterally coupled filters using thinned single crystal lithium niobate layers acting as HBAR transducer, bonded on lithium niobate or quartz substrates. The idea consists in exploiting dispersion properties as well as the presence of a thick resonating structures (i.e. the substrate) to exhaust the lateral coupling between two resonators placed very close one another to favor their modal interaction. (YXl)/36° and (YXl)/163° $LiNbO_3$ cuts have been chosen to allow for the excitation of very high overtones, taking advantage of their strong electromechanical coupling (k_s^2 respectively equal to 20 and 45%), and to excite either longitudinal or shear waves. The substrates were respectively $LiNbO_3$ (YXl)/36° for filter demonstration and AT Quartz for temperature compensation purpose [4].

Many structures have been manufactured and have allowed for filter characterization in the vicinity of 300, 500 and 700 MHz (respectively the third, fifth and seventh harmonics of the transducer thinned film alone). Rejection in excess of 20 dB are demonstrated with a single filter cell. Corresponding Insertion losses of about 15 dB are emphasized and could be easily improved by impedance matching. Coupling two cells consequently provide rejection in excess of 40 dB with insertion losses of about 10 dB once matched. In the case of the Niobate/Quartz structure however, a single-mode response is obtained near 1.7 GHz, as the coupling conditions obviously are less efficiently met at this frequency than in the previous case. As a shear mode is exploited here, temperature compensation is obtained, yielding favorable conditions for using the device to stabilize an oscillator loop near this frequency. The four-port structure was easily looped using a wide band amplifier and the corresponding phase noise was found at 130 dB at 10 kHz from the carrier,

978-1-61284-110-6/11/$26.00 ©2011 IEEE

competing well with state-of-the-art solutions. In the following sections, the paper then shows the analysis of the proposed device, its implementation and tests and concludes on its interest for signal processing applications

II. Device description and analysis

The principle of the device is inspired from the so-called monolithic filters based on coupled bulk waves in single crystals and illustrated in Fig.1. This is achieved by setting two resonators very close one another. The gap between these resonators must be narrow enough to allow for evanescent waves apart the resonator electrodes to recover and hence yielding mode coupling conditions. As both resonators operate similarly, they do present the same spectral distribution and each harmonic bulk mode of the structure is identically coupled at almost exactly the same frequency. The above mentioned coupling condition then favors for each coupled mode the existence of two distinct resonances corresponding to symmetrical and anti-symmetrical coupled modes in the structure composed by the two resonators. The width of the separation area then controls the spectral distance between the two coupled modes. This operation principle only holds if the resonator actually traps the wave under its electrode, and therefore this condition must be validated by the analysis of dispersion modes, as shown in many text books and papers (see for instance [5]). Although this analysis was not conducted for this paper, the evanescent conditions are attested by the extremely high Q factor obtained using the tested configuration ((YX*l*)/36° $LiNbO_3$ thinned plate on (YX*l*)/36° $LiNbO_3$ substrate [5]), i.e. more than 50 000 at 1.5 GHz.

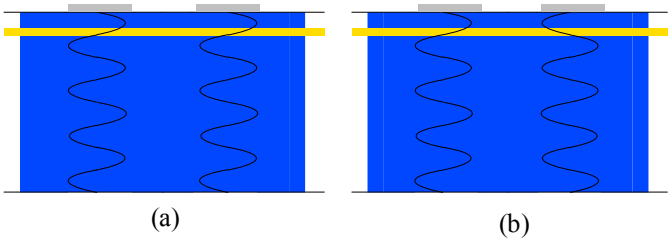
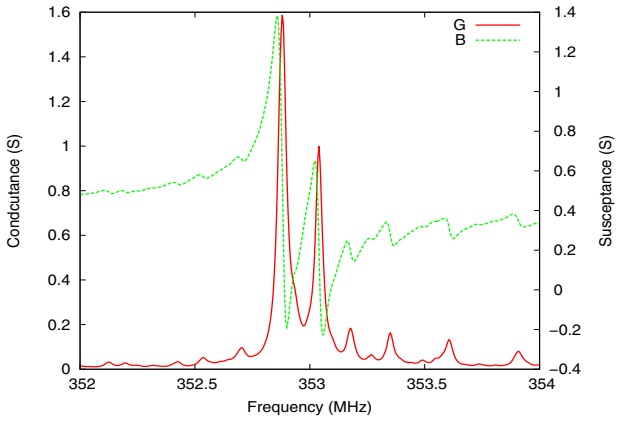
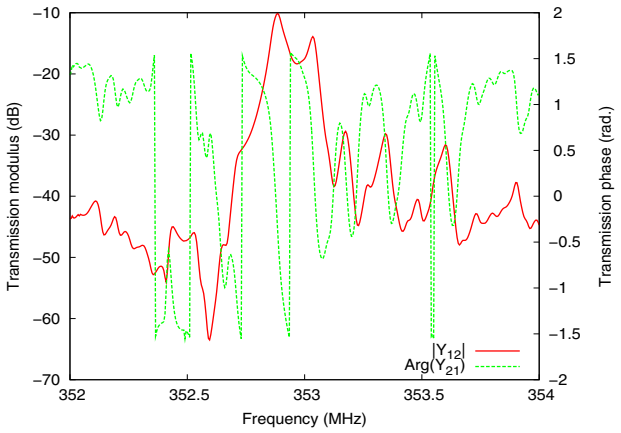


Figure 1. Principle scheme of the laterally-coupled-mode HBAR filter (a) symmetrical mode (b) antisymmetrical mode

The above presented principle has been applied for frequency filters in the rang 1-30 MHz using standard bulk wave resonators exploiting the fundamental mode of a single-crystal or ceramic piezoelectric plate compliant with the above-discussed coupling conditions. Higher frequencies are hardly reachable using these devices because of the thinning conditions required for such a purpose, yielding fragile devices poorly suited to industrial processes together with significant difficulties to match the frequency target. Recent developments have shown however that the principle could be applied in some extent to film bulk acoustic resonators (FBARs) by adapting the resonator structures for meeting the expected coupling conditions. In this work, we exploit the possibility for manufacturing HBARs based on thinned single crystal layers atop any substrate to check the possibility for fabricating such filters at high frequencies, i.e. near or above 1GHz. Preliminary to manufacturing and test, computation work has been conducted to check the actual capability of such a structure to behave as expected and to allow for coupling modes similarly to what happens in monolithic filters. Therefore, a periodic 2D-finite-element-analysis (FEA) coupled with a boundary radiation condition (so_called BEM) has been disorientated to address the problem of non periodic structures, adding absorbing conditions (Perfectly Matched Layers – PML) at the edge of the meshed period [8]. The thinned $LiNbO_3$ plate then is meshed with the two top Al electrodes separated by a 10µm gap. The BEM parts represents the finite thickness substrate accounting for the bottom Au electrode (the bonding layer). Mainly based on PML, the model requires some parametric adjustment but realistic operating conditions were tested in the vicinity of 300 MHz. Figure 2 shows the admittance parameters of the corresponding contribution (Y_{11} and Y_{21}). The first one clearly shows the existence of two modes very close one another whereas the second gives a prediction of the pass band of the 2-pole filter when using the structure as a quadrupole. To definitely confirm the actual operation of the device, the deformed mesh corresponding to each mode also has been plotted in Figure 3, showing the mode phase alternation.

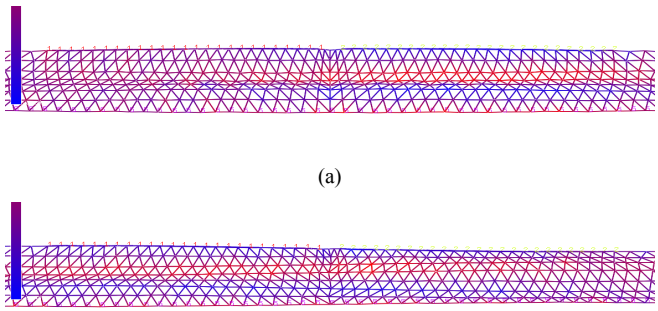


(a)



(b)

Figure 2. Theoretical response of a (YX*l*)/36°/(YX*l*)/36° $LiNbO_3$ laterally-coupled-mode HBAR filter (a) admittance (b) transfer



(a)

(b)

Figure 3. Deformed meshes corresponding to the excitation of the two modes of the filter (a) symmetrical mode (b) anti-symmetrical mode

This theoretical work tends to demonstrate the actual possibility for the fabrication of laterally-coupled HBAR filters. The manufacture of such devices is addressed in the next sections.

III. Implementation and Characterization

The fabrication of single crystal wafers devoted to HBAR applications has been already published several times and is not described here in details. Information about the process can be found in [4]. We have been using first a material stack consisting of (YX*l*)/36° lithium niobate wafers, one being thinned down to 30µm and operating as the HBAR transducer, the other acting as resonator (mechanical tank) because of the very high Q factor obtained with such structures, assuring a high quality of energy trapping under the top electrode. These first tests have been achieved by splitting the top electrode (aluminum), thus giving rise to two resonators separated by an insulating strip of 10µm, with dimensions 145 × 200 Mm ². The contact mass is obtained by returning it on the upper surface of the substrate, to actually achieve the quadrupole of Figure 4 shows a SEM photograph of this achievement.

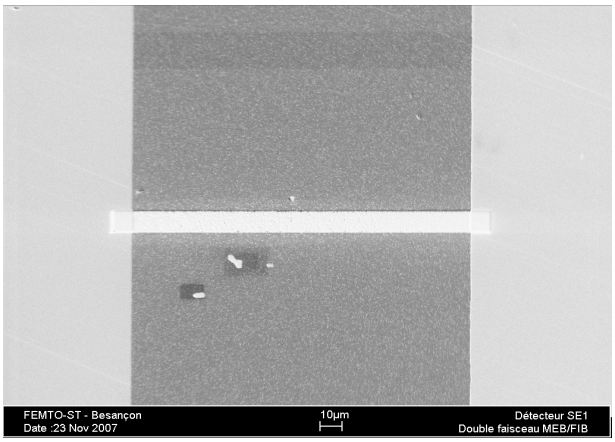


Figure 4. SEM picture of the filter structure with two coupled resonators stacked laterally-coupled LiNbO3 (34µm) / Au (300nm) / LiNbO3 (350µm). The dividing line drawn on the FIB

We were able to characterize the filter response in terms of transfer function for different frequency bands, corresponding course to areas of high spectral coupling of the longitudinal mode. Figure 5 shows the type of spectral response in the vicinity of the 3rd harmonic of the fundamental mode of the transducer.

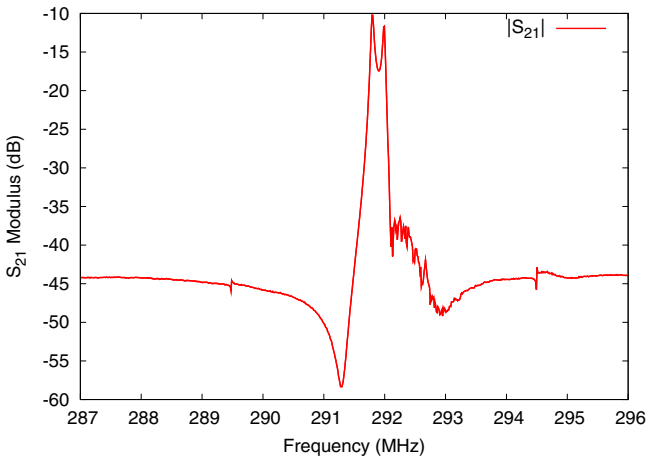


Figure 5. Transfer function of the filter at around 300 MHz.

Several filtering functions are visible, separated by 10 MHz from each other, with a maximum dynamic range around 500 MHz as shown in Figure 6. We then see a rejection out of band approaching 30 dB (for a single filtering cell) and losses approaching 10 dB (no impedance matching), but also a strong ripple in the band rendering inoperative the filter given specifications generally effective for this application. However, there are very clearly the two poles forming the filter, which seems to prove clearly enough by the operation

mode coupling of the device. We also conducted these measurements in the vicinity of 500 MHz and 700 MHz, as shown in Figures 6, 7 and 8. In particular, we see that with the increasing frequency, filter bandwidth is less hollow and more regular, with a slight reduction of-band rejection (25 dB at 500 MHz and 22 dB at 720 MHz) recalling all the same it is only a single cell. The losses in the band from 12 to 15 dB, but we recall once again that every effort has been dedicated to the impedance matching of the two resonators. We finally remark on many measures a contribution of undesired shear mode (each 6 MHz) although very low. The transfer function of the filter at 720 MHz yet reveals a parasitic contribution more pronounced in low band whose origin is not yet identified. The 3dB bandwidth of such filters covers 200 kHz, a value fairly modest to link to the coupling measured for the corresponding modes, reaching at most a few thousand. Efforts remain to be done to improve this situation. We attempted to identify the upper limit in operating frequency of the structure. Figure 9 shows the transfer function of the filter in the vicinity of 1.5 GHz. It is clear that at this frequency, the filtering function is no longer provided by the structure, which appear with difficulty out of band rejection of 8 dB loss clearly increased (again without effort impedance matching) .

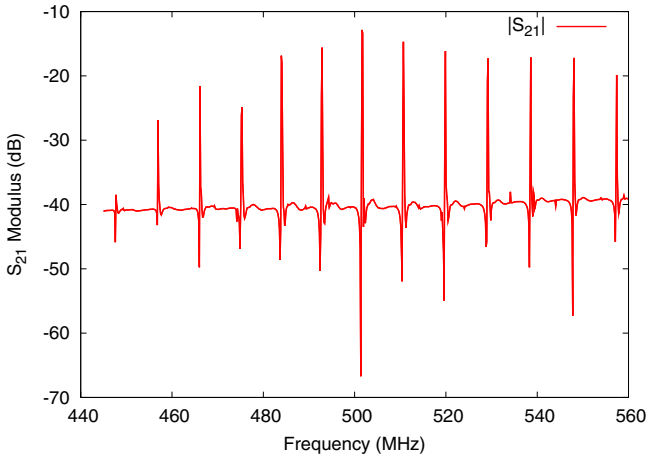


Figure 6. Transfer function in wide band of the filter in the vicinity of 500 MHz (5th harmonic of the fundamental mode of the layer of transduction).

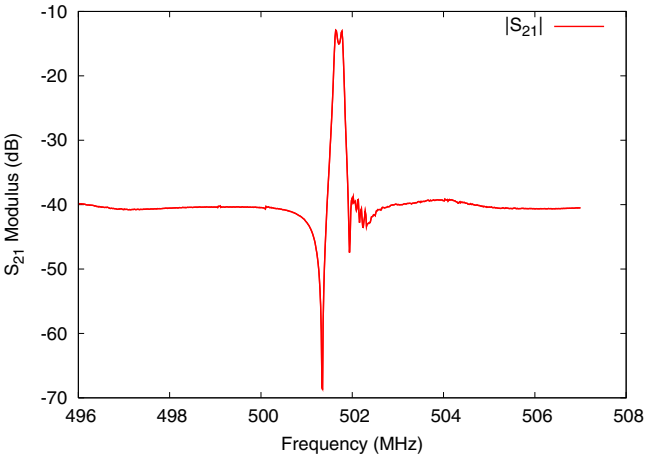


Figure 7. Transfer function of the filter in the vicinity of 500 MHz, improved in-band ripple, reduced out of band rejection and frequency of the radiated high.

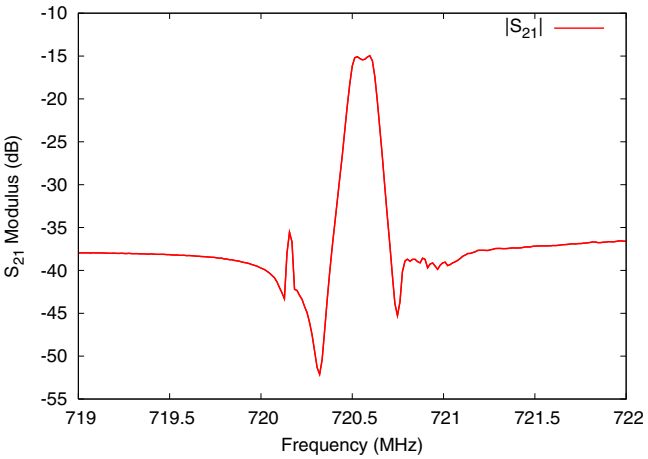


Figure 8. Transfer function of the filter in the vicinity of 720 MHz, further improvement of the ripple in the band, reduced out of band rejection and radiated high frequency

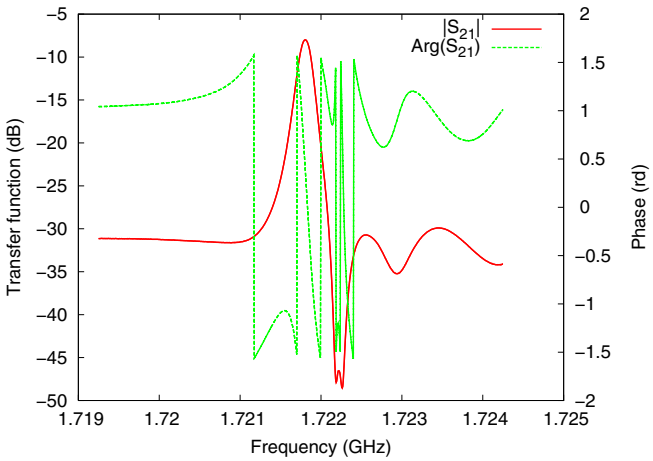


Figure 9. Transfer function of the filter in the vicinity of 1.5 GHz, operational limits of the structure (low-band rejection)

IV. Conclusion

In this paper we showed the development of very narrow band filters based on HBAR. This contribution brings an original and simple to implement cross-coupled filters. Indeed, the quadrupole structure lends itself well known in the manufacture of filters whose parameters may be imposed by the planar shape of the electrodes and the spacing between two coupled resonators. It offers a significant simplification of the structure of the filters by its essentially planar, although it requires access to the underlying electrode in one way or another. The quality factor displayed by the quadrupoles show that it would be possible to stabilize an oscillator. However before using this solution, it is necessary to identify the parametric influences using simple models.

Acknowledgement

The authors would like to thank Roland Salut for the FIB etching process steps exploited in the presented work.

This work has been supported by the Centre National d'Etudes Spatiales (CNES) under grant #04/CNES/1941/00-DCT094, and by the Direction Generale de l'Armement (DGA) under grants #05.34.016 and #07-34-020.

References

[1] C.C.W. Rüppel, T.A. Fjeldly (Eds). – Advances In Surface Acoustic Wave Technology, Systems And Applications. Vol. 1 and 2, Selected Topics in Electronics and Systems, World Scientific Publishing Co Pte Ltd (2001).

[2] K. M. Lakin, G.R Kline, K.T. McCarron, High Q microwave acoustic resonators and Filters, Microwave Symposium Digest,1993, pp.1517-1520

[3] T. Pensala, J. Meltaus, K. Kokkonen, M. Ylilammi, 2-D modeling of laterally acoustically coupled thin film bulk acoustic wave resonator filters. IEEE Trans Ultrason Ferroelectr Freq Control. 2010 V.57, n°11, 2010,pp. 2537-2549

[4] T.Baron, D. Gachon, G. Martin, S. Alzuaga, D. Hermelin, J.P. Romand, S. Ballandras, Temperature compensated Radio-Frequency Harmonic Bulk Acoustic Resonators, Proc.of IEEE IFCS, 2010

[5] D.Gachon, E. Courjon, J. Masson, V. Petrini, J.Y. Rauch, S. Balandras, LiNbO3-LiNbO3 High Overtone Bulk Acoustic Resonator Exhibiting High Q.f Product, Proc. of Ultrasonics Symposium, 2007, pp1417-1420

Experimental Demonstration of a Phase-Modulated Optoelectronic Oscillator Using Balanced Detection

Patrick T. Callahan, Michael L. Dennis, and Thomas R. Clark
Johns Hopkins University Applied Physics Laboratory
Laurel, MD, USA
Email: patrick.callahan@ieee.org

***Abstract*—The use of phase modulation in lieu of the typical amplitude modulator in an optoelectronic oscillator (OEO) has not yet been extensively investigated. In this paper we present the first experimental demonstration of a phase-modulated OEO using a fiber delay-line demodulator with balanced detection. Preliminary phase noise and frequency stability measurements are also demonstrated.**

I. Introduction

The optoelectronic oscillator (OEO) has attracted a great deal of interest for its potential to generate ultra-low phase noise microwave signals. In the years since the publication of the seminal work on this topic by Yao and Maleki [1,2], there has been a rapid proliferation of new OEO systems widely varying in design and sophistication. However, the vast majority of previous demonstrations have relied upon an intensity-modulated direct-detection (IMDD) approach within the resonator. An architecture based on phase modulation instead of intensity modulation offers several advantages that can enable improved phase noise performance. In this paper, we introduce a novel architecture for an OEO based on phase modulation, and present phase noise and frequency stability characterization of the oscillator at 10 GHz. To our knowledge, this work constitutes the first experimental demonstration of a phase-modulated OEO using balanced detection.

II. Phase-Modulated OEO: A Novel Architecture

A simplified diagram of a conventional IMDD-based OEO is shown in Figure 1(a). One main advantage of using a phase modulator instead of an amplitude modulator is that $LiNbO_3$ phase modulators have lower optical insertion loss. Typically the difference is approximately 2-3 dB. In order to generate a feedback signal from the photodiode, there must be a phase-to-amplitude conversion element within the loop. One previous demonstration of a phase-modulated OEO has used fiber-Bragg-gratings (FBG) to filter off one of the phase-modulation sidebands, thereby allowing the other sideband to beat directly with the optical carrier to produce an intensity modulated signal [3]. A more efficient method is to make use of a dispersive element within the loop, which changes the phase relationship between the two modulation sidebands. The sidebands on either side of the carrier are initially exactly out of phase; introducing a dispersive element allows the mixing products between each sideband and the carrier to add constructively instead of destructively, thus generating an intensity modulation. One group has achieved this using a chirped FBG [4], while another has used a whispering-gallery-mode (WGM) resonator as the dispersive element [5]. Both of these methods suffer from high optical insertion loss, in the first case due to the necessity of using an optical circulator (which typically have 1-3 dB of excess loss), and in the second case due to the losses incurred when coupling to the WGM. Our system instead uses a delay-line demodulator to convert from phase to amplitude modulation, an interferometric process that can be made nearly lossless by using fiber couplers.

The intensity modulation is then detected using photodiodes; for our system this is accomplished via balanced detection, in which the two complementary interference products from the demodulator are each detected and then

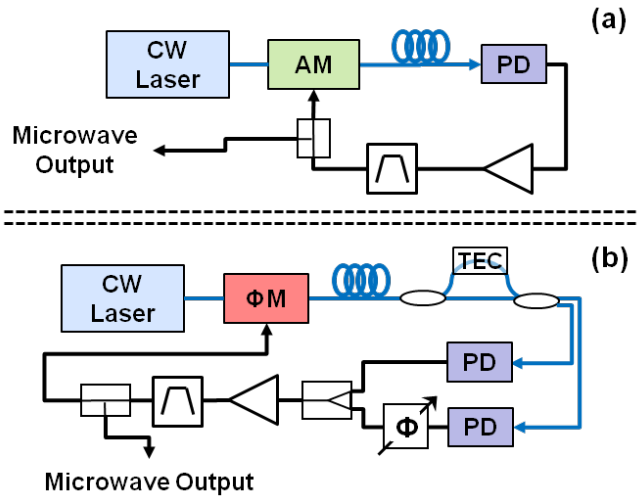


Figure 1. Diagram of (a) the typical intensity-modulated, direct-detection OEO architecture; and (b) the phase-modulated, balanced detection architecture using a delay-line demodulator. Abbreviations are as follows: CW – continuous-wave, AM – amplitude modulator, PD – photodiode, ΦM – phase modulator and TEC – thermoelectric heater.

This material is based on independent work supported in part by the Stuart S. Janney Program at the Johns Hopkins University Applied Physics Laboratory.

combined in phase using an RF phase shifter. This architecture is shown in Figure 1(b), and is similar to the design proposed (but not implemented) in [6]. The immediate advantage of balanced detection over the standard IMDD architecture is that for the same laser power, twice as much optical power is converted to RF signal, so the total RF power generation within the loop is greater by 6 dB. Including the reduced insertion loss for phase modulators compared to intensity modulators, this yields a net advantage of approximately 10 dB. This increase in loop gain enables the possibility of operating the OEO without an RF amplifier [7]. This in turn can significantly reduce the phase noise of the oscillator.

III. Theory: Threshold and Expected Phase Noise

A. Oscillation Threshold Condition

The mathematical expression for the threshold condition can be easily derived for the architecture described above. The derivation proceeds very similarly to that described in [1], and so is not presented here. For a phase-modulated OEO using a delay-line demodulator with balanced detection, the oscillation threshold is given by:

$$\sqrt{G}\alpha\eta P_0 R \sin(\omega_0 \Delta t_d) \geq V_\pi / \pi , \qquad (1)$$

where G is the net power gain (including RF losses) of the amplifier, α is the optical insertion loss through the loop, η is the photodiode responsivity, P_0 is the optical power at the input to the phase modulator, R is the photodiode resistance, ω_0 is the optical angular frequency, Δt_d is the differential delay of the demodulator, and V_π is the half-wave voltage of the modulator. When the delay-line demodulator is held at the quadrature condition using the proper temperature control settings, the sine term becomes unity, and this expression differs from the canonical case in [1] only by a factor of two. This reduction in the required threshold power with respect to an IMDD architecture is due to the fact that balanced detection captures both interference products from the demodulator and adds them in phase.

B. Expected Phase Noise Performance

The expression for the RF power spectral density can be used as a good approximation for the phase noise spectral density when phase noise processes dominate over amplitude noise processes. For the loop lengths used in our system, this is a reasonable assumption. Therefore using the expression derived in [2] for the RF spectrum, we can expect the phase noise to be given by:

$$\mathcal{L}(f) = \frac{G\rho_N}{(2\pi\tau_l f)^2 P_{osc}} , \qquad (2)$$

where G is again the net gain of the amplifier, ρ_N is the noise spectral density at the input to the amplifier, τ_l is the round-trip time in the loop, and P_{osc} is the RF power in the oscillating mode. In the simplified model presented here, ρ_N includes thermal noise, the photodiode shot noise and the laser relative intensity noise (RIN) evaluated at the rms photodiode operating current. We approximated the laser RIN as a white noise process, choosing a representative value of -145 dBc/Hz. Substituting the various system parameters into (2) gives an expected value of -89 dBc/Hz at 1 kHz offset for a 120 m loop, and -108 dBc/Hz at 1 kHz offset for a 1 km loop. Note that the amplifier gain acts to increase the noise in (2). Therefore if the power in the oscillating mode was sufficient to achieve threshold without the amplifier, a considerable reduction in the phase noise would be possible.

IV. Experimental Setup

A picture of the experimental setup for our oscillator is shown in Figure 2. The optical input was a 1550 nm laser followed by an erbium-doped fiber amplifier. The RIN for this laser was specified to be <145 dBc/Hz at 100 Hz offset. For our system, the required input optical power to reach threshold is approximately 20 mW. The typical operating power used was 60 mW at the input to the phase modulator. The modulator had a half-wave voltage of approximately 3 V at 10 GHz, and an optical insertion loss of 2 dB. For the 120-m loop configuration, the output of the modulator was connected to a fiber stretcher, for use in length stabilization. For the 1-km loop configuration the fiber stretcher was omitted from the setup. The main fiber delay line and the delay-line demodulator were housed in a box lined with foam to provide a limited degree of acoustic and thermal isolation. The delay-line demodulator was of a type widely used in differential phase-shift-keyed optical receivers, and had a differential delay of 125 ps. The phase was manually controlled by heating one arm of the interferometer with a thermoelectric element. The two outputs of the demodulator were then detected by the photodiodes, which had a responsivity of approximately 0.7 A/W. The photodiodes were internally terminated with a 50 Ω resistor, and had a 3-dB bandwidth of 12 GHz. The RF outputs were combined in phase using a microwave phase shifter, and passed to the RF amplifier. The amplifier had 20 dB of gain at 10 GHz. Following the amplifier was an RF filter with a 3-dB bandwidth of 100 MHz centered at 10 GHz, and 4 dB of RF insertion loss. The output of the filter was then passed through a 6-dB tap coupler to provide access to the 10 GHz microwave signal, and finally fed back onto the RF input port of the modulator.

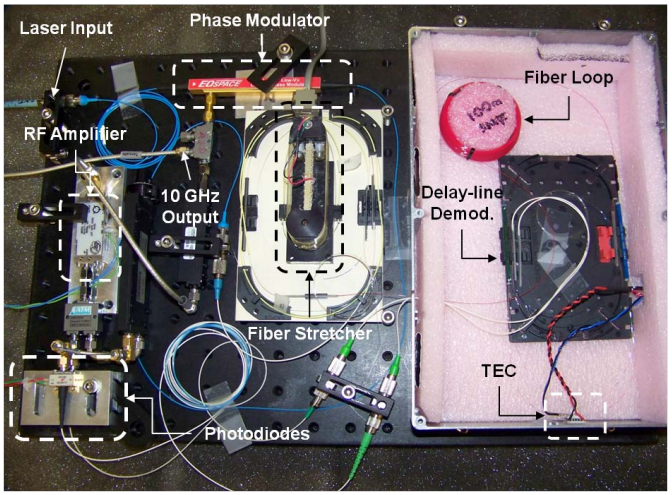


Figure 2. Picture of the oscillator setup in the 120-m fiber loop configuration. For the 1-km loop configuration, the fiber stretcher was not in the loop.

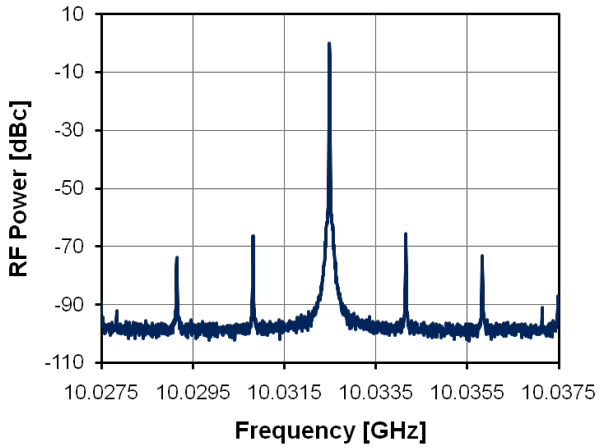


Figure 3. RF spectrum of the OEO with a 120-m fiber loop. The frequency span was 10 MHz, the resolution bandwidth was 10 kHz, and 16 averages were taken.

V. Results & Discussion

A. Phase Noise for a 120-m Fiber Loop Resonator

Figure 3 shows the RF spectrum of the OEO with a 120-m loop of fiber in the resonator. The central tone of the oscillator is at 10.032 GHz, and the spurs are side modes of the resonator separated by the inverse of the round-trip time. For 120 meters of fiber this corresponds to 1.67 MHz, as can be seen from the spectrum. The stability of the OEO in this configuration was poor; thermal fluctuations in the surrounding environment caused the oscillator frequency to drift on the order of a few hundred hertz per second. This was greater than the capture range of the phase-locking hardware for our commercial phase noise test set (Agilent E5500). While a measurement of the absolute phase noise with the two-oscillator technique was not possible, we were nonetheless able to characterize the phase noise of the OEO by closely examining the RF spectrum. Strictly speaking a measurement with a spectrum analyzer will yield the total integrated phase and amplitude noise of the system, but in our system the oscillator phase noise is far dominant over amplitude noise processes.

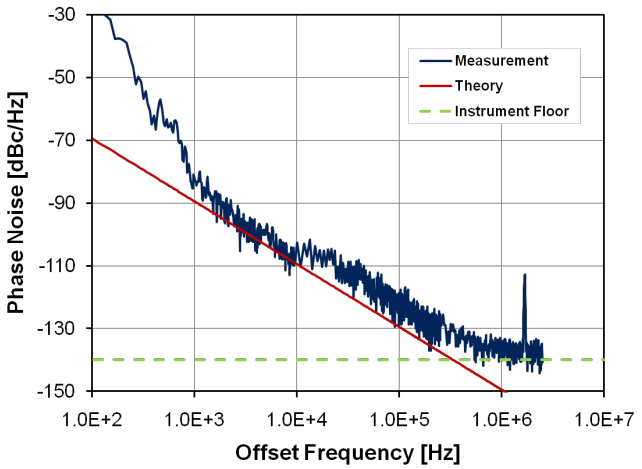


Figure 4. Composite spectrum for phase noise approximation of the OEO with a 120-m fiber loop (solid blue). Predicted phase noise performance from theory (solid red) and instrument noise floor (dashed green) are also shown.

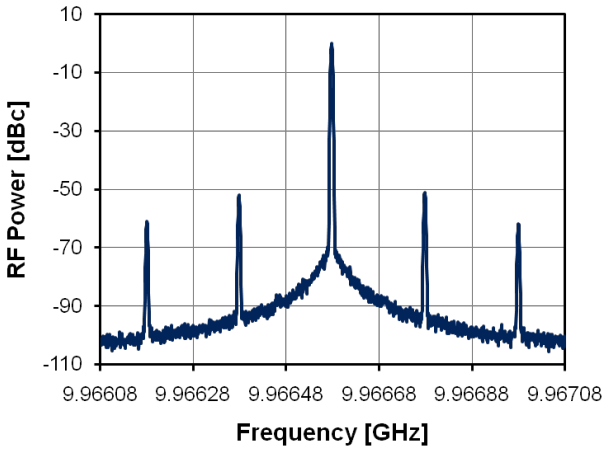


Figure 5. RF spectrum of the OEO with a 1-km fiber loop. The frequency span is 1 MHz, the resolution bandwidth is 3 kHz, and 16 averages were taken.

Shown in Figure 4 is a composite phase noise picture constructed from several measurements of the spectrum at different frequency spans and resolution bandwidth settings. The composite is normalized to a 1-Hz resolution bandwidth for ease of comparison with traditional phase noise measurements. Shown in red is the phase noise performance predicted by the theory in Section III. Below 1-kHz offset from the carrier, the random-walk and frequency drift induced by thermal fluctuations is readily apparent, and manifests as a $1/f^4$ slope. Above 500-kHz offset the measurement begins to be limited by the noise floor of the spectrum analyzer. The first side mode of the resonator appears again at 1.67 MHz offset, although for the measurements of that portion of the spectrum, the resolution bandwidth (30 kHz) was too broad to properly resolve the peak. This accounts for the discrepancy in the spur level between Figures 3 and 4.

B. Phase Noise for a 1-km Fiber Loop Resonator

The oscillator was reconfigured using a 1-km long fiber delay instead of 120 m, with the expectation that a longer loop would result in a more stable output frequency. We posit that the dominant source of frequency instability for the 120-m system was most likely thermal fluctuations. This implies that a longer loop delay would lead to increased stability due to statistical averaging of local temperature variations over the length of the fiber delay. A system with a longer delay would thus benefit from a greater degree of averaging, decreasing the net variance in frequency due to inhomogeneity in the temperature distribution. From (2), we also expect to achieve an improvement in the phase noise of ~18 dB. Figure 5 shows the RF spectrum for the 1-km loop system. Here the side modes appear at 200-kHz offset, corresponding to the longer round-trip time. The significant improvement in stability over the shorter loop allowed us to phase-lock a microwave reference oscillator to the OEO, enabling a measurement of the absolute phase noise. The method used for this measurement was the standard two-oscillator technique, in which the device under test is held in quadrature with a stable reference source via a phase-locked loop (PLL). Using a double-balanced mixer as a phase detector, the phase error between the two oscillators is then measured with a spectrum analyzer for offset frequencies outside the bandwidth of the

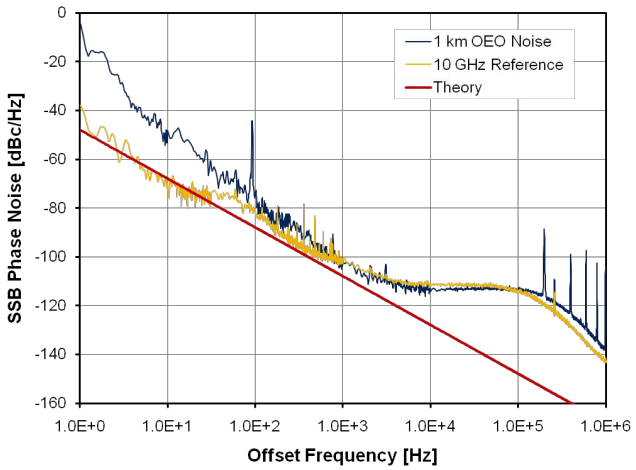


Figure 6. Single-sideband phase noise measurement of the OEO noise for a 1-km fiber loop (dark blue) and the 10-GHz reference oscillator (orange). Expected phase noise performance from the theory is shown in red.

PLL. Within the bandwidth of the PLL, a fast Fourier transform (FFT) analyzer measures the FFT of the correction signal that is applied to the reference oscillator in order to keep the two signals phase-locked. The results are shown in Figure 6, where the OEO noise is in dark blue. To measure the contributions of the reference oscillator, we used two frequency synthesizers that had similar performance and measured them against one another. Unfortunately the noise of the reference oscillator dominates over the OEO noise for almost the entire measurement, as is seen by the orange trace. Shown in red is the theoretical performance that we would expect to measure for the OEO if a higher-quality reference was available. Below 100-Hz offset, the $1/f^3$ slope indicates flicker frequency noise, which could be caused by acoustic vibrations, added noise due to the RF amplifier, or potentially by the photodiodes.

C. *Frequency Stability for a 1-km Fiber Loop Resonator*

To characterize the free-running stability of the oscillator, we measured the output frequency with a spectrum analyzer, operating as a frequency counter. The counter recorded the peak frequency in 1-s intervals for approximately 30 minutes.

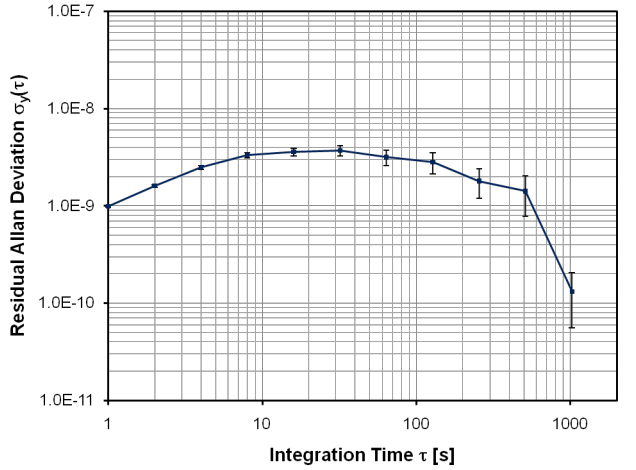


Figure 7. Allan deviation of fractional frequency fluctuations for the 10 GHz output of the 1-km OEO. The Allan deviation was calculated after subtracting out the effect of linear frequency drift.

To calculate the Allan variance, we first fit the data to a line, and subtracted out the linear frequency drift caused by global temperature changes in the laboratory. The drift rate was found to be ~17 Hz/s, corresponding to an approximate temperature drift rate of 2 mK/s during the length of the measurement. The Allan deviation of the residual frequency fluctuations is shown in Figure 7. The predominantly flat shape of the curve is consistent with the flicker frequency noise effects observed in the phase noise plot. The average frequency stability (without linear drift) is $\sigma_y = 3\times10^{-9}$.

VI. Conclusion

We have presented the first experimental demonstration of a phase-modulated optoelectronic oscillator using a delay-line demodulator with balanced detection. The noise characteristics of the OEO were investigated in two configurations, with a 120 m fiber delay and with a 1 km fiber delay. Measurement of the phase noise of the 1-km loop OEO was primarily limited by the noise of the microwave reference oscillator, as well as flicker noise sources close to the carrier. The RF amplifier is likely to be a significant contributor of flicker noise. As such, a primary goal of future work on this project will be to generate sufficient gain within the OEO loop such that the amplifier is unnecessary. This can be readily achieved with the proper choice of photodiodes and phase modulator. Photodiodes without internal terminations will provide 6 dB greater RF output power, and high-optical-power-handling photodiodes are also commercially available. Another method for improvement would be to use a phase modulator with a lower half-wave voltage, such as a modulator with a double-pass waveguide architecture. The frequency stability of the oscillator is limited primarily by thermal drift and fluctuations, which in turn change the length of the fiber delay. Thermal isolation and length stabilization techniques are potential areas of further investigation.

References

[1] X. S. Yao and L. Maleki, "Optoelectronic oscillator for photonic systems," *IEEE J. Quantum Electron.*, vol. 32, no. 7, pp. 1141-1149, July 1996.

[2] X. S. Yao and L. Maleki, "Optoelectronic microwave oscillator," *J. Opt. Soc. Am. B*, vol. 13, no. 8, pp. 1725-1735, August 1996.

[3] T. Sakamoto, T. Kawanishi and M. Izutsu, "Optoelectronic oscillator using a $LiNbO_3$ phase modulator for self-oscillating frequency comb generation," *Opt. Lett.*, vol. 31, no. 6, pp. 811-813, March 2006.

[4] W. Li and J. Yao, "An Optically Tunable Optoelectronic Oscillator," *J. Lightw. Technol.*, vol. 28, no. 18, pp. 2640-2645, September 2010.

[5] K. Volyanskiy *et al.*, "Compact optoelectronic microwave oscillators using ultra-high Q whispering gallery mode disk-resonators and phase modulation," *Opt. Express*, vol. 18, no. 21, pp. 22358-22363, October 2010.

[6] G. J. Dick and N. Yu, "A new OEO design using optical phase modulation and modulation suppression," *IEEE International Frequency Control Symposium 2006*, pp. 507-512, 2006.

[7] C. W. Nelson, A. Hati and D. A. Howe, "Relative intensity noise suppression for RF photonic links," *IEEE Photon. Technol. Lett.*, vol. 20, no. 18, pp. 1542-1544, September 2008.

Characterization of Surface Acoustic Wave Optomechanical Oscillators

Gaurav Bahl, John Zehnpfennig, Matthew Tomes, and Tal Carmon
Optics and Photonics, Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan, USA 48105
Email: bahlg@umich.edu

***Abstract*—We describe and experimentally demonstrate an optomechanical oscillator where light drives a surface mechanical mode of a spherical resonator, using a combination of photoelastic scattering and optical electrostriction. These oscillators are shown to have discretely-selectable frequencies ranging from 50 MHz to 1.5 GHz on a single silica microsphere device. We also report on initial measurements of phase noise and continuous frequency tuning of these oscillators.**

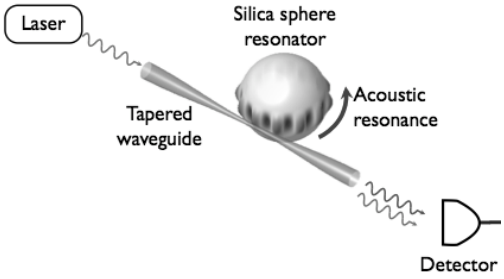
I. Introduction

In recent years, oscillators based on parametric mechanical vibration in optical microresonators have emerged as promising candidates for compact and low power photonic frequency references [1]–[4]. While prior work has focused on the use of radiation pressure and gradient forces [5] in order to drive mechanical vibration, here we exploit optical electrostriction [6] to stimulate (as in a laser) an acoustical whispering gallery resonance of the microsphere. These acoustical modes are concentrated at the equator of the sphere, and have been calculated through finite element modeling to be shear type and Rayleigh type [7] surface acoustic waves (SAW).

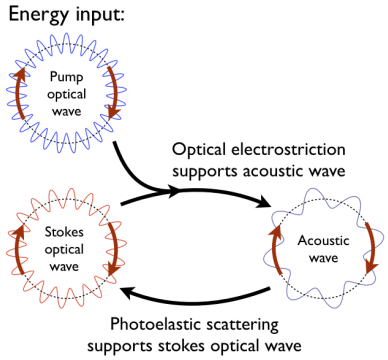
Such SAW-based optomechanical oscillators (OMO) have been theoretically proposed in the past [8]. In this new class of optomechanical oscillator, incoming light drives and interacts with the surface acoustic mode, resulting in two optical signals at the output of the device (Figure 1(a)). The beat note between these optical signals is measured with a photodetector, and is at the acoustical frequency of the surface mode that we excited.

II. Surface Acoustic Wave Optomechanical Oscillators

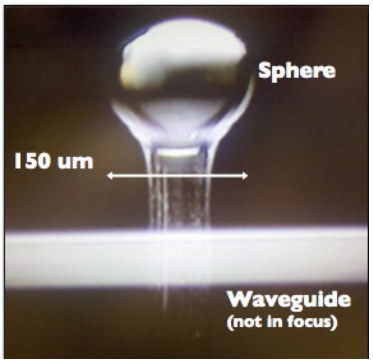
We will now describe the process of exciting optomechanical oscillations within these spherical resonators. An incoming pump photon scatters from an acoustic phonon occupying a mechanical eigenmode of the spherical resonator, resulting in a scattered Stokes photon that is down-shifted to a lower frequency. The scattered Stokes light together with the incoming pump light generates electrostrictive pressure which amplifies the initial acoustic wave. Energy and momentum are conserved in this process through the generation of phonons at the same acoustical frequency from which the pump light was initially scattered. The process is illustrated in Figure 1(b) and is similar to stimulated Brillouin scattering that takes place in bulk materials [6].



(a) Schematic of the oscillator. The detector outputs the beat note between the two optical signals, which is at the mechanical frequency.



(b) Optical electrostriction and photoelastic scattering drive the mechanical oscillations of a traveling acoustical mode



(c) Microscope image of the oscillator

Fig. 1. Surface acoustic wave optomechanical oscillator: Theory and practical realization

978-1-61284-110-6/11/$26.00 ©2011 IEEE

The experimental realization of a SAW-OMO (Figure 1(a)) requires two optical whispering gallery modes that have extremely close frequencies (in relative optical terms) but at the same time have different wavevectors. This requirement can be expressed as follows

$$M_P = M_S + M_A \tag{1}$$

$$\omega_P = \omega_S + \omega_A \ . \tag{2}$$

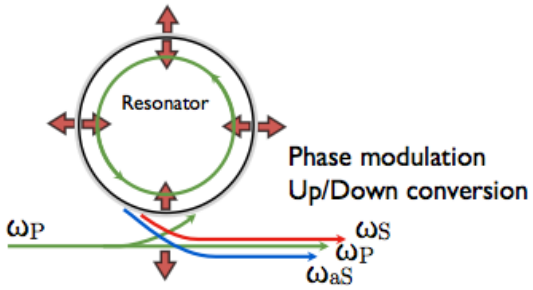
Here the M_i parameters are the number of periods of the optical modes ('P', 'S' subscripts) and of the acoustical mode ('A' subscript) around the circumference of the sphere, and their relationship describes momentum conservation. The simultaneous frequency relationship ω_i describes energy conservation. This phase matching requirement is the result of solving for a synchronous solution [9] to the optical and acoustical wave equations that describe the processes of photoelastic scattering and optical electrostriction within the sphere. Light is coupled in and out of the microsphere resonator [10] using a tapered optical fiber [11]–[13]. Optical quality factors are in the range of 10^8.

Previous demonstrations of oscillations induced via Brillouin scattering, shown in CaF_2 [14] and in silica [15], depended on back-scattering from the acoustical wave to close the internal feedback loop in the above described process. When we consider momentum conservation in back-scattering, the resulting phonons have a wavelength that is approximately half of the pump light used. For a 1.5 um wavelength pump propagating in silica, this translates to an optical wavelength of 1 um within the resonator, and an acoustical wavelength of 0.5 um resulting in frequencies in the 11 GHz range [15] for when silica is the host material. However, the lifetime for such hypersonic phonons is extremely short, resulting in low acoustical finesse. The acoustical frequencies are primarily dictated by material properties in this case.

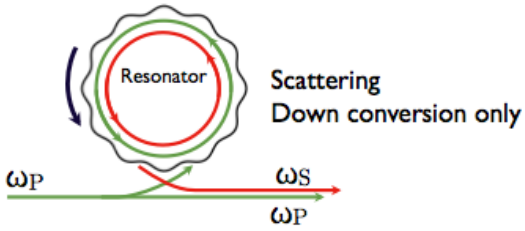
A question arises whether lower frequencies, dictated to some extent by geometry, can be targeted instead. Lower frequency phonons would have much longer lifetimes within the host material, and would have a much higher acoustical finesse within the microsphere. As opposed to previous demonstrations of backward scattering, here we rely on forward scattering [8], [16] to generate lower frequency phonons. Momentum conservation for forward scattering dictates that the acoustic frequency is significantly smaller, and is of the order of 100 MHz.

III. Identification of Mechanical Modes

Previous OMO demonstrations, have primarily been based on mechanical parametric oscillation of optical microcavities. The standard varying parameter in these cases is the optical path length and the mechanical modes can be described as breathing modes (Figure 2(a)). The resulting optical spectrum at the output thus shows modulation effects (phase and amplitude) with sidebands at both higher and lower frequencies. Higher frequency harmonics of these sidebands are also present [4].



(a) Breathe modes generate both upper and lower sidebands (with higher harmonics) to the optical input.



(b) Surface acoustic modes only generate a single red-downshifted sideband.

Fig. 2. Differentiating mechanical modes through optical output signals.

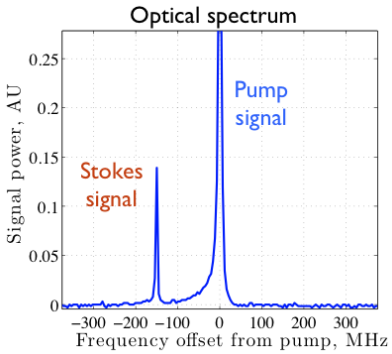
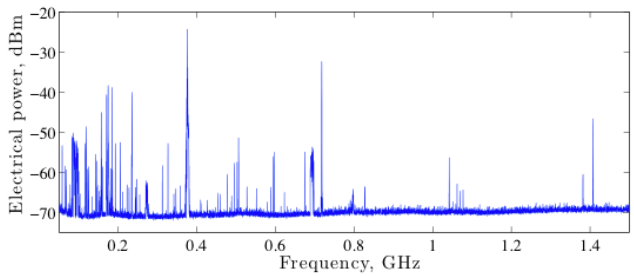


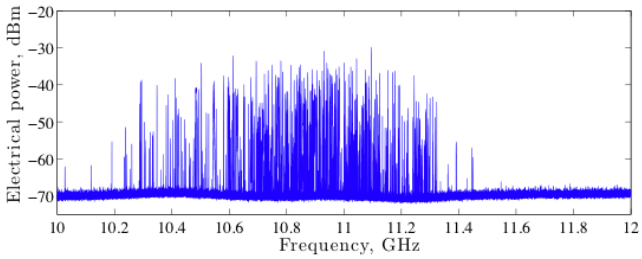
Fig. 3. Verification of the surface-wave nature of a 150 MHz oscillation. Only one Stokes red-shifted optical line is present adjacent to the pump.

In the case described in this work, the mechanical excitation is a surface mode that travels unidirectionally along the equator (Figure 2(b)). Light that is scattered from this traveling mode is always downshifted from the pump due to the always-receding nature of the mode. The optical path length does not change. As a result, there is only a single Stokes-frequency optical sideband that appears at the output.

Since there is a difference between the optical signatures of the two types of mechanical modes, we can distinguish between them by resolving the spectra of the output light using an optical spectrum analyzer. We have experimentally studied the optical spectra to confirm the generation of surface mechanical modes (Figure 3).



(a) Many low frequency oscillations (50 MHz - 1.4 GHz) can be excited (this work)



(b) High frequency oscillations around 11 GHz can also be excited (prior art)

Fig. 4. Excitation of various oscillation frequencies while slowly scanning the pump laser from 1520 - 1570 nm.

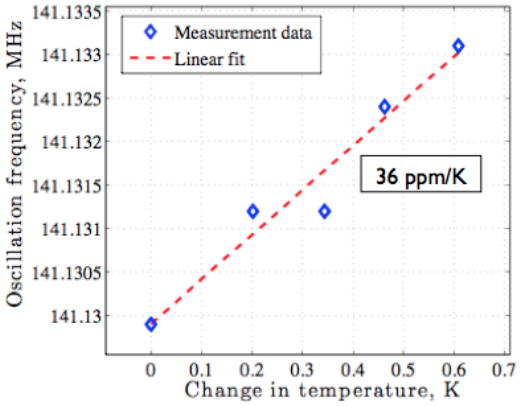


Fig. 5. The capability for continuous frequency tuning using temperature exists. The tuning slope for a given oscillation depends on the specific combination of optical and mechanical modes that participate in the process. The pictured measurement is for a 141 MHz oscillation.

IV. Experimental results

A. Oscillation frequency and tuning

We have experimentally determined (Figure 4) that a single device can show hundreds of discretely-selectable oscillation frequencies through the described process. In the experiment shown in Figure 4(a), we scanned the pump laser from 1520 nm – 1570 nm in order to probe many pump modes, while keeping the electrical spectrum analyzer on the 'peak hold' setting. In the low frequency regime that this paper describes, we were able to capture oscillations ranging from 50 MHz up to 1.4 GHz.

We can combine this result with the previously demonstrated ability to excite acoustical modes in the X-band (11.5 GHz) frequency regime via Brillouin back-scattering (Figure 4(b)). The implications of this combined result are that a single device can produce oscillations ranging from 50 MHz–1.4 GHz, and also from 10 GHz–11.5 GHz. Indeed there is a large gap between these frequency ranges where oscillations have not been experimentally demonstrated. We can, however, begin to fill this gap by using different pump wavelengths and different materials.

Beyond the capability for discrete frequency selection with this oscillator, we can also thermally tune the oscillation frequency (Figure 5). Here, the major effect responsible for tuning is the positive temperature coefficient of material stiffness for silica. The tunability however does appear to depend on the specific mechanical and optical modes that participate in any given oscillation. It must be noted that temperature change also affects the optical modes.

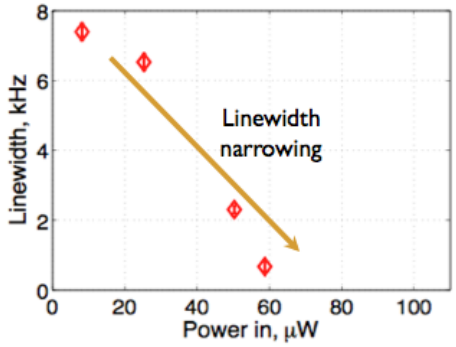


Fig. 6. Oscillation linewidth narrows with increasing optical pump power, consistent with the narrowing of laser linewidth.

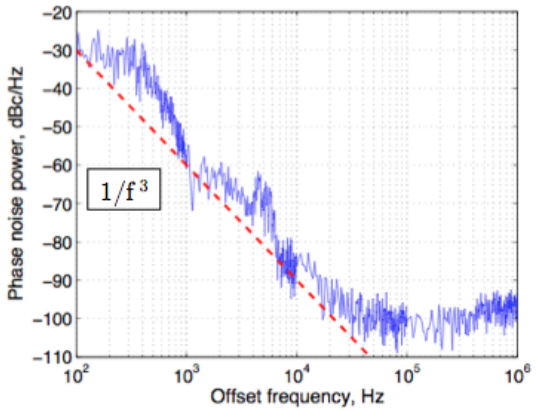


Fig. 7. Initial phase noise measurements indicate a $1/f^3$ slope. The phase noise measurement from a 94 MHz oscillation at -35 dBm signal power is pictured here.

B. Noise performance

We have performed several experiments studying the noise performance of these oscillators. Our data shows narrowing of the oscillation linewidth as a function of increasing pump optical power (Figure 6). The linewidth narrowing is consistent

with lasing processes. This line narrowing behavior should continue until the system reaches the Schawlow-Townes limit. We also have preliminary phase noise measurements which indicate a $1/f^3$ characteristic close to carrier (Figure 7).

V. Conclusions

We have experimentally demonstrated a class of optomechanical oscillators that excite surface acoustic waves on silica microspheres through electrostrictive pressure generated by light. While SAW are typically generated on planar piezoelectric substrates, with metal electrodes used for applying electric fields, here we apply the electric field through a set of virtual electrodes formed by light which co-propagates with the traveling acoustical mode. Additionally, the excitation mechanism of electrostriction is available in any dielectric material, opening up broad possibilities in material selection. Since SAW-based sensors are well-studied, the potential applications of this technology extend beyond oscillators into various sensing techniques.

References

[1] T. Carmon, H. Rokhsari, L. Yang, T. Kippenberg, and K. Vahala, "Temporal behavior of radiation-pressure-induced vibrations of an optical microcavity phonon mode," *Phys. Rev. Lett.*, vol. 94, no. 22, p. 223902, 2005.

[2] H. Rokhsari, T. Kippenberg, T. Carmon, and K. Vahala, "Radiation-pressure-driven micro-mechanical oscillator," *Opt. Express*, vol. 13, no. 14, pp. 5293–5301, Jul 2005.

[3] T. J. Kippenberg, H. Rokhsari, T. Carmon, A. Scherer, and K. J. Vahala, "Analysis of radiation-pressure induced mechanical oscillation of an optical microcavity," *Phys. Rev. Lett.*, vol. 95, no. 3, p. 033901, Jul 2005.

[4] T. Carmon and K. J. Vahala, "Modal Spectroscopy of Optoexcited Vibrations of a Micron-Scale On-Chip Resonator at Greater than 1 GHz Frequency," *Phys. Rev. Lett.*, vol. 98, no. 12, p. 123901, Mar 2007.

[5] J. Rosenberg, Q. Lin, and O. Painter, "Static and dynamic wavelength routing via the gradient optical force," *Nature Photonics*, vol. 3, pp. 478 – 483, July 2009.

[6] R. W. Boyd, *Nonlinear Optics*, 3rd ed. Elsevier, 2008, Chapter 9.

[7] L. Rayleigh, "On waves propagated along the plane surface of an elastic solid," *Proc. R. Soc. Lond.*, vol. A17, pp. 4 – 11, 1885.

[8] A. B. Matsko, A. A. Savchenkov, V. S. Ilchenko, D. Seidel, and L. Maleki, "Optomechanics with surface-acoustic-wave whispering-gallery modes," *Phys. Rev. Lett.*, vol. 103, no. 25, p. 257403, Dec 2009.

[9] H. A. Haus and W. Huang, "Coupled-mode theory," *Proceedings of the IEEE*, vol. 79, no. 10, pp. 1505 –1518, Oct. 1991.

[10] M. Gorodetsky and V. Ilchenko, "High-Q optical whispering-gallery microresonators: precession approach for spherical mode analysis and emission patterns with prism couplers," *Optics Communications*, vol. 113, no. 1-3, pp. 133 – 143, 1994.

[11] M. Cai and K. Vahala, "Highly efficient hybrid fiber taper coupled microsphere laser," *Opt. Lett.*, vol. 26, no. 12, pp. 884–886, Jun 2001.

[12] S. Spillane, T. Kippenberg, O. Painter, and K. Vahala, "Ideality in a Fiber-Taper-Coupled Microresonator System for Application to Cavity Quantum Electrodynamics," *Phys. Rev. Lett.*, vol. 91, no. 4, p. 043902, Jul 2003.

[13] J. C. Knight, G. Cheung, F. Jacques, and T. A. Birks, "Phase-matched excitation of whispering-gallery-mode resonances by a fiber taper," *Optics Letters*, vol. 22, no. 15, pp. 1129–1131, 1997.

[14] I. S. Grudinin, A. B. Matsko, and L. Maleki, "Brillouin lasing with a CaF_2 whispering gallery mode resonator," *Phys. Rev. Lett.*, vol. 102, no. 4, p. 043902, Jan 2009.

[15] R. Y. Chiao, C. H. Townes, and B. P. Stoicheff, "Stimulated brillouin scattering and coherent generation of intense hypersonic waves," *Phys. Rev. Lett.*, vol. 12, no. 21, pp. 592–595, May 1964.

[16] R. Shelby, M. Levenson, and P. Bayer, "Resolved forward Brillouin scattering in optical fibers," *Physical review letters*, vol. 54, no. 9, pp. 939–942, 1985.

New strategies for fiber-based femtosecond lasers low-noise microwave generation

W. Zhang, A. Haboucha, T. Li, S. Seidelin[§], A. Luiten*, R. Holzwarth[+], M. Lours, Y. Le Coq, G. Santarelli

LNE-SYRTE, Observatoire de Paris, CNRS, UPMC, 75014 Paris, France
[§] Institut Néel, CNRS, Université Joseph Fourier, 38042 Grenoble, France
* School of Physics, University of Western Australia, Crawley 6009, Australia
[+]MenloSytems GmbH D-82152 Martinsried, Germany

Email: yann.lecoq@obspm.fr

Low-phase-noise and stable microwave signals are of prime importance in a variety of scientific and technological fields, such as, for example, atomic frequency standards[1,2], radar and remote sensing, and very long baseline interferometry. The combination of ultra-stable lasers [3] and low noise optical frequency division by use a femtosecond laser presents a possibility to realize extremely low phase noise microwave source with reliable and operational-level technology. We have previously reported phase noise levels at -120 dBc/Hz at 1Hz from a 12 GHz carrier, with white phase noise plateau approaching -140dBc/Hz [4,5,6].

We present here new strategies to improve the phase noise of our generated microwave signal and increase reliability of the system. A 250MHz repetition rate fiber-based femto-second laser system with an Electro-Optic Modulator inside its cavity is realized at a quasi-industrial level, in close collaboration with RD teams from MenloSystems GmbH. This system was characterized to allow about 500 kHz control bandwidth with moderate amplitude modulation cross-talk. Using this unique feature, we were able to reach the ultra-stable regime, where the spectral linewidth of the optical comb is limited by that of the continuous laser reference it's locked to, instead of its intrinsic noise properties. This very large bandwidth improves significantly our capability to generate low-noise microwave signals with a scheme simpler than previously reported. Incidently, it also renders direct optics-to-optics comparission simpler and more stable.

We have also improved significantly the relative intensity noise (RIN) of the fiber-based optical frequency comb, and studied extensively how to decrease its impact on low-phase noise microwave generation via photo-detection. A thorough characterization of the amplitude-phase conversion factor of several commonly used photodetector has been realized (see [7] for more details). Amplitude-phase conversion factor below 0.03 rad per relative intensity change can be realized reliably for a given set of operating conditions when generating 12 GHz microwave signals. The currently achieved intrinsic RIN (below -135 dBc/Hz in the 10Hz-100kHz Fourier frequency

[1] J. Millo et al. Appl. Phys. Lett. **94**, 141105 (2009)

[2] S. Weyers et al. Phys. Rev A, **79**, 031803R (2009)

[3] J. Millo et al. Phys. Rev. A 79, 053829 (2009)

[4] J. Millo et al. Opt. Lett, **34**, 3707 (2009)

[5] W. Zhang et al. Appl. Phys. Lett. **96**, 211105 (2010)

[6] W. Zhang et al. arXiv:1012.0852, accepted in IEEE Trans. UFFC

[7] W. Zhang et al. arXiv:1104.4495, submitted to App. Phys. B

range) [8] of our laser leads us to state with confidence than the residual amplitude phase conversion is a negligible compared to other sources of excess noise, in particular shot noise which arise from limited microwave signal to noise ratio.

To finish, and in order to improve the shot noise limit of our system, we have implemented experimentally a repetition rate multiplication technique [9] which allows us to go from 250 MHz to 4 GHz repetition rate with a simple and compact device based on a ladder Mach-Zender Interferometr topology. This multiplication is realized with minute degradation in low Fourier frequencies phase noise spectrum as we confirmend experimentally. The resulting increase in useful signal at 12 GHz (up to -4 dBm obtained with a High Linearity Photodiode from Discovery Semiconductor) allows us to obtain a white residual phase noise plateau at ~156dBc/Hz (for a very preliminary result), at the very best level of performance ever achieved for microwave generation with femtosecond frequency combs.

These three techniques put together pave the way to a very simple and compact microwave generator with ultra-low phase noise on a very broad part of the Fourier frequencies spectrum. We can therefore envision in a close future to generate a microwave signal at 12 GHz with a phase noise at -120 dBc/Hz flicker noise at 1Hz from the carrier, $10^{-16}\tau^{-1}$ Allan deviation (with $\tau>1$ the number of averaging seconds) and a white phase noise plateau approaching -160dBc/Hz.

[8] W. Zhang et al. arXiv:1104.4495, submitted to App. Phys. B

[9] D. Diddams et al. Opt. Exp. **17**, 3331 (2009)

Optical Scattering Induced Noise in RF-Photonic Systems

O. Okusaga, J. Cahill, and W. Zhou
Microphotonics Branch,
U. S. Army Research Laboratory,
Adelphi, MD, U.S.A
Olukayode.okusaga@us.army.mil

A. Docherty, G. M. Carter, and C. R. Menyuk
Electrical Engineering Department,
University of Maryland Baltimore County
Baltimore, MD, U.S.A

***Abstract*—Scattering mechanisms in optical fiber induce noise in the spectra of optical signals travelling through the fiber. This fiber-induced noise limits the quality factor of optical fiber resonators and degrades low-noise signals transmitted via optical fiber. In this work, we present preliminary data from our systematic study of optical scattering in fiber. We present noise spectra for Rayleigh and Brillouin scattering. We also demonstrate suppression of these scattering mechanisms via laser frequency modulation.**

I. INTRODUCTION

Length dependent noise in optical fiber limits the performance of radio frequency (RF) photonic systems [1,2]. The sources of this optical-fiber-induced noise must be indentified and mitigated in order to optimize RF-photonic systems. Optical scattering represents one class of fiber-length-dependent mechanisms that can induce noise in optical signals transmitted via fiber. Interactions between incident light and density fluctuations in optical fiber can alter the wave vector and frequency of portions of the incident light. These interactions alter the phase delay and loss experienced by light in the fiber in a random fashion. These random variations in phase and amplitude amount to noise imprinted on the incident optical signal.

In this work, we present preliminary data from our systematic study of optical scattering in fiber. The data we present exhibits evidence of both Brillouin and Rayleigh scattering in fiber. Our data demonstrates that both mechanisms can contribute to the noise induced on optical signals transmitted via fiber. Furthermore, we demonstrate that this optical noise converts to electrical noise after photodetection, thereby increasing the noise of any RF signals imprinted on the optical beam. Finally, we demonstrate the suppression of both Brillouin and Rayleigh scattering by frequency modulating the incident optical signal.

II. OPTICAL SCATTERING MECHANISMS

A. Rayleigh Scattering

Rayleigh scattering can be described as scattering of incident light off of stationary density fluctuations in the optical fiber [3]. Whenever light approaches the boundary between two regions with different indices of refraction, a portion of the light is reflected off of the interface. Since the dimensions of these density fluctuations are on the order of the wavelength of the incident light, the light can be scattered in many directions. Since these density fluctuations are located randomly along the length of the fiber, the induced scattering generates phase and amplitude noise.

The density fluctuations can themselves be induced by the incident light via electrostriction leading to stimulated Rayleigh scattering (STRS) [4]. The limited data in the literature on Rayleigh scattering in optical fiber suggests that the Rayleigh gain bandwidth is on the order of 10 kHz [5]. In addition, because the density fluctuations are nearly stationary, Rayleigh scattering does not induce a significant frequency shift in the scattered light. The scattering data we present at power levels below 10 dBm are consistent with the above characteristics of Rayleigh scattering.

B. Brillouin Scattering

Brillouin scattering is caused by the interaction between the incident light and phonons within the fiber [6]. These phonons can be thought of as moving density fluctuations as opposed to the stationary density fluctuations that cause Rayleigh scattering. Because phonons have a finite energy, the interaction induces a frequency shift in the scattered light. This frequency shift is approximately 11 GHz in optical fiber. Due to the conservation of both energy and momentum, the finite phonon momentum causes Brillouin scattering to occur primarily in the counter-propagating direction.

At approximately 10 MHz, the Brillouin scattering bandwidth is significantly greater than that of Rayleigh scattering. As in the case of Rayleigh scattering, the incident light can amplify phonons via electrostriction thereby leading to stimulated Brillouin scattering (SBS). The data we present in this work shows that SBS requires significantly greater power levels than STRS. This result is consistent with the literature [5].

This work was funded by the U.S. Defense Department Coalition Warfare Program.

C. *Backscattered versus transmitted spectra*

Our ultimate goal is to determine the effects of both Rayleigh and Brillouin scattering on the spectrum of the light transmitted through a length of optical fiber. It is this transmitted spectrum that is converted into electrical phase noise in RF-photonic systems. However, we first characterize both scattering mechanisms by measuring the spectrum of the backwards scattered light. Both Rayleigh and Brillouin scattering contribute to the backward scattered light whereas most other fiber length dependent noise mechanisms do not. Therefore, the backward scattered spectrum allows us to analyze the effects of both scattering mechanisms separately from other noise phenomena such as Kerr interaction or dispersion.

III. Optical Domain Measurements

We began our study with broadband optical domain measurements. These measurements were performed with an optical spectrum analyzer with a bandwidth greater than 300 nm at 1550 nm. Performing broadband measurements allowed us to determine the spectral regions in which the backscattered light was concentrated.

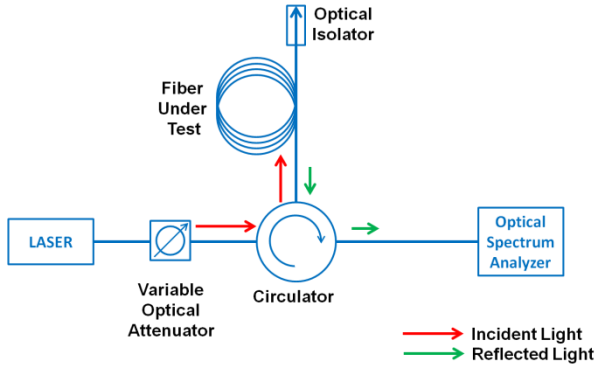


Figure 1. A schematic diagram of our optical-domain backscatter measuremets.

A. *Experimental Setup*

Fig. 1 shows a schematic diagram of our optical-domain backscatter measurement setup. We used a 100 mW distributed feedback (DFB) semiconductor diode laser. The variable optical attenuator (VOA) was used to alter the optical power into the fiber without changing the relative power spectrum of the incident light. The circulator separated the incident light from the backscattered light. The latter was sent to our optical spectrum analyzer.

B. *Experimental Results*

Fig. 2 shows the optical spectra of the back-scattered light from a 6 km spool of single-mode fiber with varying input power levels. The data highlights three regions of interest: the spectral region around the wavelength of the incident light, and the regions approximately 0.1 nm above and below the wavelength of the incident light. The latter two regions correspond to the anti-Stokes and Stokes wavelengths induced by Brillouin scattering [6]. The increased optical power measured in the Stokes region for input power levels above 11 dBm is due to SBS. However, the bandwidth of the Stokes spectrum shown in Fig. 2 is limited by the resolution bandwidth of the optical spectrum analyzer (0.01 nm). We require higher resolution measurements to properly resolve the SBS bandwidth and determine the effects of Rayleigh scattering.

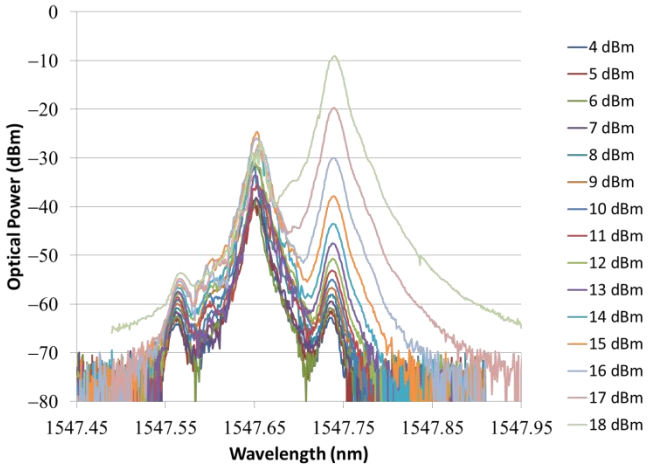


Figure 2. A plot of the optical spectra of the backscattered light from a 6 km spool of single-mode fiber with varying input power levels.

IV. Electrical Domain Measurements

In order to gain greater resolution, we performed electrical-domain measurements on the backscattered optical spectrum. By down-converting the backscattered signal into the electrical domain, we sacrifice bandwidth but increase spectral resolution. Since we had already determined the spectral regions of interest using our optical domain measurements, we could perform narrow-band high-resolution measurements around these regions.

A. *Experimental Setup*

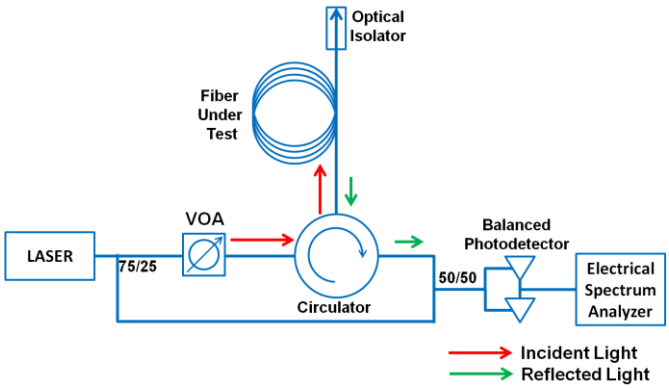


Figure 3. A schematic diagram of our electrical-domain coherent backscatter measuremets

We use the same laser, VOA, and circulator used in the optical-domain measurements. We use photodetectors to convert the backscattered light into electrical currents which we then transmit to a high-resolution electrical spectrum analyzer (ESA). However, the backscattered optical power can be as low as –60 dBm. At such low power levels, the photodetector output is dominated by shot noise. To increase the sensitivity of our electrical domain measurement, we used a coherent detection scheme. We split off 25% of the laser

power and sent it into one input of a 2-by-2 optical combiner. This portion of the laser power was used as a local oscillator (LO) signal. The backscattered light from the circulator was sent into the other input of the combiner. The combined optical signal was then sent to a pair of balanced detectors. The output current from the balanced detector pair was proportional to the product of the backscattered and LO field amplitudes. In other words, we used the LO signal to coherently amplify the backscattered signal above the shot noise level of the photodetectors.

B. Experimental Results

In the electrical domain, the spectral regions of interest are: the region around the Stokes-frequency shift (approximately 11 GHz), and the region around DC where which corresponds to the region around the incident light central frequency in the optical domain.

1) Brillouin Scattering Data

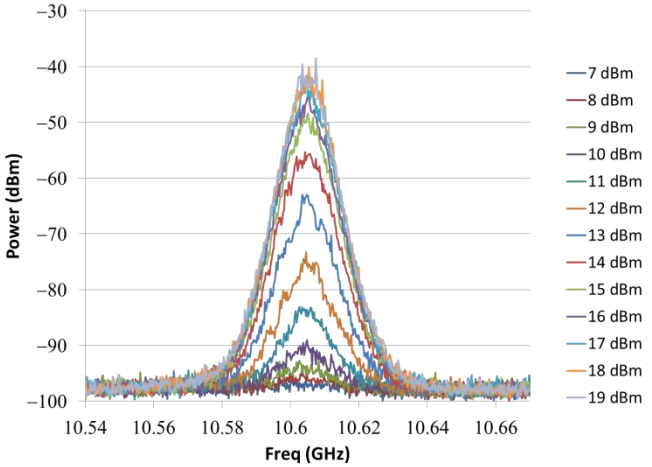


Figure 4. A plot of the backscattered power spectra from a 6 km spool of single-mode fiber measured in the electrical domain for varying input power levels. These spectra was measured in the frequency range around the Brillouin-induced Stokes shift.

Fig. 4 shows a plot of the backscattered power from a 6 km spool of single-mode fiber measured in the electrical domain. The peak backscattered power for all input power levels was approximately 10.61 GHz. The bandwidth of the Brillouin scattered spectra was on the order of 10 MHz. We were able to measure Stokes-shifted signals above the shot noise level for input power levels above 8 dBm. Due to the limited bandwidth of the SBS spectrum, we expect that the Brillouin scattered light will only directly interfere with RF-signals with frequencies within 10-20 MHz of the 10.6 GHz Stokes frequency.

2) Rayleigh Scattering Data

To investigate the Rayleigh spectra, we measured the spectrum of the electrical signal from the photodetectors around DC. In particular, we focused on the range of frequencies between 10 Hz and 1 MHz. This range corresponds to the frequency range shifted between 10 Hz and 1 MHz from the central frequency of the incident optical beam.

Fig. 5 shows a plot of the backscattered power spectra from a 6 km spool of single-mode fiber. We measure significant levels of power-dependent noise for incident power levels as low as –10 dBm.

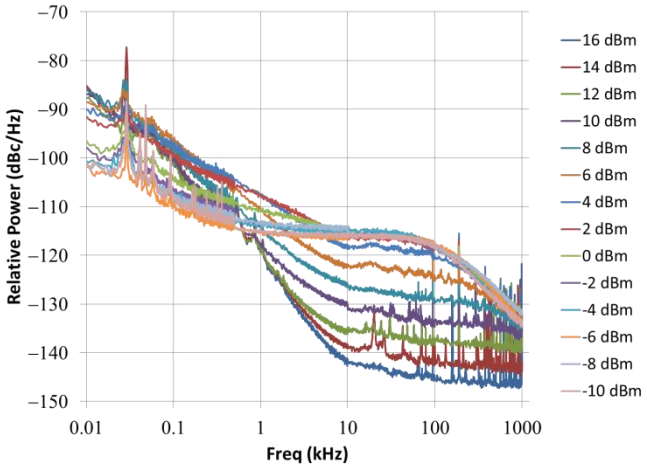


Figure 5. A plot of the backscattered power spectra from a 6 km spool of single-mode fiber measured in the electrical domain for varying input power levels. These spectra were measured in the frequency range around DC.

At offset frequencies less than 10 kHz, the relative noise power increases with increasing input power. However, for offset frequencies greater than 10 kHz, the relative noise power decreases with increasing optical power. This gain filtering suggests a Rayleigh scattering bandwidth of approximately 10 kHz. This result is consistent with the literature on Rayleigh scattering in optical fiber [5].

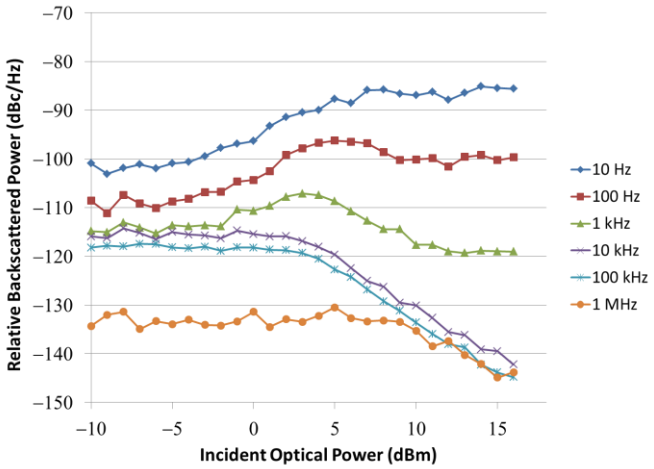


Figure 6. A plot of the relative backscattered power versus incident optical power for varying offset frequencies.

Fig. 6 shows a plot of relative backscattered power versus incident optical power for varying offset frequencies. These level curves highlight the gain filtering induced by STRS. At 10 and 100 Hz offset from the central frequency of the incident light, the scattered power increases with the incident power until approximately 8 dBm. At incident power levels greater than 8 dBm, the relative backscattered power spectral density remains fixed at approximately –85 dBc/Hz at 10 Hz

offset frequency. It is at this power range that we begin to observe significant Brillouin scattering. This data suggests that at power levels above 8 dBm, SBS draws power away from the Rayleigh scattering process, thereby clamping the STRS gain.

V. Transmission Spectra

Having characterized both Brillouin and Rayleigh scattering in our fiber spools, we then measured the spectra of the transmitted light through the same spools. We focused on the region between 10 Hz and 1 MHz from the central frequency of the incident light. This region will be converted to electrical noise at the photodetector regardless of the RF frequency of the specific RF-photonic system. The Stokes frequency range, on the other hand, will only be relevant for RF-photonic systems operating at or near 10. 6 GHz.

A. Experimental Setup

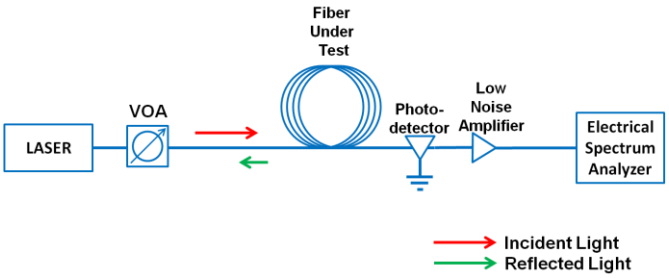


Figure 7. A schematic diagram of our electrical-domain transmitted spectrum measurement setup.

Fig. 7 shows a schematic diagram of our transmitted spectrum measurement setup. We use the same laser and VOA used in our previous measurements. However, we remove the circulator and measure the transmitted optical power from the fiber under test. After the photodetector, a low noise amplifier (LNA) is used to increase the sensitivity of our measurement system. The LNA is AC-coupled with a 0.2 Hz corner frequency. This AC-coupling effectively filters out the DC current from the incident light beam. Suppressing this DC current allows us to measure the much lower noise levels in the range between 10 Hz and 1 MHz around the central carrier.

B. Experimental Data

We varied the optical power into the fiber from –4 dBm to 17 dBm. We found two qualitative regimes: a low-power regime where the noise spectra were dominated by Rayleigh scattering; and a high-power regime where Brillouin scattering affected the noise spectra. We present experimental data for these two regimes separately in the following subsections.

1) Rayleigh Scattering Regime

Fig. 8 shows a plot of transmitted power measured in the electrical domain from a 6 km spool of single-mode fiber. The input power levels ranged from –4 to 11 dBm. For input power levels below 11 dBm, the spectra resemble the Rayleigh backscatter data we presented in previous sections. We observe the same gain-filtered noise spectrum with a 10 kHz bandwidth. This is consistent with our initial hypothesis that Rayleigh scattering in optical fiber is largely isotropic, producing scattered light in both the forward and backward directions.

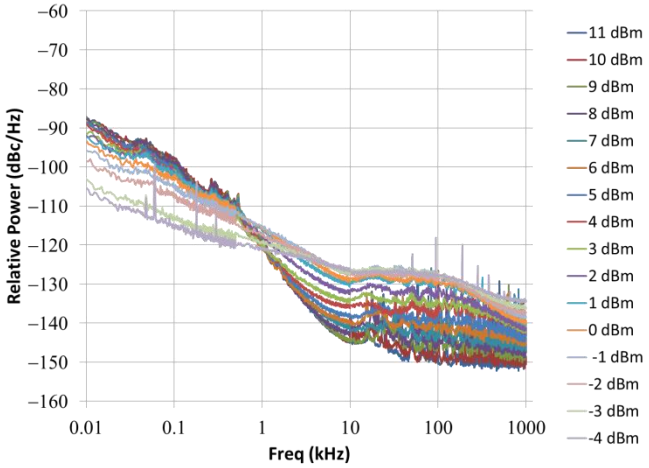


Figure 8. A plot of transmitted power spectra through a 6 km single-mode fiber spool for varying input power levels. This input power range from –4 to 11 dBm is dominted by stimulated Rayleigh scattering.

2) Brillouin Scattering Regime

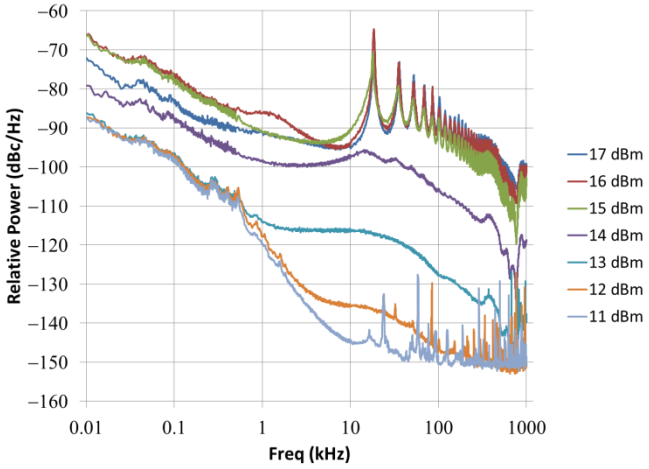


Figure 9. A plot of transmitted power spectra through a 6 km singl-mode fiber spool for varying input power levels. This input power range from 11 to 17 dBm is dominated by stimulated Brillouin scattering.

Fig. 9 shows a plot of the transmitted power measured in the electrical domain from a 6 km spool of single-mode fiber. The input power levels ranged from 11 to 17 dBm. For this range of input power levels, we found that the spectra deviate significantly form backscattered spectra for the same spool. We find that for offset frequencies beyond the Rayleigh scattering bandwidth, the transmitted noise increases with increasing optical power. This increased noise can be explained by pump depletion due to SBS. The SBS effect randomly removes power from the incident light and transfers it to the Stokes beam. This random power modulation results in an increase in the noise spectra of the incident light beam.

At power levels above 13 dBm, we observe Fresnel fringes in the transmitted spectra. The frequency spacing of these Fresnel fringes (~17 kHz) corresponds to a cavity length equal twice the length of the fiber spool under test. These Fresnel

fringes can be explained by double Brillouin scattering. As the incident light travels down the length of the fiber, SBS generates a Stokes beam that travels in the opposite direction. This backscattered Stokes beam in turn generates an anti-Stokes beam with the same frequency as the original incident light beam. This Stokes-induced anti-Stokes beam then reflects off connectors at the other end of the fiber and travels back down to the photodetector. At the detector, the double-scattered beam interferes with the original incident beam to generate Fresnel fringes. Our data demonstrates that, despite the Stokes frequency shift, SBS can contribute indirectly to the noise in the frequency range close to the original incident optical beam's central frequency.

VI. Scattering Suppression

Our ultimate goal is to characterize fiber-length-dependent noise in RF-photonic systems and investigate means of suppressing this noise. It is well known that length dependent noise is present in optoelectronic oscillators (OEOs) [1]. Maleki et al. have demonstrated that frequency modulation of the laser in an OEO can suppress some of this fiber-length dependent noise [7]. Fig. 10 shows the phase noise from a 6 km OEO with and without frequency modulation. The data shows that frequency modulation can suppress the phase noise of the OEO by more than 10 dB.

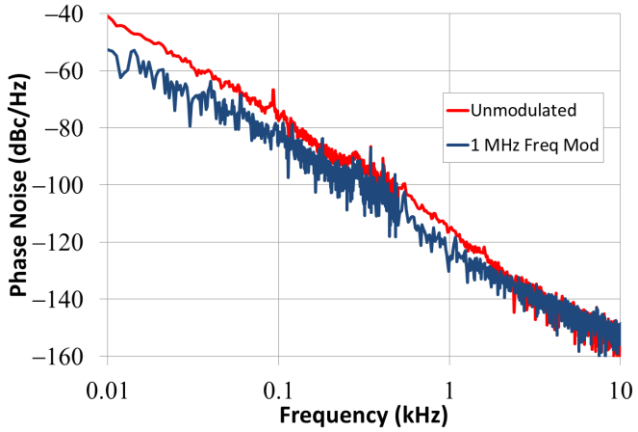


Figure 10. A plot of the phase noise from a 6 km OEO with and without laser frequency modulation.

If the fiber-length-dependent noise observed in the OEO is caused in part by either Brillouin or Rayleigh scattering, we would expect that laser frequency modulation would also suppress the Rayleigh or Brillouin scattered power.

A. Brillouin Scattering Suppression

To determine the effect of frequency modulation on the Brillouin scattered spectrum, we repeat our optical-domain measurements while modulating the incident light into the fiber. The incident light is modulated in two ways. We modulate the drive current into the laser. Doing so modulates both the amplitude and the frequency of the laser signal. To separate the effect of laser amplitude modulation from frequency modulation, we also amplitude modulate the laser output using an external amplitude modulator.

Fig. 11 shows the peak optical Stokes power (relative to the input power) versus input optical power for three cases: no modulation, external amplitude modulation, and direct frequency modulation via modulation of the laser drive current. The data shows that frequency modulation is the most effect means of suppressing SBS. In both cases, a 10 MHz RF modulation tone with 10 mW of power was used. Frequency modulating the laser led to a SBS suppression factor of 50 dB, whereas amplitude modulation only reduced the SBS gain by 15 dB.

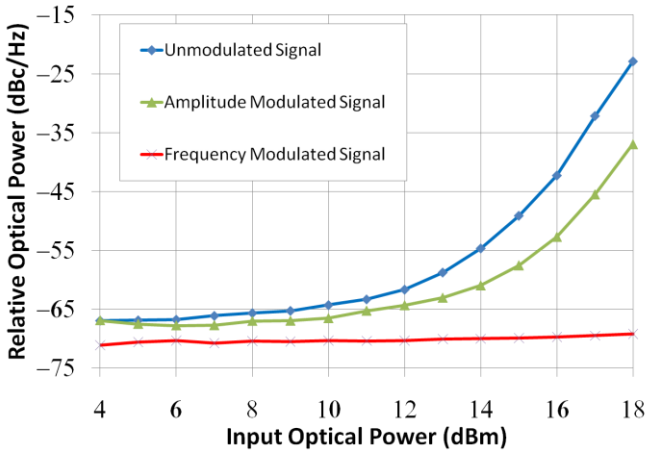


Figure 11. A plot of the peak Stoke power from a 6 km spool versus input optical power for various modulation schemes.

B. Rayleigh Scattering Suppression

To determine the effect of laser frequency modulation we measure the backscattered power in the electrical domain at 10 Hz offset frequency for varying input power levels with and without laser frequency modulation. This measurement allows us to estimate the peak Rayleigh scattered noise power versus input power in a 6 km single-mode fiber spool.

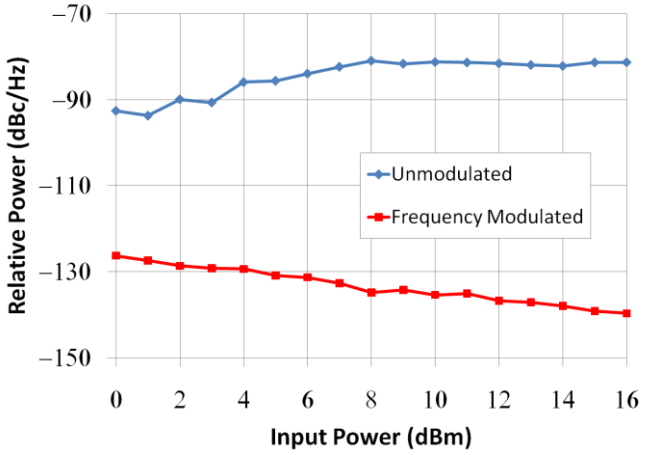


Figure 12. A plot of the peak Rayleigh-scattered power from a 6 km spool versus input optical power with and without laser frequency modulation.

Fig. 12 shows a plot of the peak Rayleigh-scattered power form a 6 km spool of single-mode fiber versus input optical power with and without laser frequency modulation. The data shows that frequency modulation suppresses the Rayleigh-scattered power by more than 50 dB. Like SBS, STRS is

suppressed by frequency modulation of the incident optical beam.

VII. Conslusions

In this work, we present data demonstrating the effects of Brillouin and Rayleigh scattering on both the backscattered and transmitted spectra of light incident on a length of single-mode optical fiber. Our data shows that both Rayleigh and Brillouin scattering contribute to the noise within 1 MHz of the central frequency of the transmitted optical beam. This scattering-induced optical noise is converted to RF noise after photodetection. Our data shows that Rayleigh scattering dominates the scattering-induced noise at input power levels below 11 dBm. Stimulated Brillouin scattering dominates the noise at input power levels above 11 dBm. At input power levels above 13 dBm, we observe evidence of noise induced by double-Brillouin scattering in a 6 km fiber spool. Finally, we show that laser frequency modulation reduces the magnitude of both Rayleigh and Brillouin-induced noise in optical fiber.

References

[1] O. Okusaga, W. Zhou, E. Levy, M. Horowitz, G. Carter, and C. Menyuk, "Non-ideal loop-length-dependence of phase noise in OEOs," in Conference on Lasers and Electro-Optics/International Quantum Electronics Conference, OSA Technical Digest (CD) (Optical Society of America, 2009), paper CFB3.

[2] S. M. Foreman, K. W. Holman, D. D. Hudson, D. J. Jones, and J. Ye, "Remote transfer of ultrastable frequency references via fiber networks," Rev. Sci. Instrum. **78**, 021101 (2007).

[3] M. Froggatt and J. Moore, "High-Spatial-Resolution Distributed Strain Measurement in Optical Fiber with Rayleigh Scatter," Appl. Opt. **37**, 1735-1740 (1998).

[4] A. Melloni, M. Frasca, A. Garavaglia, A. Tonini, and M. Martinelli, "Direct measurement of electrostriction in optical fibers," Opt. Lett. **23**, 691-693 (1998).

[5] T. Zhu, X. Bao, L. Chen, H. Liang, and Y. Dong, "Experimental study on stimulated Rayleigh scattering in optical fibers," Opt. Express **18**, 22958-22963 (2010).

[6] R. Jenkins, R. Sova, and R. Joseph, "Steady-State Noise Analysis of Spontaneous and Stimulated Brillouin Scattering in Optical Fibers," J. Lightwave Technol. **25**, 763-770 (2007).

[7] X. Yao, L. Davis, and L. Maleki, "Coupled Optoelectronic Oscillators for Generating Both RF Signal and Optical Pulses," J. Lightwave Technol. **18**, 73- (2000).

Theoretical Investigation of Optical Fiber-Length-Dependent Phase Noise in Opto-Electronic Oscillators

Andrew Docherty*, Olukayode Okusaga†, Curtis R. Menyuk*, Weimin Zhou†, Gary M. Carter*
*University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250
†U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD 20783

***Abstract*— Optical fiber-length-dependent flicker phase noise has been identified as a major limit to the reduction of phase noise in opto-electronic oscillators (OEOs). We outline a systematic approach to model the optical path of the OEO, and investigate the sources of length-dependent phase noise in the optical path and how they are effected by the dispersion and Kerr nonlinearity of the optical fiber.**

I. Introduction

Optical fiber-length-dependent flicker phase noise is a major limit in opto-electronic oscillators (OEOs) and other RF photonic devices [1]. Identifying the sources of this effect and mitigating them is important to increase the performance of RF photonic systems using long fiber lengths. A number of different phenomena can in principle contribute to this effect, these include: the conversion of laser frequency noise into RF phase noise via fiber dispersion [2]; the conversion of laser amplitude noise into RF phase noise via third order dispersion and Kerr nonlinearity; parametric amplification of the laser noise, which is then converted into RF phase noise in the photodetector or RF amplifiers; and the conversion of laser noise into RF phase noise via scattering processes in the optical fiber, including Brillouin and Rayleigh scattering and scattering from end-faces and connectors [3].

Due to the increase in the system Q as the length of the optical fiber is increased, the RF phase noise of an OEO should decrease with increasing fiber length; however, in experimental systems this expected decrease is not observed at lower frequencies. Fig. 1 shows the experimentally measured RF phase noise in an OEO with three different fiber lengths. While the phase noise around 10 kHz decreases with increasing length, the phase noise below 1 kHz is considerably higher than that the expected phase noise (shown as the dashed lines); indeed, in the measured system at frequencies around 10 Hz there was essentially no improvement in the phase noise when the fiber length was increased by an order of magnitude, from 500 m to 6 km. This has been found to be consistent with a flicker noise source that increases as the fiber length is increased [5].

In order to understand the sources of this length-dependent flicker noise and their relative effect on the RF phase noise in the OEO we model the optical transfer of both signal and noise through the optical path, including the effects of laser noise, fiber loss, dispersion, and Kerr nonlinearity. We show that for physically realistic noise powers the effects of laser amplitude noise amplification and conversion to RF phase noise are much smaller than the white noise floor of the system and are not candidates for sources of length-dependent flicker noise.

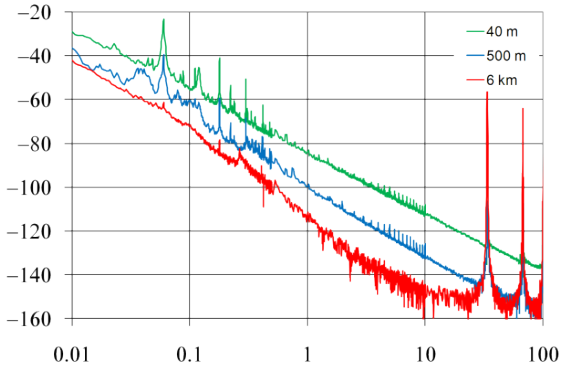


Fig. 1. The experimentally measured phase noise in an OEO (solid lines) and the theoretically expected phase noise (dashed lines) for three different lengths of fiber illustrating the length-dependent flicker noise in the OEO [4].

II. Modeling the Optical Path

A typical OEO system configuration is illustrated in Fig. 2 which shows the evolution of the signal and the noise sources in both the optical and electronic domains. These noise sources include the shot noise of the photodetector, the thermal and flicker noise in the amplifier, the laser noise, and spontaneous scattering noise induced in the optical fiber.

The RF signal undergoes a complex nonlinear evolution as it travels through the optical path from point (1) in Fig. 2 where it is AM modulated onto the optical carrier (2) by the electro-optic modulator (EOM) and travels through the optical fiber (3). The photodetector (4) converts the optical signal to the RF domain again, and mixes the multiple harmonics of the RF signal around the optical carrier, giving multiples of the RF signal frequency, which are then amplified and filtered in the RF domain to produce a single RF frequency, and lastly fed back into the modulator.

A. The electo-optic modulator

Firstly, the electro-optic modulator imprints the electronic input signal onto the laser output. The generic expression for

978-1-61284-110-6/11/$26.00 ©2011 IEEE

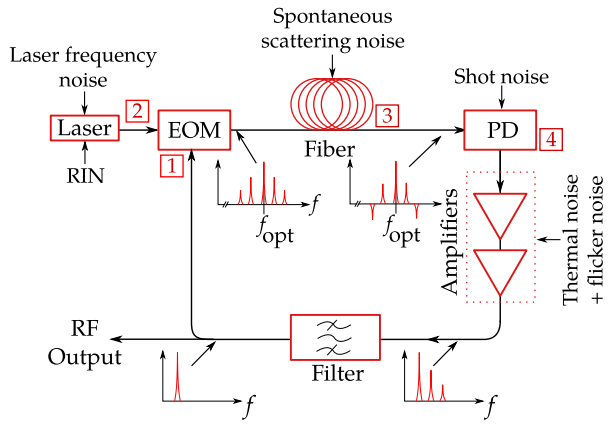


Fig. 2. The OEO system showing the sources of noise and the harmonics of the RF signal at different points in the loop.

the electric field at the output of the modulator is given by the equation,

$$E_{\text{opt}}(t) = \frac{1}{2}E_{\text{laser}}(t)\{\eta_1 \exp[jv_1 A_{\text{in}}(t)] + \eta_2 \exp[jv_2 A_{\text{in}}(t) + j\psi]\}, \quad (1)$$

where $A_{\text{in}}(t)$ is the RF input signal, $\eta_{1,2}$ are determined by the extinction ratio, $v_{1,2}$ are determined by V_π and the modulator chirp, and ψ is determined by the bias.

Directly after modulator the optical power is given by,

$$P(t) = \frac{P_0}{4}\left(\eta_1^2 + \eta_2^2 + 2\eta_1\eta_2 \cos[(v_2 - v_1)A_{in}(t) + \psi]\right), \quad (2)$$

where P_0 is the laser input power.

These modulator parameters of Eq. 1 can be related to the physical device parameters the half-wave voltage V_π, the bias voltage V_B, and the modulation depth η by the following relations

$$\begin{aligned} \eta &= 2\eta_1\eta_2 \\ V_\pi &= \frac{\pi}{v_1 - v_2} \\ V_B &= \frac{V_\pi}{\pi}(\psi - \frac{\pi}{2}) \end{aligned} \quad (3)$$

with the restriction that $\eta_1^2 + \eta_2^2 = 1$.

Using these parameter definitions Eq. 2 is equal to the standard formula for optical power after the modulator that was given in Yao-Maleki [6], Furthermore, we assume the RF input to the modulator is a cosine with amplitude $V_{\text{in}}(t)$ noise and $\phi(t)$ phase noise given by

$$A_{\text{in}}(t) = V_{\text{in}}(t)\cos\big[\omega_0 t + \phi(t)\big].$$

Then we use the Jacobi-Anger expansion and write the optical signal as a multi-scale expression for the harmonics of the RF oscillator around the optical carrier,

$$E_{\text{opt}}(t) = \frac{E_{\text{laser}}}{2}\sum_{m=-\infty}^{\infty} j^m \Big[\eta_1 J_m(v_1 V_{\text{in}}) + \eta_2 J_m(v_2 V_{\text{in}})\exp(j\psi)\Big]\exp\Big\{jm[\omega_0 t + \phi(t)]\Big\}. \quad (4)$$

B. Laser phase and amplitude noise

A likely source of length dependent flicker noise comes from the laser noise converted to RF phase noise by some process in the fiber. It is therefore important to take the laser frequency and amplitude noise into account in any noise model of the OEO. The laser frequency and amplitude noise in an OEO was recently measured by Volyanskiy et al. [2]. Their measured spectra showed the laser frequency noise with a $1/f$ slope at frequencies below 1 kHz and RIN with a slope of $1/f$ to $1/f^{1.5}$ at lower frequencies. In this paper we use laser frequency noise and RIN noise power spectra that are within the range of their experimentally measured results. The experimental results also depended strongly upon the power of the laser; in the simulations we do not change the laser noise so that the effects of fiber nonlinearity can be isolated. We model the $1/f$ noise using a spectral filtering algorithm as described in Levy et al. [5].

The laser output including RIN and laser phase noise is given by,

$$E_{\text{laser}}(t) = E_0\big[1 + \alpha_{\text{RIN}}(t)\big]\exp\left[j\omega_c t + j\int_0^t \Delta\omega(t')dt'\right], \quad (5)$$

where E_0 is the field amplitude of the laser, ω_c is the optical carrier frequency, $\alpha_{\text{RIN}}(t)$ is the laser amplitude noise, and $\Delta\omega(t)$ is the laser frequency noise.

C. The optical fiber

The signal from the modulator is directly coupled into an optical fiber. We model the propagation of the optical signal through the fiber including the effects of loss, dispersion, third-order dispersion, and Kerr nonlinearity using the following equation [7],

$$\frac{\partial E}{\partial z} = -\frac{\alpha}{2}E - \beta_1\frac{\partial E}{\partial t} - j\frac{\beta_2}{2}\frac{\partial^2 E}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 E}{\partial t^3}, + j\gamma|E|^2 E \quad (6)$$

where α is the fiber loss, $\beta_1 = 1/v_g$ with v_g being the group velocity, $\beta_2 = \beta''(\omega_c)$ the dispersion, $\beta_3 = \beta'''(\omega_c)$ is the third-order dispersion, and γ is the nonlinear coefficient. Here we take typical parameters for the Corning® SMF-28 fiber of $\alpha = 0.16$ dB/km, $\beta_2 = -20$ ps^2/km, $\beta_3 = 0.1$ ps^3/km, $\gamma = 0.0015$ (W m)$^{-1}$.

The laser frequency noise is inherently large at optical frequencies due to the large operating frequency of the laser. The optical phase noise is therefore extremely large, which greatly restricts the time discretization of the numerical simulations if implemented directly. For efficient simulations we do not explicitly calculate with the laser phase noise, instead we make the substitution $E(t,z) = E_{\text{laser}}(t)A(t - z/v_g, z)$; where $E_{\text{laser}}(t)$ is given by Eq. 5 and $A(t,z)$ is the propagating

envelope initially given by the modulator response, delayed by the group velocity. This leads to an equation in $A(t,z)$

$$\frac{\partial A}{\partial z} \simeq -\frac{\alpha}{2}A - j\frac{\beta_2}{2}\left[\partial_t + j\Delta\omega(t)\right]^2 A + \frac{\beta_3}{6}\left[\partial_t + j\Delta\omega(t)\right]^3 A + j\gamma P_0\left[1+\alpha_{\mathrm{RIN}}(t)\right]^2 |A|^2 A, \quad (7)$$

where we no longer need to directly model the large laser phase noise.

Furthermore, to model the frequencies close to the RF harmonics we can expand the field $A(z,t)$ near the harmonics of the RF oscillation frequency, w_0. Namely, we write

$$A(z,t) = \sum_{m=-M}^{M} a_m(z,t)\exp(jm\omega_0 t), \quad (8)$$

where M is the number of harmonics that are modeled in the optical domain. This assumption is equivalent to assuming the noise spectrum is slowly varying compared to the RF oscillation frequency. Substituting Eq. 8 into Eq. 7 we arrive at a set of coupled PDEs

$$\frac{\partial a_m}{\partial z} = -\frac{\alpha}{2}a_m - j\frac{\beta_2}{2}\left[\partial_t + j\Delta\omega(t)\right]^2 a_m + \frac{\beta_3}{6}\left[\partial_t + j\Delta\omega(t)\right]^3 a_m + j\gamma \sum_{k=-M}^{M}\sum_{l=-M}^{M} a_j a_k a^*_{j+k-m}. \quad (9)$$

The nonlinear term contains all cross-phase modulation and four wave mixing terms. The numerical solution of these coupled equations allows the noise signals close-in to the carrier to be propagated with full nonlinear dynamics.

D. Detection

The detector completes the optical path and converts the optical signal to the electronic domain. Using the notation of Eq. 8 the detected signal at the fundamental RF frequency is given by the following equation

$$V_{\mathrm{RF}}(t) = \rho R \sum_{m=-\infty}^{\infty} a_m(L,t)a^*_{m-1}(L,t)\exp(j\omega_0 t)$$

All other harmonics will be filtered in the electronic domain, so we need only consider in this fundamental RF frequency.

If the optical fiber is modeled as a perfect delay with no other effects then the RF signal reduces to the equation given by Yao and Maleki [6]

$$V_{\mathrm{RF}}^{(\mathrm{ideal})}(t) = \frac{P_0}{2}R\rho\eta\cos\left(\frac{\pi V_B}{V_\pi}\right)J_1\left(\frac{\pi V_{\mathrm{in}}}{V_\pi}\right)\cos\left[\omega_0 t + \phi(t)\right]$$

where R is the impedence. We use this ideal response of the optical path as a comparison for the results of the different optical effects.

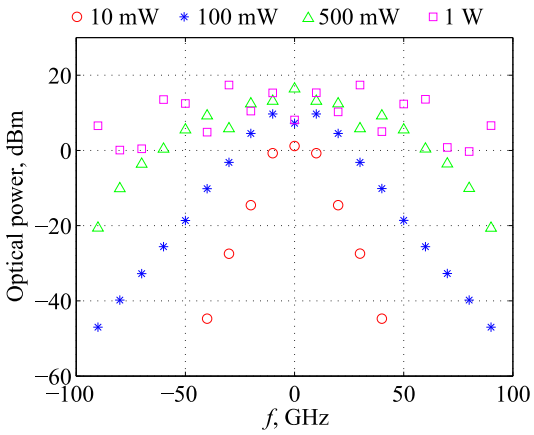


Fig. 3. Theoretical optical power in the harmonics for a 10 GHz OEO after 6 km of transmission through SMF-28 for different input laser powers.

III. The Large Signal Behavior

As a first step, it is necessary to find the stationary state of the system. Once the steady state is identified, it is possible to calculate the impact of the laser, photodetector, and RF amplifier noise on the OEO's phase noise.

In general there will be two main effects on the evolving harmonics in the optical domain: Firstly, increasing the nonlinearity will increase the power transferred to the harmonics further from the carrier in the optical domain due to spectral broadening. Secondly, the phase of the harmonics will rotate due to both dispersion and nonlinear effects leading to reduction of the detected signal, if a perfect modulator is considered.

The calculated power in the optical harmonics at the end of a 6 km optical fiber is shown in Fig. 3. As expected as the optical power is increased the power in the signal broadens in the frequency domain. For power levels below 100 mW the majority of the power is in the three lowest order harmonics; however, above this cutoff power in the higher-order harmonics increases quickly.

The phase of the harmonics is changed by dispersion and nonlinearity, if we neglect the amplitude changes just described, which is approximately valid for laser powers less than 100 mW, we can write the phasor at the end of the fiber as the input phasor rotated by a phase θ_m

$$a_m(L) = a_m(0)\exp\left[-\frac{1}{2}\alpha z + j\theta_m\right]$$

Assuming that the majority of the signal comes from the three lowest order harmonics ($m=-1,0,1$) the relative depahsing factor between the beating harmonics is given by

$$\delta = \frac{\theta_1(L)+\theta_{-1}(L)}{2} - \theta_0(L)$$

For dispersion alone this dephasing factor would be given by $\delta = \frac{\beta_2}{2}\omega_0 L$; however, in the presence of nonlinearity it must be calculated numerically. Assuming an ideal modulator any

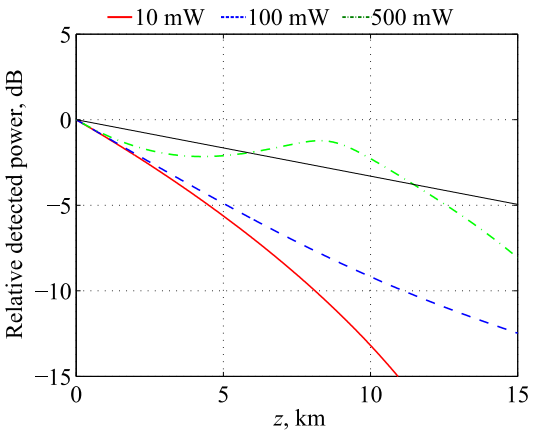


Fig. 4. Calculated detected power for a 10 GHz OEO for a modulator with zero chirp, $\eta = 0.7$, $V_\pi = 5.1$ V and $V_B = 2.6$ V. The solid black line is the expected detector power if only fiber loss was considered.

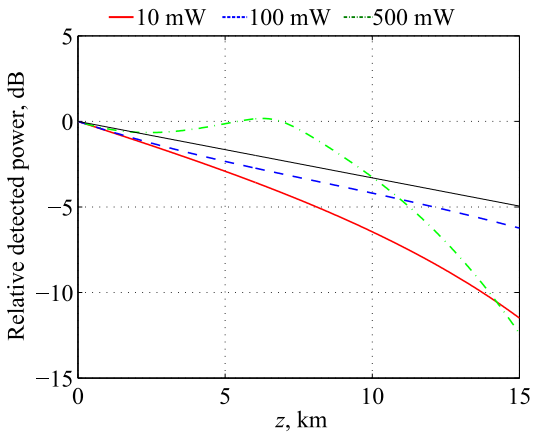


Fig. 5. Calculated detected power for a 10 GHz OEO for a modulator with a chirp of $\alpha = 0.6$, $\eta = 0.7$, $V_\pi = 5.1$ V and $V_B = 2.6$ V. The solid black line is the expected detector power if only fiber loss was considered.

dephasing would cause a decrease in the detected signal that is given by

$$V_{\mathrm{RF}}(t) = \exp(-\alpha L)\cos(\delta)V_{\mathrm{RF}}^{(\mathrm{ideal})}(t)$$

This is shown in Fig. 4 where the detected power for an unchirped modulator for different laser powers is shown. The lower power curve is considerably decreased over the case where there is only fiber loss, and increasing the power causes the Kerr nonlinearity to increasingly oppose the phase rotation due to the dispersion.

For a modulator with non-ideal extinction ratio and chirp the RF harmonics will have different phases immediately after modulation which will be further changed by optical propagation. This can increase or decrease the detected signal depending on the chirp of the modulator. Most commercial modulators have chirps that cancel the effect of anomalous dispersion; therefore having modulator chirp will increase the detected signal in the presence of optical dispersion. This is shown in Fig. 5. The detected power for low laser powers is higher than that in Fig. 4 due to the modulator chirp.

IV. RF and Optical Noise Behavior

A. Dispersion

Expanding Eq. 7 and neglecting third order dispersion, nonlinearity, higher order noise terms and assuming the laser frequency noise is a slow signal compared to the frequencies of interest we obtain

$$\frac{\partial A}{\partial z} - \beta_2 \Delta\omega(t)\frac{\partial A}{\partial t} \simeq -\frac{\alpha}{2}A - j\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} \qquad (10)$$

the left hand side can now be seen as a random timing jitter driven by the laser frequency noise, and the terms on the right-hand side of Eq. 10 give the fiber loss and a phase rotation in the frequency domain. The timing jitter of $\beta_2\Delta\omega(t)z$ can be converted to a phase shift of the RF frequency leading to a RF phase noise due to propagation through a fiber of length L of,

$$\phi_{\mathrm{RF}}(z) = \beta_2\omega_0 L\Delta\omega(t), \qquad (11)$$

which is the same equation given by Volyanskiy et al. [2].

B. Kerr nonlinearity

In recent experimental measurements the RF phase noise in an OEO using standard fiber as well as dispersion shifted fiber (DSF) were compared [4]. The experimental results failed to show the expected decrease in length-dependent flicker noise that would be expected if the conversion of laser frequency noise to RF phase noise discussed in the previous section was dominant. Furthermore, the results were highly power dependent indicating that there is a dominant nonlinear process that drives the length-dependent flicker noise. In this section we investigate the Kerr nonlinearity of the fiber and determine its role in length-dependent noise in the OEO.

Considering Kerr nonlinearity alone the optical signal only experiences a nonlinear phase rotation. After direct detection this has no effect on the noise or large signal dynamics of the OEO. However, the combination of nonlinearity and dispersion can have complex effects. In the previous section these effects were investigated with regards to the large signal behavior. The effect of nonlinearity and dispersion can also affect the noise, specifically it can cause noise mixing of noise between channels through cross-phase modulation and in addition can cause parametric amplification of noise, where power from the pump is transferred via four-wave mixing to the noise.

To investigate their effects on the noise after detection we use numerical simulations of the optical path described in the preceding sections, from modulator input to detector output.

We begin taking a system with a noise-free RF input and adding laser phase noise alone to the simulations. As discussed previously, dispersion directly converts laser frequency noise to RF phase noise after detection. This is shown in Fig. 6 the dashed lines indicate different laser powers, while assuming

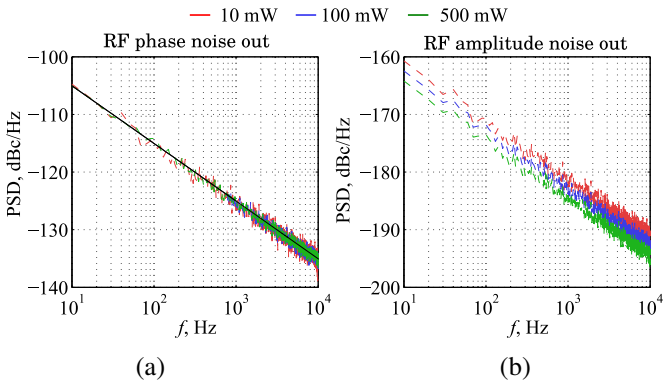


Fig. 6. The theoretical detected RF (a) frequency and (b) amplitude noise converted from a typical laser frequency noise spectrum by dispersion and nonlinearity in a 6 km length of SMF-28 fiber.

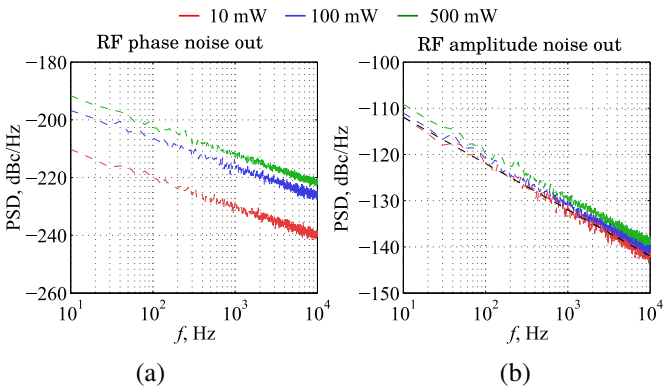


Fig. 7. The theoretical detected RF (a) phase and (b) amplitude noise spectra after optical propagation with a typical laser RIN at the input for a 6 km length of SMF-28 fiber.

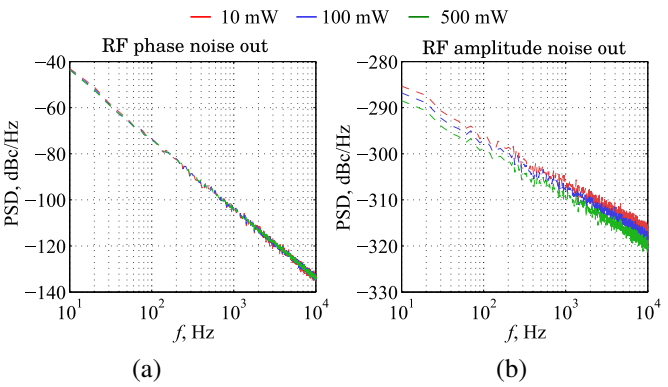


Fig. 8. The theoretical detected RF (a) phase and (b) amplitude noise spectra after optical propagation with an RF phase noise input for a 6 km length of SMF-28 fiber.

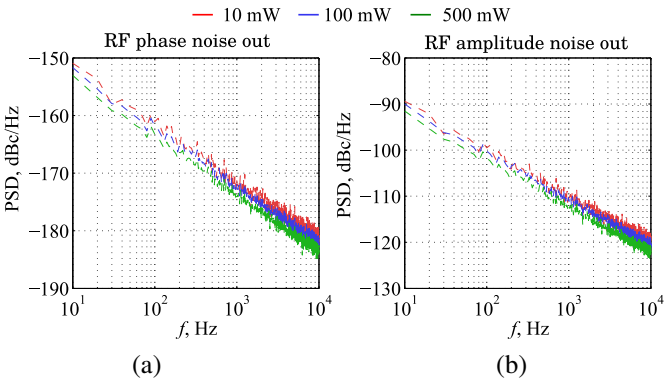


Fig. 9. The theoretical detected RF phase and amplitude noise spectra after optical propagation with an RF amplitude noise input for a 6 km length of SMF-28 fiber.

the same laser frequency noise at each power. We see that the RF phase noise is the same independent of power, and has excellent agreement with the linear calculation of Eq. 11 shown as the solid line.

Figure 7 shows the results of similar numerical experiments where in this case the only noise in the simulations is the laser RIN noise. The RIN converts directly to RF amplitude noise as expected, and is also parametrically amplified through propagation in the fiber, but only at powers which are considerably higher than that used in experimental OEOs and only by a small amount. Kerr nonlinearity and third order dispersion can also convert RIN to RF phase noise as seen in Fig. 7a; however, this is expected to be negligible — it is well below the white noise floor of the system.

With zero laser noise the conversion of a typical RF phase noise on the input signal before the modulator to RF phase and amplitude noise at the detector output is shown in Fig. 8. The RF phase noise out is not affected by Kerr nonlinearity and remains constant at different powers. Similarly, with zero laser noise the conversion of a typical RF amplitude noise on the input signal before the modulator to RF phase and amplitude noise at the detector output is shown in Fig. 9. The RF amplitude noise marginally decreases with increasing laser power, due to nonlinear squeezing effects. Third order dispersion converts RF amplitude noise to RF phase noise, which is less than other sources of RF phase noise in the system, and as the amplitude noise is gain clamped by the saturable gain of the loop, is not expected to be large at low frequencies.

V. Conclusion

We are conducting a systematic investigation of the optical domain portion of OEOs with a view to modeling major optical processes that contribute to length-dependent RF phase noise. To date we have investigated the effects of dispersion and nonlinearity on both the large signal and noise in the OEO optical path. As has been previously demonstrated, dispersion converts laser frequency noise to RF phase noise; however, experiments have indicated that this is not the only source of length-dependent noise and the noise increases nonlinearly with the optical power. We have found that the Kerr nonlinearity of the fiber was not the cause of this observed nonlinear length-dependent RF phase noise. In the future we plan to theoretically investigate other possible nonlinear amplification processes in the fiber, in particular Brillouin and Rayleigh effects.

REFERENCES

[1] O. Okusaga, E. Adles, W. Zhou, C. Menyuk, G. Carter, E. Levy, and M. Horowitz, "Spurious-mode suppression in optoelectronic oscillators," in *International Frequency Control Symposium (IFCS)*. IEEE, 2010, pp. 539–543. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5556272

[2] K. Volyanskiy, Y. Chembo, L. Larger, and E. Rubiola, "Contribution of Laser Frequency and Power Fluctuations to the Microwave Phase Noise of Optoelectronic Oscillators," *J. Lightw. Technol*, vol. 28, no. 18, pp. 2730–2735, 2010. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5545364

[3] D. Eliyahu, D. Seidel, and L. Maleki, "RF amplitude and phase-noise reduction of an optical link and an opto-electronic oscillator," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 2, pp. 449–456, 2008. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4435092

[4] E. Adles, A. Docherty, C. Menyuk, G. Carter, O. Okusaga, W. Zhou, E. Levy, A. David, and M. Horowitz, "Loop-length dependent sources of phase noise in optoelectronic oscillators," in *International Frequency Control Symposium (IFCS)*. IEEE, 2010, pp. 550–553. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5556267

[5] E. C. Levy, O. Okusaga, M. Horowitz, C. R. Menyuk, W. Zhou, and G. M. Carter, "Comprehensive computational model of single- and dual-loop optoelectronic oscillators with experimental verification." *Opt. Express*, vol. 18, no. 20, pp. 21 461–76, Sept. 2010. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/20941042

[6] X. Yao and L. Maleki, "Optoelectronic microwave oscillator," *JOSA B*, vol. 13, no. 8, pp. 1725–1735, 1996. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=josab-13-8-1725

[7] G. Agrawal, *Nonlinear fiber optics*, 3rd ed. Academic Press, 2001. [Online]. Available: http://books.google.com/books?id=wjHP0oAVcScC&pgis=1

Generation of Kerr Combs in MgF_2 and CaF_2 Microresonators

W. Liang, A. B. Matsko, A. A. Savchenkov, V. S. Ilchenko, D. Seidel, and L. Maleki
OEwaves Inc., 2555 E. Colorado Blvd., Ste 400, Pasadena, California, 91107 U.S.A.
Email: Andrey.Matsko@oewaves.com

***Abstract*— We report on a study of hyper-parametric oscillation in high-Q calcium fluoride and magnesium fluoride optical microresonators, and show that, magnesium fluoride microresonators generate wider Raman-free Kerr frequency combs at telecom wavelengths, in contrast with calcium fluoride ones. We observed 30 nm wide frequency combs in magnesium fluoride microresonators pumped with only 2 mW of continuous wave 1543 nm light.**

I. Introduction

Optically pumped solid state microresonators allow generating optical frequency combs resulting from four-wave mixing (FWM) in the resonator host material (frequently dubbed as Kerr combs) [1]-[11]. The efficiency of the FWM process depends on group velocity dispersion (GVD) and the thermo-refractive properties of the resonator modes. Anomalous GVD as well as small thermo-refractivity is desirable for generation of octave-spanning Kerr combs [8]. On the other hand, stable frequency combs are also produced in resonators possessing normal GVD [2], [3], [12]. In this work we discuss results of our recent efforts on optimization of the resonator morphology and host material selection for generation of stable Kerr combs at telecom wavelengths. We have compared optical combs generated in calcium (CaF_2) and magnesium (MgF_2) fluoride whispering gallery mode (WGM) resonators. The resonators are characterized with the opposite signs of GVD at 1543 nm. We found that the Kerr comb generation in magnesium fluoride resonators suffer less from the stimulated Raman [13], [14] as well as Brillouin [15], [16] scattering, compared to the comb generation in calcium fluoride WGM resonators. The power efficiency of the comb generation in MgF_2 WGM resonators is also superior.

II. Kerr combs in MgF_2 resonators

Generation of Kerr frequency combs usually starts with so called hyper-parametric oscillation resulting from conversion of two pump photons into two Stokes and anti-Stokes photons [12]. Having competing hyper-parametric oscillation [17], [18] and Raman lasing [13], [14] is an undesirable feature in the Kerr comb generators. Usage of a resonator host material possessing anomalous GVD is one of the ways for suppressing Raman lasing, since the anomalous GVD relaxes phase matching conditions required for the efficient comb generation, and the threshold of the hyper-parametric oscillation becomes lower than that of Raman lasing. The resonator host material also should have a low thermo-refractive constant to reduce related low frequency instabilities. We found that magnesium fluoride possesses both of these properties. In addition, crystalline MgF_2 features an excellent mechanical stability, hardness, and optical transparency, making it promising for production of very high-Q WGM microresonators [19], [20], [21], [22], [23]. Resonators made with MgF_2 are attractive because they allow dual mode thermal stabilization that can lead to stable and power efficient RF oscillators without any atomic frequency references involved [24].

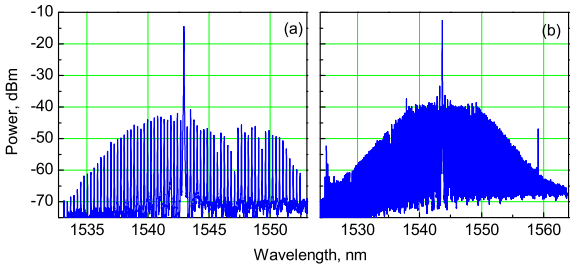


Fig. 1. (a) and (b) Spectra of optical frequency combs generated in two different MgF_2 whispering gallery mode microresonators pumped with approximately 2 mW of continuous wave light. The combs have 34.7 GHz (a) and 27 GHz (b) repetition rates corresponding to the FSR of the resonators. No SRS is observed.

We have demonstrated an optical frequency comb generated in MgF_2 WGM resonators characterized with 34.7 GHz and 27 GHz FSRs, and loaded Q-factor of about 10^9. The resonators were fabricated from a commercially available z-cut MgF_2 optical window by mechanical polishing. The shape of the resonators in the vicinity of mode localization was that of an oblate spheroid (with removed polar cups). One of the resonators (see Fig. (3)) was characterized with radius $a = 1$ mm, corresponding to 34.67 GHz free spectral range (FSR) of the fundamental mode family. The other resonator was characterized with radius $a = 1.27$ mm, corresponding to 27 GHz FSR of the fundamental mode family. The resonators were pumped optically using identical 1543 nm distributed feedback diode lasers self-injection locked [25] to WGM modes of the resonators. The locking occurred due to the intrinsic Rayleigh scattering in the resonators host material. The optical power emitted by the lasers was approximately 4 mW. The coupling efficiency of the pump light to the corresponding resonator mode was approximately 30%, while

978-1-61284-110-6/11/$26.00 ©2011 IEEE

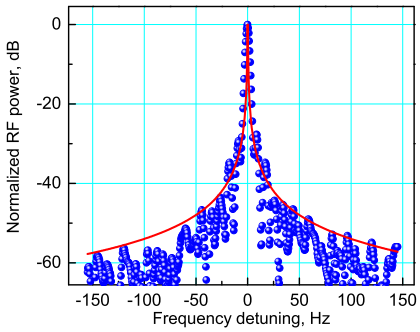


Fig. 2. Frequency spectrum of the RF signal produced by the 27 GHz frequency comb impinging (Fig. 1b) on a fast photodiode, and recorded using an RF spectrum analyzer (blue circles). The resolution bandwidth of the measurement and the span are 300 Hz. The instantaneous linewidth determined by fitting the skirts of the RF spectrum with a Lorentzian line (red solid line) is 0.4 Hz. The center frequency of the RF line is drifting in time determining the long term (in)stability of the signal resulting from the thermal drifts in the setup, while the instantaneous linewidth is a measure of the short term stability intrinsically determined by the oscillation process (see [25] and references therein). The measured RF carrier frequency is 27 GHz, and the corresponding frequency detuning is calculated from this frequency.

50% of the light exiting the resonators is collected into a multimode fiber coupled to an optical spectrum analyzer. An example of the observed frequency spectra of the generated combs is shown in Fig. 2. No stimulated Raman scattering (SRS) was observed with the resonators.

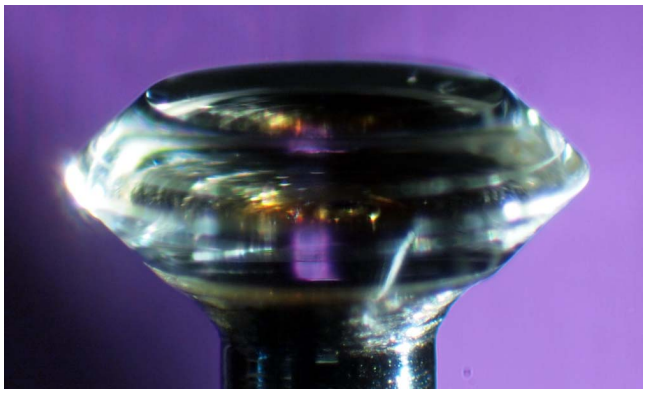


Fig. 3. A picture of MgF_2 resonator characterized with 34.7 GHz free spectral range on the manufacture post.

To our knowledge this is the most power-efficient, in terms of spectral width and pump power ratio, frequency comb generator reported so far (see Table I).

This is also the first demonstration of Kerr comb generation in an anisotropic crystalline material [29]. Our previous attempts to produce frequency combs in crystalline SiO_2, $LiNbO_3$, and $LiTaO_3$ WGM resonators were unsuccessful, even though those materials are characterized with large values of cubic nonlinearity and low values of optical attenuation that facilitate generation of Kerr combs. Fast cubic nonlinearity is approximately the same in transparent dielectrics

Material	Pump, mW,	Width, THz	Repetition, GHz
Silica [1]	130	63	895
Silica [8]	2,500	165	850
CaF_2 [2]	50	24	25
CaF_2 [3]	25	4.3	14
Si_3N_4 [26]	310	35	403
Si_3N_4 [27]	300	75	204
MgF_2 [28]	2	2.4	35
MgF_2 [this work]	2	4.4	27

TABLE I

PARAMETERS OF VARIOUS KERR COMBS.

with several exceptions. Nonlinear refractive index n_2 and corresponding XPM/SPM interaction constant g for various transparent materials is shown in Table II. Coupling parameter characterizing comb generation $g = \hbar\omega_0^2 n_2 c/(\mathcal{V} n_0^2)$ (where ω_0 is the frequency of the cw pumping light, n_0 is the linear refractive index of the material; see also [30]) is found for a resonator with mode volume $\mathcal{V} = 2 \times 10^{-7}$ cm^3 (1 mm mode radius, 3 μm radial mode depth, 10 μm mode height) pumped at 1,550 nm. Notations "o" and "e" stay for ordinary and extraordinary polarization of light. The results vary from research to research within 20% margin.

Data presented in Table 1 suggests that the four-wave mixing should have similar efficiency in the dielectrics provided that losses and phase matching conditions are identical. Some materials, like chalcogenide glasses and semiconductors have much larger nonlinearity. However, this nonlinearity is usually accompanied by nonlinear absorption that reduces the efficiency of the four-wave mixing.

Material	n_2, 10^{-16} cm^2/W	g, 10^{-3} rad/s
Al_2O_3	2.8 [31]	2
CaF_2	1.9 [32]	2
MgF_2	0.9 (e,o) [32], [33]	1 (e,o)
SiO_2	3.4 [34]	3
Fused silica	2.6 [35]	3
$LiNbO_3$	20 (o)/ 5 (e) [36]	9 (o)/ 2.5 (e)
Si_3N_4	25 [37]	14
As_2Se_3	1200 [38]	500

TABLE II

CUBIC NONLINEARITY OF OPTICAL MATERIALS

We see the reason for this in the interplay of the not optimal GVD, leading to increase of the comb generation threshold, and the low enough threshold of the SRS, competing with the Kerr comb generation, in the materials. SRS results in the pump depletion suppressing Kerr comb generation. Kerr combs always have lower threshold compared with SRS in the MgF_2 WGM resonator we studied.

We studied the dependence of the comb optical spectrum on resonator loading, while keeping the laser current constant (Fig. 4). The detuning of the frequency of the pumping light from the corresponding optical resonance was kept as small as possible, though we did not control this parameter directly.

As expected, no comb was observed when the coupling prism touched the resonator surface and the loaded Q-factor became too small to support the comb generation. The loading was then decreased until hyper-parametric oscillation started and several spectral lines of the comb separated by almost a THz appeared. A further decrease of the resonator loading resulted in the reduction of the repetition rate and creation of higher density combs. After stabilization of the resonator loading, adjustment of the laser current produced stable Kerr frequency combs with desirable repetition rates and spectral shapes.

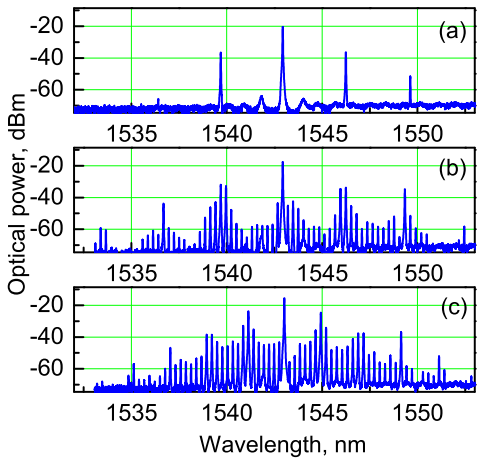


Fig. 4. Growth of the frequency comb generated in the MgF_2 resonator (FSR=34.67 GHz) pumped with cw light. The loading of the resonator decreases (the Q-factor increases) from (a) to (c). The comb with higher repetition rate is generated in the overloaded resonator. Increase of the loaded Q-factor results in filling the gaps between the modes. Finally the comb transforms to that in Fig. (1a).

III. Dynamics of the comb growth and its analysis

According to our experiments with WGM resonators characterized with anomalous GVD, Kerr frequency combs with higher frequency repetition rates have lower threshold compared with smaller repetition rate combs. To explain the observation and to find the dependence of the threshold conditions on the repetition rate we consider the case of two sidebands only and the pump (a model of a hyper-parametric oscillator [30]), and assume the power of generated sidebands is much smaller than the pump power. The set of equations linearized with respect to sidebands can be written as follows

$$\left[\Gamma_0 - ig|A|^2\right]A = F_0, \tag{1}$$

$$\dot{B}_+ = -\Gamma_+ B_+ + 2ig|A|^2 B_+ + igB_-^* A^2, \tag{2}$$

$$\dot{B}_-^* = -\Gamma_-^* B_-^* - 2ig|A|^2 B_-^* - igB_+(A^*)^2, \tag{3}$$

where $\Gamma_0 = i(\omega_0 - \omega) + \gamma_0$, $\Gamma_\pm = i(\omega_\pm - \tilde{\omega}_\pm) + \gamma_\pm$, F_0 stands for the cw pump amplitude; ω_0, ω_+, and ω_- are the eigenfrequencies of the modes; ω is the frequency of the pump, $\tilde{\omega}_+$ and $\tilde{\omega}_-$ are the frequencies of the generated sidebands. Decay rates γ_0, γ_+, and γ_- represent both coupling and internal losses. Due to the energy and photon number conservation we find $\tilde{\omega}_- + \tilde{\omega}_- = 2\omega$.

We introduce dimensionless variables $\xi = g|A|^2/\gamma_0$ where $A = |A|\exp i\phi_0$; $\mathcal{B}_+ = |B_+|/|A|$ where $B_+ = |B_+|\exp i\phi_+$; $\mathcal{B}_- = |B_-|/|A|$ where $B_- = |B_-|\exp i\phi_-$; $\psi = \phi_{F_0} - \phi_0$ where $F_0 = |F_0|\exp i\phi_{F_0}$; $\phi = 2\phi_0 - \phi_+ - \phi_-$; $\Delta_+ = (\omega_+ - \tilde{\omega}_+)/\gamma_0$; and $\Delta_- = (\omega_- - \tilde{\omega}_-)/\gamma_0$. Furthermore, we introduce $f = (g/\gamma_0)^{1/2}|F_0|/\gamma_0$, $\Delta_0 = (\omega_0 - \omega)/\gamma_0$, and $D = (2\omega_0 - \omega_+ - \omega_-)/\gamma_0$. For the sake of simplicity we assume $\gamma_\pm = \gamma_0$.

The characteristic equation for Eqs. (2) and (3) is

$$\begin{vmatrix} \lambda + 1 + i(\Delta_+ - 2\xi) & -i\xi e^{i\phi} \\ i\xi e^{-i\phi} & \lambda + 1 - i(\Delta_- - 2\xi) \end{vmatrix} = 0. \tag{4}$$

which has two solutions

$$\lambda_\pm = -1 \pm \frac{1}{2}\sqrt{4\xi^2 - (2\Delta_0 - D - 4\xi)^2} - \frac{i}{2}(\Delta_+ - \Delta_-). \tag{5}$$

The instability condition for the sidebands (oscillation condition) is

$$4(\xi^2 - 1) - (2\Delta_0 - D - 4\xi)^2 > 0 \tag{6}$$

where ξ is given by steady state solution of Eq. (1)

$$\xi\left[1 + (\Delta_0 - \xi)^2\right] = f^2. \tag{7}$$

Eq. (6) results in the threshold condition for the normalized amplitude of the pump field inside the resonator

$$\xi_\pm = \frac{1}{3}\left(2\Delta_0 - D \pm \sqrt{\frac{1}{4}(2\Delta_0 - D)^2 - 3}\right). \tag{8}$$

Since, by definition $\xi \geq 0$, we find $2\Delta_0 - D > 2\sqrt{3}$. In the case of anomalous GVD the pump mode can be dynamically stable (no bistability observed) and modulationally unstable [39]. The conditions are: $1 < f^2 < 8/(3\sqrt{3})$ and $\lambda_+ > 0$.

This analysis allows us to find the dependence of the preferred repetition rate of the comb on the loaded bandwidth of the resonator. In the experiment we kept the optical pump at the corresponding resonance, selecting $\Delta_0 = \xi$, and $\xi = f^2$ (Eq. (7)). The maximal gain (5) ($\lambda_+ = \xi - 1$) occurs when $D = -2\xi$. The value of D depends on the comb repetition rate and is a free parameter in our case. Since the oscillation threshold power corresponds to $\xi = f^2 = 1$ ($\lambda_+ > 0$), the hyper-parametric oscillation that eventually leads to Kerr frequency comb generation starts in the mode pair for which $D = -2$, or

$$\nu_{i+j} + \nu_{i-j} - 2\nu_i \simeq \gamma_{FWHM}. \tag{9}$$

where ν_i is the frequency of ith mode, and $\gamma_{FWHM} = 2\gamma_0$ is the full width at half maximum (FWHM) of the loaded resonance.

The condition (9) is valid for resonant tuning of the external pump and can be fulfilled only in the case of anomalous GVD [4]. The resonant tuning here is determined in terms of the optically pumped, not the "cold", resonator. Because the

modes are nonlinear, the SPM results in a shift of the optically pumped resonance when the optical pump power increases. To stay resonant, i.e. to maintain the maximal optical power within the resonator for a fixed power of the external pump, the pump frequency should be properly adjusted such that the condition $\Delta_0 = \xi$ is preserved.

The simplest theoretical model of the Kerr comb growth can be built based on the model described above of three-mode hyper-parametric oscillator [30]. Increase of the optical pump power above the oscillation threshold results in generation of the two preferable optical sidebands. Those sidebands beat with themselves and the pump light and produce higher order sidebands via the XPM effect. The threshold of the comb generation is determined by the threshold of the hyper-parametric oscillation, and the higher order sidebands are always phase locked to the lower order sidebands [12].

The comb repetition rate depends on the switching dynamics of the system. The hyper-parametric oscillations have the pump-power-dependent gain maximum. Generation of a higher frequency repetition rate comb is expected at the higher pump power [9]. This is a result based on the notion of compensation of the nonlinear phase shift by the anomalous GVD [4]. On the other hand, an adiabatic increase of the pump power results in broadening of the lower repetition rate frequency comb started at lower pump power. This process is stable if growth of the higher-order optical sidebands properly reduce the power of light confined within the optical mode.

Since the power level of the optical pumping used in our experiment is limited, the comb generation starts only when the Q-factor of the mode exceeds a certain level. The corresponding FWHM of the maximally overloaded mode is approximately 2 MHz. Calculating the GVD of the fundamental mode sequence using Sellmeier equation for MgF_2 (Fig. 5) we find that, according to (9), the lowest threshold is reached for the modes being at 0.6 THz from the carrier, which approximately corresponds to the comb observed in Fig. (4a) for the case of highest loading of the resonator. Reduction of the loading results in the reduction of the preferred comb repetition rate and to the growth of the comb at a single FSR as in Fig. (4b and 4c). The simplest three mode analysis [30] cannot explain this process, and numerical methods [39] should be applied.

We did not observe nonlinear processes other than four-wave mixing in MgF_2 resonators. It does not mean that the processes are absent there. The most probable reason is that the four wave mixing dominates at the wavelength and power levels of the pumping light as well as the WGM resonator size and shape we utilized. The spectrum of the observed nonlinear phenomena is much broader in the case of CaF_2 resonators. In what follows we discuss two cases of competition of the comb generation process with SRS and SBS processes.

IV. Competition of Kerr combs and SRS in CaF_2 resonators

GVD depends on the size of the resonator and the wavelength of the pumping light, so there is an optimal combination

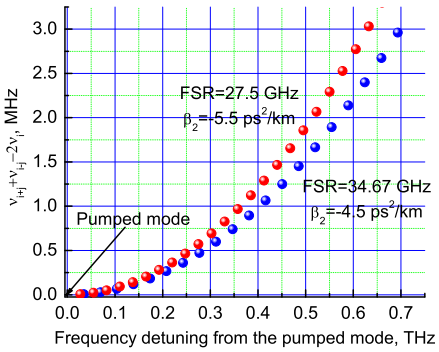


Fig. 5. GVD of the WGM spectrum found in terms of frequency difference $\nu_{i+j} + \nu_{i-j} - 2\nu_i$, where i is the azimuthal number of the pumped mode ($i = 5,548$ in the case of the resonator with 36.67 GHz FSR). The FSR of the resonator is defined as $\nu_{FSR} = \nu_i - \nu_{i-1} = 34.67$ GHz. The optical pump is extraordinarily polarized with respect to the crystalline axis (electric filed vector is collinear with the resonator axis and the symmetry axis of the crystal). The estimated value of GVD of the modes of the resonator with 36.67 GHz FSR is $\beta_2 = -(\nu_{i-1} + \nu_{i+1} - 2\nu_i)(n_e/2\pi c\nu^2_{FSR}) = -4.5$ ps^2/km, where $n_e = 1.38$ is the extraordinary index of refraction of the material, c is the speed of light in the vacuum. Similar calculations can be done for the resonator with 27 GHz FSR.

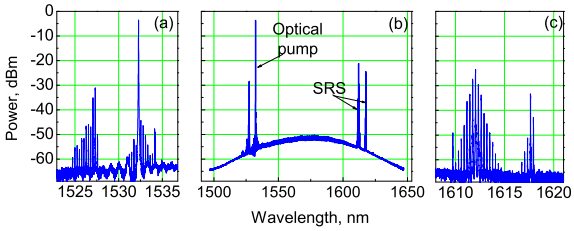


Fig. 6. Competition of stimulated Raman scattering (SRS) and hyper-parametric oscillation in a fluorite microresonator. (b) Optical spectrum generated in a CaF_2 resonator having 34.6 GHz FSR, and pumped with 1 mW of continuous wave 1532 nm light. (a) and (c) show details of the spectrum. The observed SRS is more efficient than the hyperparametric oscillation process. Nonlinear interaction of the Stokes Raman comb with the carrier results in generation of "anti-Stokes" sidebands in the vicinity of the carrier as well as the second lower frequency manifold of the Raman comb. It worth noting that the manifold also experiences Raman gain since the Raman line is broad enough.

of those parameters for Kerr comb generation. For instance, CaF_2 WGM resonators demonstrate generation of Raman-free, stable, and wide enough Kerr fundamental combs with repetition rates around 10 GHz and centered at 1550 nm, in the vicinity of zero GVD point of the material [12]. On the other hand, SRS [13] becomes preferable in smaller CaF_2 resonators (35 GHz FSR) pumped at 1532 nm, and phase locked Raman frequency combs are readily observed there [14]. One of the reasons is that the GVD of CaF_2 is normal in the resonators at 1532 nm, while it is nearly zero at the Raman wavelength. The resonator prefers to oscillate at smaller normal FSR.

An example of competition of the FWM and SRS processes in those resonators is illustrated by Fig. (6). Raman Stokes line is generated as a dual frequency comb around 1611 nm.

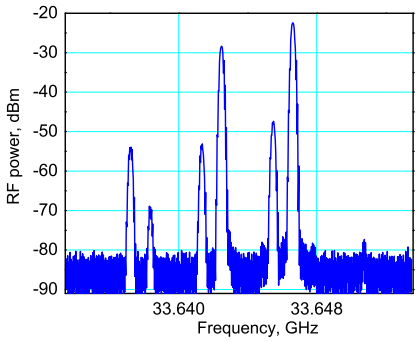


Fig. 7. Microwave spectrum generated by the optical signal shown in Figure 6 on a fast photodiode. The existence of the RF lines possessing different frequencies can be explained by the different values of the FSR at different optical wavelengths. FSR is about 4 MHz smaller at the carrier wavelength (1532 nm) than at the Raman wavelength (1611 nm). This is verified by using an optical high-pass filter. The repetition rates of the combs are determined by the different FSR values.

At the same time the Kerr comb is generated around the optical pump line (1532 nm). FWM of the Raman combs and pumping light results in generation of an anti-Stokes frequency comb in the vicinity of the pump line, as in dual-pump fiber-optic parametric amplifiers [40]. To verify the phase locking feature of the generated optical combs we sent them to a fast photodiode and found that the combs produce several narrow RF signals separated by several MHz, Fig. (7).

V. Competition of Kerr combs and SBS in CaF_2 resonators

Stimulated Brillouin scattering (SBS) is usually forbidden in solid state ring microresonators with high quality factors. The reason for this is the phase mismatch between the optical phonons and light. If the pump light is resonant with a mode of the resonator, the scattered light has no optical mode to interact with. Moreover, if occasionally such a mode is present, the overlap integral between the acoustic mode as well as the pump and Stokes optical modes tends to be zero. The absence of SBS in optical microresonators is used as an advantage in studies of weaker nonlinear effects such as four wave mixing and stimulated Raman scattering. On the other hand, it is interesting to understand for which morphology as well as dimensions of the microresonators the SBS is not forbidden. An SBS laser based on a microresonator should possess a low threshold as well as narrow linewidth.

A triply-resonant SBS was demonstrated recently in over-moded WGM resonators [15], [16]. Since the enhanced SBS uses bulk, not surface, acoustic modes, the observed SBS frequency was given by the properties of the resonator host material, and shape of the resonator was less important. Due to the high absorption of the hypersound waves the finesse of the acoustic "mode" can be less than one, which simplified phase matching of the process.

We observed interplay between SBS and four-wave mixing in an overmoded CaF_2 resonator pumped with 1550 nm light (Figure 8a). Because we studied only the light scattered in the forward direction, we have observed pronounced second order SBS harmonic. The first order harmonic was visible due to Rayleigh scattering within the resonator modes. The Anti-Stokes line was generated due to four-wave mixing process. All three generated harmonics are phase locked with the pump light, since they produce single RF tone at the fast photodiode (Figure 8b). The locking occurs due to four-wave mixing.

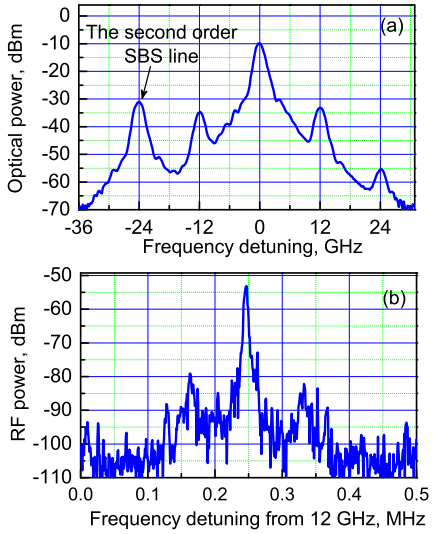


Fig. 8. (a) Optical signal resulting from the competing stimulated Brillouin scattering (SBS) and four wave mixing in a CaF_2 WGM resonator. The resonator is pumped at 1550 nm. The frequency of the SBS Stokes line is 12.1 GHz. (b) RF signal produced by the optical signal on a fast photodiode.

The Brillouin frequency shift for the case of backscattering along the axis of a cubic crystal can be found approximately as

$$\Delta\nu_B = \frac{2n_0}{\lambda_0}\sqrt{\frac{C_{11}+C_{12}+2C_{44}}{2\rho}}. \qquad (10)$$

where $n_0 = 1.424$ is the refractive index of the material at $\lambda_0 = 1550$ nm, $\rho = 3.18$ g/cm^3 is the density of the material. Calcium fluoride is a cubic crystal described by three independent elastic constants: $C_{11} = 1.6420$, $C_{12} = 0.4398$, $C_{44} = 0.3370$ [$\times 10^{11}$ Pa]. Substituting these values into Eq. (10) we find $\Delta\nu_B \simeq$ 12.1GHz. Further analysis shows that the polarization of the phonons is nearly collinear with their wave vector in the case of the highest–frequency branch. Hence, the observed SBS process results from the interaction of light and longitudinal hypersound wave.

VI. Conclusion

We have observed very efficient generation of Kerr frequency combs in optically pumped magnesium fluoride whispering gallery mode resonators and found that the resonators are free from the undesirable competition of the four-wave

mixing and Raman as well as Brillouin stimulated scattering observed in calcium fluoride resonators under similar conditions. The MgF_2 kerr comb generators are useful for spectroscopy, production of spectrally pure RF signals, and for generation of nonclassical states of light.

ACKNOWLEDGEMENTS

The work was supported in part by DARPA MTO.

REFERENCES

[1] P. Del'Haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, and T. J. Kippenberg, "Optical frequency comb generation from a monolithic microresonator," Nature (London) **450**, 1214-1217 (2007).

[2] A. A. Savchenkov, A. B. Matsko, V. S. Ilchenko, I. Solomatine, D. Seidel, and L. Maleki, "Tunable Optical Frequency Comb with a Crystalline Whispering Gallery Mode Resonator," Phys. Rev. Lett. **101**, 093902 (2008).

[3] I. S. Grudinin, N. Yu, and L. Maleki, "Generation of optical frequency combs with a CaF_2 resonator," Opt. Lett. **34**, 878-880 (2009).

[4] I. H. Agha, Y. Okawachi, and A. L. Gaeta, "Theoretical and experimental investigation of broadband cascaded four-wave mixing in high-Q microspheres," Opt. Express **17**, 16209-16215 (2009).

[5] J. S. Levy, A. Gondarenko, M. A. Foster, A. C. Turner-Foster, A. L. Gaeta, and M. Lipson, "CMOS-compatible multiple-wavelength oscillator for on-chip optical interconnects," Nature Photonics **4**, 37-40 (2009).

[6] L. Razzari, D. Duchesne, M. Ferrera, R. Morandotti, S. Chu, B. E. Little, and D. J. Moss, "CMOS-compatible integrated optical hyper-parametric oscillator," Nature Photonics **4**, 41-45 (2009).

[7] D. Braje, L. Hollberg, and S. Diddams, "Brillouin-enhanced hyperparametric generation of an optical frequency comb in a monolithic highly nonlinear fiber cavity pumped by a cw laser," Phys. Rev. Lett. **102**, 193902 (2009).

[8] P. Del'Haye, T. Herr, E. Gavartin, R. Holzwarth, and T. J. Kippenberg, "Octave spanning frequency comb on a chip," arXiv:0912.4890v1 [physics.optics] 24 Dec 2009.

[9] O. Arcizet, A. Schliesser, P. DelHaye, R. Holzwarth, and T. J. Kippenberg, "Optical frequency comb generation in monolithic microresonators," in *Practical Applications of Microresonators in Optics and Photonics*, A. B. Matsko, ed. (CRC Press, 2009), Chap. 11.

[10] A. B. Matsko, A. A. Savchenkov, W. Liang, V. S. Ilchenko, D. Seidel, and L. Maleki, "Whispering gallery mode oscillators and optical comb generators," Proc. of 7^{th} Symp. Frequency Standards and Metrology, ed. L. Maeki, pp. 539-558 (World Scientific, New Jersey, 2009).

[11] Y. K. Chembo, D. V. Strekalov, and N. Yu, "Spectrum and dynamics of optical frequency combs generated with monolithic whispering gallery mode resonators," Phys. Rev. Lett. **104**, 103902 (2010).

[12] L. Maleki, V. S. Ilchenko, A. A. Savchenkov, W. Liang, D. Seidel, and A. B. Matsko, "High performance, miniature hyper-parametric microwave photonic oscillator," Proc. of IEEE IFCS2010, 558–563 (2010).

[13] S. M. Spillane, T. J. Kippenberg, and K. J. Vahala, "Ultralowthreshold Raman laser using a spherical dielectric microcavity," Nature (London) **415**, 621-623 (2002).

[14] W. Liang, V. S. Ilchenko, A. A. Savchenkov, A. B. Matsko, D. Seidel, and L. Maleki, "Passively mode-locked Raman laser," Phys. Rev. Lett. **105**, 143903 (2010) & arXiv:1006.5077.

[15] I. S. Grudinin, A. B. Matsko, and L. Maleki, "Brillouin lasing with a CaF_2 whispering gallery mode resonator," Phys. Rev. Lett. **102**, 043902 (2009).

[16] M. Tomes and T. Carmon, "Photonic micro-electromechanical systems vibrating at X-band (11-GHz) rates," Phys. Rev. Lett. **102**, 113601 (2009).

[17] T. J. Kippenberg, S. M. Spillane, and K. J. Vahala, "Kerr-Nonlinearity Optical Parametric Oscillation in an Ultrahigh-Q Toroid Microcavity," Phys. Rev. Lett. **93**, 083904 (2004).

[18] A. A. Savchenkov, A. B. Matsko, D. Strekalov, M. Mohageg, V. S. Ilchenko, and L. Maleki, "Low Threshold Optical Oscillations in a Whispering Gallery Mode CaF_2 Resonator," Phys. Rev. Lett. **93**, 243905 (2004).

[19] I. S. Grudinin, A. B. Matsko, A. A. Savchenkov, D. Strekalov, V. S. Ilchenko, and L. Maleki, "Ultra-high Q crystalline microcavities," Opt. Commun. **265**, 3338 (2006).

[20] H. Tavernier, N. N. T. Kim, P. Feron, R. Bendoula, P. Salzenstein, E. Rubiola, and L. Larger, "Optical disk resonators with micro-wave free spectral range for optoelectronic oscillator," in Proc. 22^{nd} Eur. Time and Frequency Forum, Toulouse, France, 2008, Paper FPE-0179.

[21] H. Tavernier, P.Salzenstein, K. Volyanskiy, Y. K. Chembo, and L. Larger, "Magnesium fluoride whispering gallery mode disk-resonators for microwave photonics applications," IEEE Phot. Tech. Lett. **22**, 1629-1631 (2010).

[22] J. Hofer, A. Schliesser, and T. J. Kippenberg, "Cavity optomechanics with ultrahigh-Q crystalline microresonators," Phys. Rev. A **82**, 031804(R) (2010).

[23] J. Alnis, A. Schliesser, C. Y. Wang, J. Hofer, T. J. Kippenberg, T. W. Hänsch, "Thermal-noise limited laser stabilization to a crystalline whispering-gallery-mode resonator" $arXiv : 1102.4227$.

[24] A. A. Savchenkov, A. B.Matsko, V. S. Ilchenko, N. Yu, and L. Maleki, "Whispering-gallery-mode resonators as frequency references. II. Stabilization," J. Opt. Soc. Am. B **24**, 2988-2997 (2007).

[25] W. Liang, V. S. Ilchenko, A. A. Savchenkov, A. B. Matsko, D. Seidel, and L. Maleki, "Ultra-narrow linewidth external cavity semiconductor lasers using crystalline whispering gallery mode resonators," Opt. Lett. **35**, 2822-2824 (2010).

[26] J. S. Levy, A. Gondarenko, M. A. Foster, A. C. Turner-Foster, A. L. Gaeta, M. Lipson, "CMOS-compatible multiple wavelength oscillator for on-chip optical interconnects," Nat. Photon. **4**, 37-40 (2010).

[27] M. A. Foster, J. S. Levy, O. Kuzucu, K. Saha, M. Lipson, A. L. Gaeta, "A Silicon-Based Monolithic Optical Frequency Comb Source," $arXiv : 1102.0326$.

[28] W. Liang, A. A. Savchenkov, A. B. Matsko, V. S. Ilchenko, D. Seidel, and L. Maleki, "Generation of near-infrared frequency combs from a MgF_2 whispering gallery mode resonator," Opt. Lett. (2011), to be published.

[29] After the paper was submitted to the meeting we found that Kerr combs were observed at 2.5 μm in MgF_2 WGM resonators by T. J. Kippenberg et al. (as reported at IEEE IFCS'11).

[30] A. B. Matsko, A. A. Savchenkov, D. Strekalov, V. S. Ilchenko, and L. Maleki, "Optical hyper-parametric oscillations in a whispering gallery mode resonator: threshold and phase diffusion," Phys. Rev. A **71**, 033804 (2005).

[31] A. Major, F. Yoshino, I. Nikolakakos, J. S. Aitchison, and P. W. E. Smith, "Dispersion of the nonlinear refractive index in sapphire," Opt. Lett. **29**, 602-604 (2004).

[32] D. Milam, M. J. Weber, and A. J. Glass, "Nonlinear refractive index of fluoride crystals," Appl. Phys. Lett. **31**, 822-824 (1977).

[33] R. DeSalvo, A. A. Said, D. J. Hagan, E. W. Van Stryland, M. Sheik-Bahae, "Infrared to ultraviolet measurements of two-photon absorption and n_2 in wide bandgap solids,"IEEE J. Quantum Electron. **32**, 1324-1333 (1996).

[34] R. Adair, L. L. Chase, and S. A. Payne, "Nonlinear refractive index of optical crystals," Phys. Rev. B **39**, 33373350 (1989).

[35] D. Milam, "Review and Assessment of Measured Values of the Nonlinear Refractive-Index Coefficient of Fused Silica," Appl. Opt. **37**, 546-550 (1998).

[36] I. A. Kulagin, R. A. Ganeev, R. I. Tugushev, A. I. Ryasnyansky, and T. Usmanov, "Analysis of third-order nonlinear susceptibilities of quadratic nonlinear optical crystals," J. Opt. Soc. Am. B **23**, 75-80 (2006).

[37] K. Ikeda, R. E. Saperstein, N. Alic, Y. Fainman, "Thermal and Kerr nonlinear properties of plasma-deposited silicon nitride/ silicon dioxide waveguides" Opt. Express **16**, 12987-12994 (2008).

[38] V. Ta'eed, N. J. Baker, L. Fu, K. Finsterbusch, M. R. E. Lamont, D. J. Moss, H. C. Nguyen, B. J. Eggleton, D.-Y. Choi, S. Madden, and B. Luther-Davies, "Ultrafast all-optical chalcogenide glass photonic circuits," Opt. Express **15**, 9205-9221 (2007).

[39] Y. K. Chembo and N. Yu, "Modal expansion approach to optical-frequency-comb generation with monolithic whispering-gallery-mode resonators," Phys. Rev. A **82**, 033801 (2010).

[40] K. K. Y. Wong, M. E. Marhic, K. Uesaka, and L. G. Kazovsky, "Polarization-independent two-pump fiber optical parametric amplifier," IEEE Photon. Technol. Lett. **14**, 911913 (2002).

Testing the Gravitational Redshift with Atomic Gravimeters?

Peter Wolf[1], Luc Blanchet[2], Christian J. Bordé[1,3],
Serge Reynaud[4], Christophe Salomon[5], Claude Cohen-Tannoudji[5]

[1]LNE-SYRTE, Observatoire de Paris, CNRS, UPMC, France
[2]GRECO, Institut d'Astrophysique de Paris, CNRS, UPMC, France
[3]Laboratoire de Physique des Lasers, Université Paris 13, CNRS, France
[4]Laboratoire Kastler Brossel, CNRS, ENS, UPMC, France
[5]Laboratoire Kastler Brossel et Collège de France, CNRS, ENS, UPMC, France

***Abstract*— Atom interferometers allow the measurement of the acceleration of freely falling atoms with respect to an experimental platform at rest on Earth's surface. Such experiments have been used to test the universality of free fall by comparing the acceleration of the atoms to that of a classical freely falling object. In a recent paper, Müller, Peters and Chu [Nature 463, 926-929 (2010)] argued that atom interferometers also provide a very accurate test of the gravitational redshift (or universality of clock rates). Considering the atom as a clock operating at the Compton frequency associated with the rest mass, they claimed that the interferometer measures the gravitational redshift between the atom-clocks in the two paths of the interferometer at different values of gravitational potentials. In the present paper we analyze this claim in the frame of general relativity and of different alternative theories, and conclude that the interpretation of atom interferometers as testing the gravitational redshift at the Compton frequency is unsound. The present work is a summary of our extensive paper [Wolf *et al.*, arXiv:1012.1194, Class. Quant. Grav. in press], to which the reader is referred for more details.**

I. Introduction

The gravitational redshift, considered as one of the "classical tests" of general relativity (GR), is actually a test of one facet of the Einstein equivalence principle (EEP). The weak version of the equivalence principle has been verified with high precision using torsion balances [1] and Lunar laser ranging [2]. The gravitational redshift was predicted by Einstein but not observed for a long time. It became observable with the advent of high precision quantum spectroscopy and was first measured in 1960 by Pound and Rebka [3], who used gamma ray spectroscopy of the radiation emitted and absorbed by ^{57}Fe nuclei. The emitter and absorber were placed at the top and bottom of a $22.5\,\mathrm{m}$ high tower at Harvard and the frequency difference predicted by GR was measured with about 1% uncertainty.

In a test of the gravitational redshift using clocks, one checks that the clock rates are universal — i.e. the relative rates depend only on the difference of gravitational potentials (as determined by the trajectories of massive test bodies) but not on the nature and internal structure of the clocks. In 1976 a hydrogen-maser clock was launched on a rocket to an altitude of $10,000\,\mathrm{km}$ and its frequency compared to a similar clock on ground. This yielded a test of the gravitational redshift with about 10^{-4} accuracy [4]. The European Space Agency will fly in 2013 the Atomic Clock Ensemble in Space (ACES) [5], including a highly stable laser-cooled atomic clock, which will test (in addition to many other applications in fundamental physics and metrology) the gravitational redshift to a precision of about 10^{-6}. The gravitational redshift is also tested in null redshift experiments in which the rates of different clocks (based on different physical processes or different atoms) are compared to each other [6].

Quite generally in a modern context, tests of GR measure the difference between the predictions of GR and of some generalized alternative theory or theoretical framework (see [7] for a review). It has to be kept in mind that the classification and inter-comparison of different tests have then to be defined with respect to the framework used.

Atom interferometers have reached high sensitivities in the measurement of the gravitational acceleration [8], [9]. This yields very important tests of the weak equivalence principle (WEP) or universality of free fall (UFF) when comparing the free fall of atoms with that of classical macroscopic matter (in practice a nearby freely falling corner cube whose trajectory is monitored by lasers). The relative precision of such tests of the UFF is currently 7×10^{-9}, using Cs [8], [10] or Rb [9] atoms. Although it remains less sensitive than tests using macroscopic bodies of different composition [2], [1] which have reached a precision of 2×10^{-13}, this UFF test is interesting as it is the most sensitive one comparing the free fall of quantum objects (namely Cæsium atoms) with that of a classical test mass (the corner cube). In this contribution, which summarizes the extensive paper [11], we investigate whether such experiments can also be interpreted as tests of the gravitational redshift.

II. Overview of our extensive paper [11]

In a recent paper, Müller, Peters and Chu [10] (hereafter abbreviated as MPC) proposed a new interpretation of atom interferometry experiments as testing the gravitational redshift, that is also the universality of clock rates (UCR), with a precision 7×10^{-9}, which is several orders of magnitude better than the best present [4] and near future [5] clock tests.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

A. Analogy with clock experiments

The main argument of MPC (see also the more detailed papers [12], [13]) is based on an analogy between atom interferometry experiments and classical clock experiments. The idea of clock experiments is to synchronize a pair of clocks when they are located closely to one another, and move them to different elevations in a gravitational field. The gravitational redshift will decrease the oscillation frequency of the lower clock relative to the higher one, yielding a measurable phase shift between them. There are two methods for measuring the effect. Either we bring the clocks back together and compare the number of elapsed oscillations, or we measure the redshift by means of continuous exchanges of electromagnetic signals between the two clocks. In both methods one has to monitor carefully the trajectories of the two clocks. For example, in the second method one has to remove the Doppler shifts necessarily appearing in the exchanges of electromagnetic signals.

In the first method, the phase difference between the two clocks when they are recombined together, can be written as a difference of integrals over proper time,

$$\Delta\varphi_{\rm clock} = \omega \left[\int_{\rm I} {\rm d}\tau - \int_{\rm II} {\rm d}\tau \right] \equiv \omega \oint {\rm d}\tau \,. \qquad (1)$$

The two clocks have identical proper frequency ω. We denote by I and II the two paths (with say I being at a higher altitude, i.e. a lower gravitational potential) and use the notation $\oint {\rm d}\tau$ to mean the difference of proper times between the two paths, assumed to form a close contour. The integrals in (1) are evaluated along the paths of the clocks, and we may use the Schwarzschild metric to obtain an explicit expression of the measured phase shift in the gravitational field of the Earth.

The phase shift measured by an atom interferometer contains a contribution which is similar to the clock phase shift (1), so it is tempting to draw an analogy with clock experiments. In this analogy, the role of the clock's proper frequency is played by the atom's (de Broglie-)Compton frequency $\omega_{\rm C} = mc^2/\hbar$, where m denotes the rest mass of the atom. However the phase shift includes also another contribution $\Delta\varphi_\ell$ coming from the interaction of the laser light used in the beam-splitting process with the atoms. Thus,

$$\Delta\varphi = \omega_{\rm C} \oint {\rm d}\tau + \Delta\varphi_\ell \,. \qquad (2)$$

Here we are assuming that the two paths close up at the entry and exit of the interferometer; otherwise, additional terms have to be added to (2). The schematic view of the atom interferometer showing the two interferometer paths I and II is given by Fig. 1. The first term in (2) is proportional to the atom's mass through the Compton frequency and represents the difference of Compton phases along the two classical paths. In contrast, the second term $\Delta\varphi_\ell$ does not depend on the mass of the atoms.

At first sight, the first term in (2) could be used for a test of the gravitational redshift, in analogy with classical clock experiments, and the precision of the test could be very good,

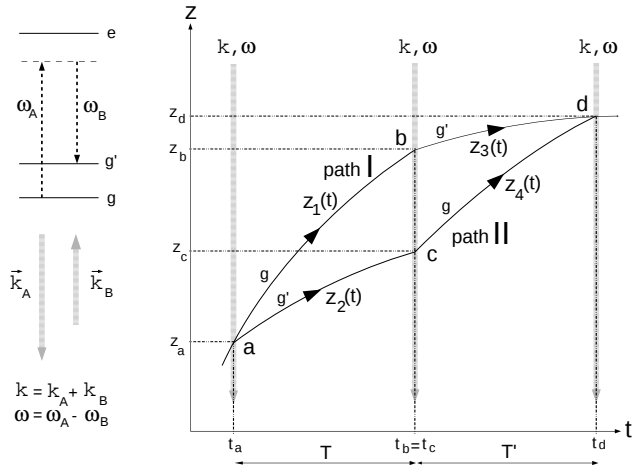


Fig. 1. Space-time trajectories followed by the atoms in the interferometer. Laser pulses occur at times t_a, $t_b = t_c$ and t_d, separated by time intervals T and T' (we have $T' = T$ when gravity gradients are neglected). The two-photon Raman transitions are between two hyperfine levels g and g' of the ground state of an alkali atom.

because the Compton frequency of the Cæsium atom is very high, $\omega_{\rm C} \approx 2\pi \times 3.0 \times 10^{25}\,$Hz. However, as shown in a previous brief comment [14] and the detailed paper [11], this re-interpretion of the atom interferometer as testing the UCR is fundamentally incorrect.[1]

At this point, we want to mention other crucial differences between atom interferometry and clock experiments, which reinforce our conclusions. In clock experiments the trajectories of the clocks are continuously controlled for instance by continuous exchange of electromagnetic signals. In atom interferometry in contrast, the trajectories of the atoms are not measured independently but theoretically derived from the Lagrangian and initial conditions. Furthermore, we expect that it is impossible to determine independently the trajectories of the wave packets without destroying the interference pattern at the exit of the interferometer.

Another important difference lies in the very notion of a clock. Atomic clocks use the extremely stable energy difference between two internal states. By varying the frequency of an interrogation signal (e.g. microwave or optical), one obtains a resonant signal when the frequency is tuned to the frequency of the atomic transition. In their re-interpretation, MPC view the entire atom as a clock ticking at the Compton frequency associated with its rest mass. But the "atom-clock" is not a real clock in the previous sense, since it does not deliver a physical signal at Compton frequency, as also recently emphasized in the same context in [16].

Note also that the phase shift (1) for clocks is valid in any gravitational field, with any gravity gradients, since it is simply the proper time elapsed along the trajectories of the clocks in

[1] See also the reply of MPC [15] to our brief comment. After completion of our work [14], [11], several independent analysis have appeared [16], [17] supporting our views and consistent with our conclusions.

a gravitational field. By contrast, the phase shift (2) is known only for quadratic Lagrangians [18] and cannot be applied in a gravitational field with large gravity gradients, or more generally with any Lagrangian that is of higher order.

B. Analysis in general relativity

The clear-cut argument showing that the atom interferometer does not measure the redshift is that the "atom-clock" contribution, i.e. the first term in (2), is in fact *zero* for a closed total path [18], [19]. As a specific example, let us consider the prediction from GR, which has been extensively treated in [20]; here we only present a very basic analysis sufficient for our purposes. The appropriate Lagrangian is given by the proper time $\mathrm{d}\tau = (-g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu/c^2)^{1/2}$, i.e. $L_{\text{GR}} = -mc^2\mathrm{d}\tau/\mathrm{d}t$, and is derived to sufficient accuracy using the Schwarzschild metric generated by the Earth,

$$L_{\text{GR}}(z,\dot{z}) = -mc^2 + \frac{GMm}{r_\oplus} - mgz + \frac{1}{2}m\dot{z}^2\,, \qquad (3)$$

where $r_\oplus$ is the Earth's radius, $g = GM/r_\oplus^2$ is the Newtonian gravitational acceleration, G is Newton's gravitational constant, M is the mass of the Earth, m the mass of the atom, c the speed of light in vacuum, z is defined by $r = r_\oplus + z$ with r the radial coordinate, and we neglect the post-Newtonian corrections. For simplicity we restrict ourselves to only radial motion, which is sufficient for the arguments in this paper.

The equations of motion are deduced from (3) using the principle of least action or the Euler-Lagrange equations, and read evidently $\ddot{z} = -g$. Then, when integrating (3) along the resulting paths, and calculating the difference of action integrals $\Delta\varphi_S$ as defined by the first term in (2), one finds

$$\Delta\varphi_S = 0\,. \qquad (4)$$

The key point about the result (4) is a consistent calculation of the two paths in the atom interferometer and of the phases along these paths, both derived from the same classical action, using in a standard way the principle of least action. At the deepest level, the principle of least action and its use in atom interferometry comes from the Feynman path integral formulation of quantum mechanics or equivalently the Schrödinger equation [18]. Thus only the second term $\Delta\varphi_\ell$ remains in (2). The final phase shift,

$$\Delta\varphi = \Delta\varphi_\ell = k\,g\,T^2\,, \qquad (5)$$

depends on the wavevector k of the lasers, on the interrogation time T and on the local gravity g. This shows that the atom interferometer is a gravimeter or accelerometer. The phase shift (5) arises entirely from the interactions with the lasers and the fact that the atoms are falling with respect to the laboratory in which the experiment is performed. The atom's Compton frequency is irrelevant.

C. Analysis in the modified Lagrangian formalism

This framework, which we call the "modified Lagrangian formalism", is an adaptation for our purpose of a powerful formalism for analyzing tests of the EEP and its various facets: the WEP, the local Lorentz invariance (LLI) and the local position invariance (LPI) [21], [22], [7]. This formalism allows deviations from general relativity and metric theories of gravity, with violations of the UFF and UCR, and permits a coherent analysis of atom interferometry experiments.

The formalism is defined by a single Lagrangian, that is however different from the GR Lagrangian. For our purposes it is sufficient to use a strongly simplified "toy" Lagrangian chosen as a particular case within the "energy conservation formalism" of Nordtvedt and Haugan [21], [22] (see [7] for a review). To keep in line with MPC we choose an expression similar to the Lagrangian of GR given by (3), namely

$$L_{\text{modified}} = -m_0\,c^2 + \frac{GMm_0}{r_\oplus} - \left(1+\beta_X^{(a)}\right) m_0\,g\,z + \frac{1}{2}m_0\,\dot{z}^2\,, \qquad (6)$$

where $\beta_X^{(a)}$ denotes a dimensionless parameter characterizing the violation of LPI, and depending on the particular type X of mass-energy or interaction under consideration; e.g. $\beta_X^{(a)}$ would be different for the electromagnetic or the nuclear interactions, with possible variations as a function of spin or the other internal properties of the atom, here labelled by the superscript (a). Thus $\beta_X^{(a)}$ would depend not only on the type of internal energy X but also on the type of atom (a).

The Lagrangian (6) describes the Newtonian limit of a large class of non-metric theories, in a way consistent with Schiff's conjecture and fundamental principles of quantum mechanics. Most alternative theories commonly considered belong to this class which encompasses a large number of models and frameworks (see [7] and references therein), like most non-metric theories (e.g. the Belinfante-Swihart theory [23]), some models motivated by string theory [24], some general parameterized frameworks like the energy conservation formalism [21], [22], the TH$\varepsilon\mu$ formalism [25], and the Lorentz violating standard model extension (SME) [26], [27].

Within this general formalism there is no fundamental distinction between UFF and UCR tests, as violation of one implies violation of the other, thus testing one implies testing the other. Depending on the theory used, different experiments test different parameters or parameter combinations at differing accuracies, so they may be complementary or redundant depending on the context. By varying (6) we obtain the equations of motion of the atom as

$$\ddot{z} = -\left(1+\beta_X^{(a)}\right) g\,, \qquad (7)$$

which shows that the trajectory of the atom is affected by the violation of LPI and is not universal. In fact we see that $\beta_X^{(a)}$ measures the non-universality of the ratio between the atom's passive gravitational mass and inertial mass. Thus, in the modified Lagrangian framework the violation of LPI implies a violation of WEP and the UFF, and $\beta_X^{(a)}$ appears to be the UFF-violating parameter. This is a classic example [21], [22] of the validity of Schiff's conjecture, namely that it is impossible in any consistent theory of gravity to violate LPI (or LLI) without also violating WEP.

The violation of LPI is best reflected in classical redshift experiments with clocks which can be analysed using a cyclic

gedanken experiment based on energy conservation [21], [22], [7]. The result for the frequency shift in a Pound-Rebka type experiment is

$$Z = \left(1 + \alpha_X^{(a)}\right) \frac{g\,\Delta z}{c^2}\,, \tag{8}$$

where the redshift violating (or UCR violating) parameter $\alpha_X^{(a)}$ is again non-universal. The important point, proven in [21], [22], [7], is that the UCR-violating parameter $\alpha_X^{(a)}$ is related in a precise way to the UFF-violating parameter $\beta_X^{(a)}$, namely

$$\beta_X^{(a)} = \alpha_X^{(a)}\, \frac{\overline{E}_X}{\overline{m}\, c^2}\,, \tag{9}$$

where $\overline{E}_X$ is the internal energy responsible for the violation of LPI, and $\overline{m}$ is the sum of the rest masses of the particles constituting the atom. Therefore we can compare the different qualitative meaning of tests of UCR and UFF. For a given set of UFF and UCR tests their relative merit is given by (9) and is dependent on the model used, i.e. the type of anomalous energy $\overline{E}_X$ and its dependence on the used materials or atoms.

We now consider the application to atom interferometry. In the experiment of [8], [10] the "atom-clock" that accumulates a phase is of identical composition to the falling object (the same atom), hence one has to consistently use the same value of $\beta_X^{(a)}$ when calculating the trajectories and the phase difference using the Lagrangian (6). It is then easy to show that $\Delta\varphi_S = 0$ with the above Lagrangian [18], [20], [28]. The vanishing of $\Delta\varphi_S$ in this case is a general property of all quadratic Lagrangians and comes from consistently using the same Lagrangian for the calculation of the trajectories and the phase shift. It is related to the cancellation between the kinetic term and the gravitational potential energy term in the Lagrangian (6).

Then the total phase shift of the atom interferometer is again given by the light interactions only, which are obtained from the phases at the interaction points evaluated using the trajectory given by (6). One then obtains

$$\Delta\varphi = \Delta\varphi_\ell = \left(1 + \beta_X^{(a)}\right) k\, g\, T^2\,. \tag{10}$$

We first note that in this class of theories the Compton frequency plays no role, as $\Delta\varphi_S = 0$. Second, we note that although $\beta_X^{(a)}$ appears in the final phase shift, this is entirely related to the light phase shift coming from the trajectory of the atoms, and thus is a measurement of the effective free fall acceleration $(1+\beta_X^{(a)})g$ of the atoms, which is given by (7). In [8], [10] the resulting phase shift is compared to $k\,\tilde{g}\,T^2$ where $\tilde{g}$ is the measured free fall acceleration of a falling macroscopic corner cube, i.e. $\tilde{g} = \left[1 + \beta_X^{(\text{corner cube})}\right] g$, also deduced from (7). In this class of theories the experiment is thus a test of the UFF, as it measures the differential gravitational acceleration of two test masses (Cæsium atom and corner cube) of different internal composition, with precision

$$\left|\beta_X^{(\text{Cs})} - \beta_X^{(\text{corner cube})}\right| \lesssim 7 \times 10^{-9}\,. \tag{11}$$

Note that this expression is equivalent to the one obtained by MPC in their recent paper [13] (p. 4), but different from the one obtained in the same paper (p. 3) for UCR tests.[2]

D. Analysis in the multiple Lagrangian formalism

In the second framework, which we call the "multiple Lagrangian formalism", the *motion* of test particles (atoms or macroscopic bodies) obeys the standard GR Lagrangian in a gravitational field, whereas the *phase* of the corresponding matter waves obeys a *different* Lagrangian. The MPC analysis in [10] belongs to this framework,[3] which raises extremely difficult problems.

The WEP is assumed to be valid for the motion of massive classical particles which thus obeys the standard Lagrangian of general relativity, $L_{\text{particle}} = L_{\text{GR}}$, i.e. to first order

$$L_{\text{particle}} = -m\,c^2 + \frac{GMm}{r_\oplus} - m\,g\,z + \frac{1}{2} m \dot{z}^2\,. \tag{12}$$

On the other hand the action integral to be used for the computation of the phase shift of the quantum matter wave is assumed to be calculated from the different Lagrangian

$$L_{\text{wave}} = -m\,c^2 + \frac{GMm}{r_\oplus} - (1+\beta) m\,g\,z + \frac{1}{2} m \dot{z}^2\,, \tag{13}$$

where the parameter β that measures the deviation from GR (i.e. $\beta = 0$ in GR) enters as a correction in the atom's gravitational potential energy. Integrating the Lagrangian (13) along the paths given by the Lagrangian (12) shows that $\Delta\varphi_S = \beta\, k\, g\, T^2$, where β takes the meaning of a redshift-violation parameter. In particular we notice that β could be "universal", in contrast with the parameter $\beta_X^{(a)}$ in (6) which depends on the type of atom (a) and on some internal energy X violating LPI. In the multiple Lagrangian formalism, because WEP is valid we can always test the value of β by comparing g as obtained from the free-fall of test bodies, to the phase difference cumulated by matter waves (as proposed by MPC [10]).

The most important problem of this formalism is that it is inconsistent to use a different Lagrangian (or metric) for the calculation of the trajectories and for the phases of the atoms or clocks. More precisely, it supposes that the fundamental Feynman path integral formulation of quantum mechanics, which is at the basis of the derivation of the phase shift in an atom interferometer [18], has to be altered in the presence of a gravitational field or could be wrong. Physically it amounts to making the distinction between the atoms when calculating their trajectories and the same atoms when calculating their phases, which is inconsistent as it

[2] In [13] atom interferometer tests set limits on the same parameter combination as classical UFF tests, namely $\beta_1 + \xi_1^{\text{bind}} - \beta_2 - \xi_2^{\text{bind}}$ ([13], p. 4), where the subscripts refer to test masses 1 and 2 (e.g. $1 =$ Cs and $2 =$ falling corner cube in [8] whilst $1 =$ Ti and $2 =$ Be in [1]), and ξ^{bind} refers to the nuclear binding energy of the test masses. On the other hand UCR tests in [13] set limits on either $\beta_1 - \xi_2^{\text{trans}}$ or $\xi_1^{\text{trans}} - \xi_2^{\text{trans}}$ ([13], p. 3), where ξ^{trans} is related to the atomic transition of the atom used in the clock and not to its nuclear binding energy.

[3] See "Methods" in [10], where they consider two scenarios. The first one clearly states that the trajectories are not modified whilst the atomic phases are. The second uses g' for the trajectories and $g(1+\beta)$ for the phases, which again corresponds to two different Lagrangians.

is the same fundamental matter field in both cases. Even more, the basis for the derivation of the atom interferometry phase shifts is unjustified because the Feynman formalism is violated. To remain coherent some alternative formalism for the atomic phase shift calculations (presumably modifying quantum mechanics) should be developed and used. More generally the multiple Lagrangian formalism supposes that the duality between particles and waves in quantum mechanics gets somehow violated in a gravitational field.

The above violation of the "particle-wave duality", implied by the multiple Lagrangian formalism, is very different from a violation of the equivalence principle in the ordinary sense. In this formalism a single physical object, the atom, is assumed to be described by two different Lagrangians, L_{wave} applying to its phase shift, and L_{particle} applying to its trajectory. By contrast, in tests of the equivalence principle, one looks for the modification of the free fall trajectories or clock rates as a function of composition or clock type. Thus, different bodies or clocks are described by different Lagrangians, but of course for any single type of body we always have $L_{\text{wave}} = L_{\text{particle}}$. The pertinent test-bed for equivalence principle violations is the modified Lagrangian formalism reviewed in Section II-C, which does not imply any particle-wave duality violation.

The second problem of theories in the multiple Lagrangian formalism and of the interpretation of MPC is the violation of Schiff's conjecture [29]. Indeed, we have assumed in this Section that the LPI aspect of the equivalence principle is violated but that for instance the WEP aspect remains satisfied. In a complete and self-consistent theory of gravitation one expects that the three aspects of the Einstein equivalence principle (WEP, LLI and LPI) are sufficiently entangled together by the mathematical formalism of the theory that it is impossible to violate one without violating all of them. The Schiff conjecture has been proved using general arguments based upon the assumption of energy conservation [21], [22]; it is satisfied, for instance, in the TH$\varepsilon\mu$ formalism [25]. A contrario, we expect that violating the conjecture leads to some breakdown of energy conservation. In the present case this translates into postulating two Lagrangians (12) and (13) for the same physical object. Explicit theories that are logically and mathematically consistent but still violate Schiff's conjecture are very uncommon (see [7] and references therein).

We conclude that the general statement of MPC according to which the atom interferometer measures or tests the gravitational redshift at the Compton frequency is incorrect. Instead, the interpretation of the experiment needs to be considered in the light of alternative theories or frameworks. Although one can consider particular alternative frameworks in which the statement of MPC could make sense, such frameworks raise unacceptable conceptual problems which are not at the moment treated in a satisfactory manner. In particular, they break the fundamental principles of quantum mechanics which are used for calculating matter wave phases. In most common and plausible theoretical frameworks the atom interferometry experiment tests the universality of free fall with the Compton frequency being irrelevant.

REFERENCES

[1] S. Schlamminger, et al., Phys. Rev. Lett. **100**, 041101 (2008).
[2] J. Williams, S. Turyshev, and D. Boggs, Phys. Rev. Lett. **93**, 261101 (2004).
[3] R. Pound and G. Rebka, Phys. Rev. Lett. **4**, 337 (1960).
[4] R. Vessot and M. Levine, Gen. Rel. and Grav. **10**, 181 (1979).
[5] L. Cacciapuoti and C. Salomon, Eur. Phys. J. Spec. Top. **127**, 57 (2009).
[6] S. Blatt, et al., Phys. Rev. Lett. **100**, 140801 (2008).
[7] C. M. Will, *Theory and experiment in gravitational physics* (Cambridge Univ. Pr., 1993).
[8] A. Peters, K. Chung, and S. Chu, Nature **400**, 849 (1999).
[9] S. Merlet, et al., Metrologia **47**, L9 (2010).
[10] H. Müller, A. Peters, and S. Chu, Nature **463**, 926 (2010).
[11] P. Wolf, L. Blanchet, Ch.J. Bordé, S. Reynaud, C. Salomon, and C. Cohen-Tannoudji, to appear in Class. Quant. Grav. (2011), arXiv:1009.2485 [gr-qc].
[12] M. Hohensee, B. Estey, F. Monsalve, G. Kim, P.-C. Kuan, S.-Y. Lan, N. Yu, A. Peters, S. Chu, and H. Müller, J. Phys. Conf. Ser. **264**, 012009 (2011).
[13] M.A. Hohensee S. Chu, A. Peters, H. Müller Phys. Rev. Lett. **106**, 151102 (2011).
[14] P. Wolf, L. Blanchet, Ch.J. Bordé, S. Reynaud, C. Salomon, and C. Cohen-Tannoudji, Nature **467**, E1 (2010).
[15] H. Müller, A. Peters, and S. Chu, Nature **467**, E2 (2010).
[16] S. Sinha and J. Samuel (2011), arXiv:1102.2587 [gr-qc].
[17] D. Giulini (2011), arXiv:1105.0749 [gr-qc].
[18] P. Storey and C. Cohen-Tannoudji, J. Phys. II France **4**, 1999 (1994).
[19] Ch.J. Bordé, Eur. Phys. J. Spec. Top. **163**, 315 (2008).
[20] Ch.J. Bordé, C.R. Acad. Sci. Paris **IV**, 509 (2001).
[21] K. Nordtvedt, Phys. Rev. D **11**, 245 (1975).
[22] M. Haugan, Ann. Phys. (N.Y.) **118**, 156 (1979).
[23] F. Belinfante and J. Swihart, Ann. Phys. (N.Y.) **1**, 168 (1957).
[24] T. Damour and A. Polyakov, Nucl. Phys. B **423**, 532 (1994).
[25] A. Lightman and D. Lee, Phys. Rev. D **8**, 364 (1973).
[26] Q. Bailey and V. Kostelecky, Phys. Rev. D **74**, 045001 (2006).
[27] V. Kostelecky and J. Tasson (2010), arXiv:1006.4106 [gr-qc].
[28] P. Wolf and P. Tourrenc, Phys. Lett. A **251**, 241 (1999).
[29] L. Schiff, Am. J. Phys. **28**, 340 (1960).

Continuous g monitoring with atom interferometry

Tristan Farah, André Clairon, Arnaud Landragin,
Sébastien Merlet and Franck Pereira Dos Santos
LNE-SYRTE
Observatoire de Paris, CNRS UMR 8630, UPMC
61 avenue de l'Observatoire, 75014 Paris, France
e-mail: sebastien.merlet@obspm.fr
web site: http : //syrte.obspm.fr/tfc/capteurs$_{\text{i}}$nertiels/

Anne Louchet-Chauvet
Laboratoire Aimé Cotton
CNRS UPR3321, Univ. Paris Sud
bâtiment 505, Campus Universitaire
91405 Orsay, France

***Abstract*— We present our cold atom interferometer using ^{87}Rb developed to perform absolute g measurement similar to the Stanford experiment [1] which used Cs. The device and the measurement protocol are described and an example of 12 days continuous measurement is presented. This measurement shows the capabilities of the instrument in the frame of a watt balance experiment, avoiding the use of environmental models to extrapolate the accurate value of g needed to link the Planck constant to the kilogram [2]. Limits in terms of sensitivity and accuracy are finally presented.**

I. Introduction

Based on techniques used on atomic clocks at LNE-SYRTE, we use atom interferometry to perform an absolute measurement of the gravitational acceleration g with ^{87}Rb free falling atoms [3]. Mainly used by geophysicists, absolute gravimeters allow geodesists to determine the distance between the geoid and the referential ellipsoid. Since the end of 80's and the proposition of Brian Kibble [4], absolute gravimeters are involved in fundamental metrology in the frame of watt balance experiments, participating in the kilogram redefinition and in which an absolute g determination is needed [5]. The most widely used absolute gravimeters measure the free fall of a corner cube using an optical interferometer [6]. The aging of the mobile part of these devices depends directly of the number of drops of the test mass. To preserve the instrument, most operators use a relatively low repetition rate and include dead time periods in g determinations which degrade the sensitivity of the measurement. In contrast, our instrument allows performing continuous g measurement at a large cycling rate of 3 Hz. It has recently participated in several comparisons with corner cube gravimeters [7], [8], [9].

We present here the device and the algorithm we use to perform continuous g measurement rejecting most of the systematic phase shifts. A measurement of 12 days is presented and the efficiency of the algorithm is discussed.

II. Atomic gravimeter CAG

The cold atom gravimeter (CAG) developed at LNE-SYRTE uses atom interferometry to perform a cyclic absolute measurement of g. At each cycle, a new cold cloud of Rb atoms is prepared in a UHV chamber, to be used as a test mass. During their free fall, these atoms undergo three stimulated Raman transitions that respectively separate, redirect, and recombine the atomic wave function, resulting in an atomic interferometer. The total phase shift between the two paths of this atomic interferometer depends on g, and scales with the square of the time interval between two consecutive Raman pulses. The phase difference is given by $\Delta\Phi_g = \vec{k}_{\text{eff}} \cdot \vec{g} T^2$ [10], where $\vec{k}_{\text{eff}} = \vec{k}_1 - \vec{k}_2$ is the effective wave vector with $|\vec{k}_{\text{eff}}| = |\vec{k}_1| + |\vec{k}_2|$ for counter-propagating beams.

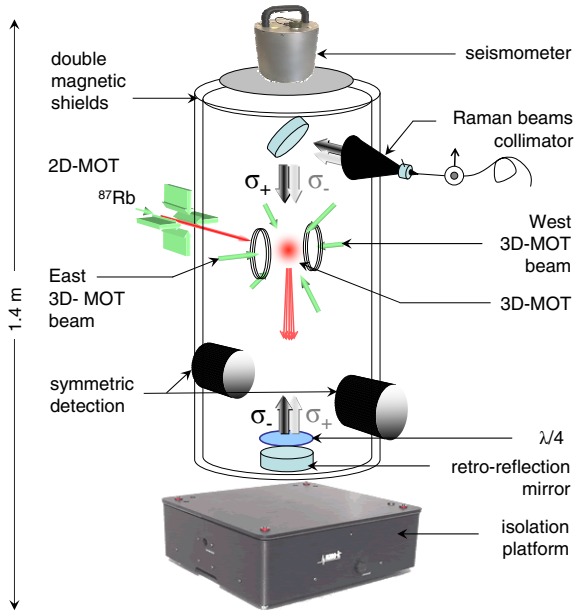


Fig. 1. Schematic of the cold atom gravimeter (CAG). A cold ^{87}Rb atomic sample is prepared by 2D- and 3D-magneto-optical trapping. After a short sub-Doppler cooling phase, where the atoms reach a temperature of $\sim 2\mu$K, the trapping beams are switched off and the atoms start their free fall in the vacuum chamber. During the free fall, a sequence of three stimulated Raman transitions is used to realize the atomic interferometer. These transitions are performed with two vertical, counter-propagating, circularly polarized laser beams, addressing the hyperfine transition of rubidium at 6.834 GHz via two-photon excitation. The phase shift at the output of the interferometer is deduced from a symmetric fluorescence measurement of the atomic state at the bottom of the vacuum chamber [11].

A more complete description of the CAG is given in Figure 1. The vacuum chamber lies on top of a passive isolation platform. The non-filtered vibration noise is measured with a Guralp seismometer rigidly attached to the vacuum chamber,

and is used to post-correct the atomic signal [3]. We use a mirror to retro-reflect the Raman beams, which is placed inside the chamber, leading to reduced optical wavefront aberrations. Furthermore, the fluorescence detection is performed with a double set of detectors placed symmetrically at opposite sides of the atomic cloud [12].

III. Measurement and protocol

The Raman lasers come from two extended cavity diode lasers based on the design described in [13], and amplified by two independent tapered amplifiers. They are phase-locked onto a low phase noise microwave reference source, which is swept according to $\omega_2 - \omega_1 = \omega_2(0) - \omega_1(0) + \alpha t$ in order to compensate for the Doppler shift induced by gravity. This adds a αT^2 term to the interferometer phase, which is eventually cancelled it for a perfect Doppler compensation. The central fringe of the interferometer thus corresponds to $\alpha_0 = \vec{k}_{\text{eff}} \cdot \vec{g}$. The value of g is therefore derived from the frequency chirp and from the mean effective wave-vector. We use a positive (resp. negative) frequency chirp to build an interferometer with $k_\downarrow$ (resp. $k_\uparrow$).

A maximal sensitivity to phase fluctuations is achieved when operating the interferometer at half fringe, which corresponds to $\Delta\Phi = \pm\pi/2$. The Raman phase is modulated by $\pm\pi/2$ so that the measurement is always performed at half fringe height, alternately on both sides to the central fringe as in an atomic fountain clock. From two consecutive measurements of the transition probability P_i and P_{i+1}, the phase error can be estimated. In practice, a correction $G\times(P_i - P_{i+1})$ is added at each cycle to α, in order to stir the chirp rate onto the central fringe. This realizes an integrator, whose time constant can be set to a few cycles by adjusting the gain G. This locking technique has the advantage of rejecting offset and contrast fluctuations, while preserving maximal sensitivity to phase fluctuations.

The measurement of g is shifted by many systematic effects. These systematic phase shifts can be sorted into two classes of error sources, either dependent ($\Delta\Phi_{dep}$) or independent ($\Delta\Phi_{indep}$) on the direction of $\vec{k}_{\text{eff}}$ pointing upwards or downwards ($\uparrow$ / $\downarrow$). The difference of the atomic phases accumulated along the two paths of the interferometer can thus be expressed as $\Delta\Phi_{\uparrow/\downarrow} = \pm\vec{k}_{\text{eff}} \cdot \vec{g}T^2 \pm \Delta\Phi_{dep} + \Delta\Phi_{indep}$. Taking that into account, the measurement procedure we use interleaves g measurements with $k_\uparrow$ and $k_\downarrow$:

Half difference and half sum of successive $\Delta\Phi_\uparrow$ and $\Delta\Phi_\downarrow$ measurements allow to separate $\Delta\Phi_{indep}$ from $k_{\text{eff}}gT^2 + \Delta\Phi_{dep}$. $\Delta\Phi_{indep}$ originates from effects related to perturbations of the internal degrees of freedom of the atoms (such as magnetic field gradient, one-photon light shift) and from radio-frequency phase shifts [14]. $\Delta\Phi_{dep} = \Delta\Phi_C + \Delta\Phi_{WF} + \Delta\Phi_{LS2}$ with $\Delta\Phi_C$ the Coriolis phase shift, $\Delta\Phi_{WF}$ the wavefront aberration phase shift and $\Delta\Phi_{LS2}$ the two-photon light shift. $\Delta\Phi_{LS2}$ is due to off-resonant Raman transitions which shift the atomic levels and therefore modify the hyperfine transition frequency [15]. As it scales linearly with the Rabi frequency of the Raman lasers Ω_{eff}, additional measurements realized with half Rabi frequency allow correcting for this effect by extrapolating to zero. Finally, the algorithm contains four configurations measuring respectively $\Delta\Phi_{\uparrow,\Omega_{\text{eff}}}$, $\Delta\Phi_{\downarrow,\Omega_{\text{eff}}}$, $\Delta\Phi_{\uparrow,\Omega_{\text{eff}}/2}$ and $\Delta\Phi_{\downarrow,\Omega_{\text{eff}}/2}$. They allow to determine $\Delta\Phi_g = k_{\text{eff}}gT^2$ only affected by the Coriolis and the wavefront distortion biases.

$$\begin{aligned}\Delta\Phi_g + \Delta\Phi_C + \Delta\Phi_{WF} &= \Delta\Phi_{\uparrow,\Omega_{\text{eff}}/2} - \Delta\Phi_{\downarrow,\Omega_{\text{eff}}/2} \\ &\quad - \frac{\Delta\Phi_{\uparrow,\Omega_{\text{eff}}} - \Delta\Phi_{\downarrow,\Omega_{\text{eff}}}}{2}\end{aligned}$$

While alternating $k_\uparrow$ and $k_\downarrow$ configurations does not degrade the sensitivity in the g measurement, the arithmetic involved when canceling the two photon light shift leads to a degradation in the short term sensitivity by a factor $\sqrt{10}$ with respect to a single configuration measurement. Indeed, as the sensitivity of a single configuration measurement $\sigma_{\Delta\Phi_g}$ is independent of the pointing direction of $\vec{k}_{\text{eff}}$ and of the Rabi frequency Ω_{eff}, summing quadratically the uncorrelated sensitivities $\sigma_{\Delta\Phi}$ of the right terms in the previous equation leads to a sensitivity on $\Delta\Phi_g$ of $\sqrt{5/2}\times\sigma_{\Delta\Phi}$. As each $\Delta\Phi_g$ is obtained with four measurement points, it corresponds to an effective cycle time of $4T_c$, where T_c is the cycle time of the experiment. This leads to a global degradation of the sensitivity of $\sqrt{5/2}\times\sqrt{4} = \sqrt{10}$ with respect to a single configuration measurement.

IV. Results

Figure 2 shows a continuous g measurement during 12 days in November 2010. It was performed in our laboratory dedicated to gravity measurements [2] close to the LNE watt balance project [5] nearby Paris. The points represented on this figure are obtained switching from one configuration to the next one every 100 shots. Each point thus represents 400 drops (about 170 s). Only three outliers out of 6 000 points over the 12 days measurement have been withdrawn from the figure. The measurements have been corrected from atmospheric pressure and polar motion effect on g [16] but not from tides, which they follow in very good agreement with a local tidal model [2]. The residuals obtained by correcting the data from tides are represented at the bottom of the figure for bins of 170 and 10 000 s. We see resolved temporal changes in the noise which we attribute to variations in human activity over nights and days or week-ends.

The Allan standard deviation of the residual signal is represented as black dots in Figure 3. It averages down as white noise in $\tau^{-1/2}$ and corresponds to an equivalent sensitivity at 1 s of $\sigma_g = 70$ μGal[1]. The sensitivity remains lower than 1 μGal after 5 000 s and flickers at the level of $0.7 - 0.8$ μGal, which is comparable to the accuracy of our local tidal model [2]. It is thus difficult to assess whether the long term stability is limited by our instrument or by the tidal model itself. On the same figure, the blue squares

[1] 1 μGal = 10^{-8} m.s^{-2}

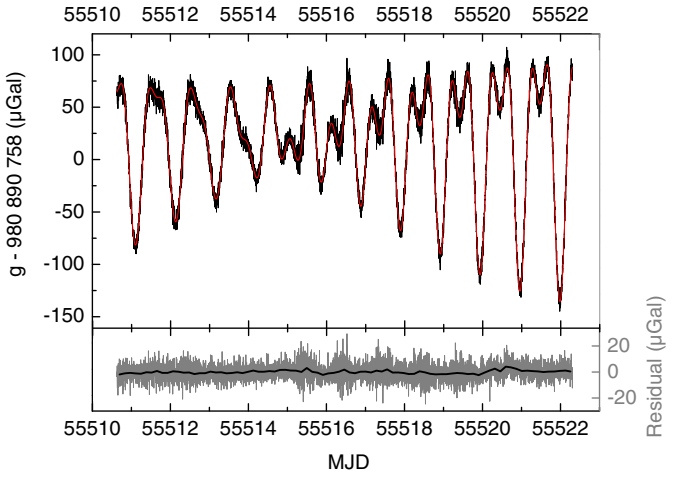


Fig. 2. Top, continuous g measurement from November 10^{th} to November 22^{nd} 2010 uncorrected from tides (bins of 400 drops). (Red solid line) Tide model. Bottom, residuals, in gray for bins of 400 drops (170 s), in black for bins of 10 000 s.

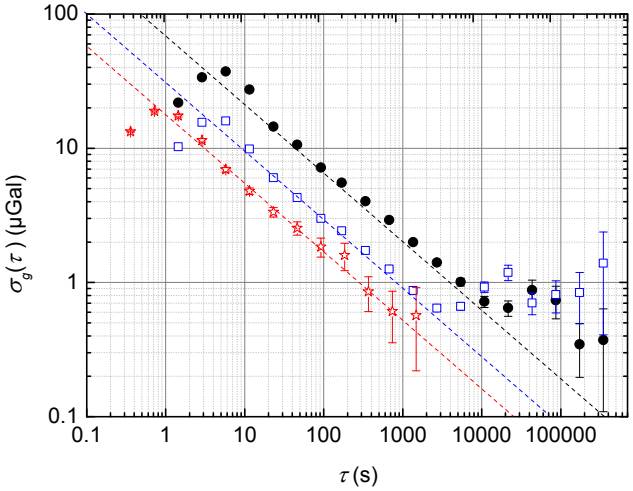


Fig. 3. Allan standard deviation of corrected g signals. Black dots: using the four configurations algorithm to reject the LS2. Blue squares: half difference of only two configurations alternating $k_\uparrow$ and $k_\downarrow$. It corresponds to a blue slope $\sqrt{5/2}$ lower than the black one. Red stars: using a single configuration measurement taken as continuous during about 3 h.

represents the Allan standard deviation of the half difference when alternating only two configurations $k_\uparrow$ and $k_\downarrow$. Note that for this plot we selected data only at full intensity, which amounts to taking only half of the data points, keeping the same effective cycle time of $4\ T_c$. The algorithm suppresses the fluctuations of $\Delta\Phi_{indep}$ and allows to obtain a sensitivity of $0.6\ \mu$Gal after $3\ 000$ s. Nevertheless the measurement is still sensitive to the two-photon light shift fluctuations represented by the bump on the red plot at about $20\ 000$ s. Finally the improvement of stability for long term measurement with four interleaved configurations clearly illustrates the efficiency of the two photon light shift rejection. We also represent as red stars the Allan standard deviation of the corrected signal obtained with a single configuration during three hours, considered here as continuous. The corresponding red slope representing the white noise should be $\sqrt{10}$ lower than the black one and the sensitivity as explained previously but it is even better as the data computed only include three hours of g measurements obtained during a night. Then the sensitivity at 1 s is $1.5\ \mu$Gal instead of $2.2\ \mu$Gal.

The total accuracy of the instrument is $5.2\ \mu$Gal, mostly due to the wave-front distortion bias [12]. It could be reduced if we limit the transverse motion of the cloud. Three comparisons with different free fall corner cube gravimeter [6], [17], [18], [19] performed from October 2009 to April 2010 [7], [8], [9] showed good agreements ($4.3 \pm 6.4\ \mu$Gal) and differences up to $15\ \mu$Gal.

V. Conclusion

We presented the LNE-SYRTE absolute atomic gravimeter, developed in the frame of the LNE watt balance experiment. It allows performing continuous g measurement at a high cycling rate of 3 Hz, avoiding the use of environmental models for the watt balance measurement sessions. The algorithm we use allows to suppress most of the phase shifts. The sensitivity of the instrument is $1\ \mu$Gal after $5\ 000$ s time measurement and flicker at $0.7\ \mu$Gal level that corresponds to the accuracy of our tidal model. The accuracy is 5 times bigger than the targeted accuracy we expect to reach when improving the uncertainty associated to the wave-front distortion correction which is the limiting effect of such a device. For that purpose, the next step consists in limiting the transverse motion of the cloud by cooling the atoms further down in a dipole trap or using horizontal Raman beams to limit the cloud expansion.

Acknowledgment

The research within this EURAMET joint research project leading to these results has received funding from the European Community's Seventh Framework Programme, ERA-NET Plus, under Grant Agreement No. 217257, and from IFRAF (Institut Francilien pour les Atomes Froids).

References

[1] M. Kasevich and S. Chu, *Atomic interferometry Using Stimulated Raman Transiitons*, Phys. Rev. Lett., vol. 67, 1991.

[2] S. Merlet, A. Kopaev, M. Diament, G. Genevès, A. Landragin, F. Pereira Dos Santos, *Micro-gravity investigations for the LNE watt balance project*, Metrologia, vol. 45, pp. 265-274, 2008.

[3] J. Le Gouët, T. E. Mehlstäubler, J. Kim, S. Merlet, A. Clairon, A. Landragin, F. Pereira Dos Santos, *Limits to the sensitivity of a low noise compact atomic gravimeter*, Appl. Phys. B, 92, pp. 133-44, 2008.

[4] B. P. Kibble, *Atomic masses and fundamental constants*, J. H. Sanders and A. H. Wapstra (Plenum, New York, 1976), vol. 5, pp. 545-551.

[5] G. Genevès, P. Gournay, A. Gosset, M. Lecollinet, F. Villar, P. Pinot, P. Juncar, A. Clairon, A. Landragin, D. Holleville, F. Pereira Dos Santos, J. David, M. Besbes, F. Alves, L. Chassagne and S. Topsu, *The BNM Watt Balance Project*, IEEE Trans. Instrum. and Meas., vol. 54, pp. 850-853, 2005.

[6] T. M. Niebauer, G. S. Sasagawa, J. E. Faller, R. Hilt, F. Klopping, *A new generation of absolute gravimeters*, Metrologia, vol. 32, pp. 159-180, 1995.

[7] ICAG'09 (International Comparison of Absolute Gravimeters) at BIPM in 2009. To be published.

[8] S. Merlet, Q. Bodart, N. Malossi, A. Landragin, F. Pereira Dos Santos, O. Gitlein and L. Timmen, *Comparison between two mobile absolute gravimeters: optical versus atomic interferometers*, Metrologia, vol. 47, L9-L11 (2010).

[9] A. Louchet-Chauvet, S. Merlet, Q. Bodart, A. Landragin, F. Pereira Dos Santos, H. Baumann, G. D'Agostino and C. Origlia, *Comparison of 3 Absolute Gravimeters Based on Different Methods for the e-MASS Project*, IEEE Trans. Instrum. and Meas, *in press*, July 2011.

[10] Ch. J. Bordé, *Theoretical tools for atom optics and interferometry*, C.R. Acad. Sci. Paris, t.2, Série IV, 509-30, 2001.

[11] Ch. J. Bordé, *Atomic interferometry with internal state labelling*, Phys. Lett. A, vol.140, pp. 10-12, 1989.

[12] A. Louchet-Chauvet, T. Farah, Q. Bodart, A. Clairon, A. Landragin, S. Merlet and F. Pereira Dos Santos, *Influence of transverse motion within an atomic gravimeter*, Submitted to NJP, 2011.

[13] X. Baillard, A. Gauguet, S. Bize, P. Lemonde, Ph. Laurent, A. Clairon and P. Rosenbusch, *Interference-filter-stabilized external-cavity diode lasers*, Optics Communications, vol. 266, 609-13, 2006.

[14] A. Peters, K. Y. Chung and S. Chu, *High-precision gravity measurement using atom interferometry*, Metrologia, vol. 38, 25-61, 2001.

[15] A. Gauguet, T. E. Mehlstäubler, T. Lévèque, J. Le Gouët, W. Chaibi, B. Canuel, A. Clairon, F. Pereira Dos Santos and A. Landragin, *Off-resonant Raman transition impact in an atom interferometer*, Phys. Rev. A, Vol. 78, 043615, 2008.

[16] Xu Guochang, *Science of Geodesy-I, Advances and Future Directions*, ISBN 978-3-642-11740-4, Springer, 2010.

[17] G. D'Agostino A. Germak, S. Desogus, C. Origlia, G. Barbato, *Reconstruction of the free-falling body trajectory in a rise-and-fall absolute ballistic gravimeter*, Metrologia, vol. 45, pp. 308-312, 2008.

[18] M.A. Zumberge, R.L. Rinker, and J.E. Faller, *A Portable Apparatus for Absolute Measurements of the Earths Gravity*, Metrologia, vol. 18, pp. 145-152, 1982.

[19] D. Schmerge and O. Francis, *Set standard deviation, repeatability and offset of absolute gravimeter A10-008*, Metrologia, vol. 43, pp. 414-418, 2006.

GaN: a Multifunctional Material Enabling MEMS Resonators Based on Amplified Piezoelectric Detection

Marc Faucher, Achraf Ben Amar, Bertrand Grimbert, Virginie Brandli, Matthieu Werquin, Lionel Buchaillot, Christophe Gaquière and Didier Théron.

Institute of Electronics, Microelectronics and Nanotechnology
Villeneuve d'Ascq, France
marc.faucher@isen.iemn.univ-lille1.fr

Yvon Cordier, Fabrice Semond
Centre de Recherches sur l'HétéroEpitaxie et ses Applications, Valbonne, France

Matthieu Werquin
MC2-Technologies,
Villeneuve d'Ascq, France.

***Abstract*— The properties of a new class of electromechanical resonators based on GaN are presented. By using the flexural modes of a doubly clamped beam, the two-dimensional electron gas (2-DEG) present at the AlGaN/GaN interface can be modulated by a field effect arising from the GaN buffer piezoelectricity. This leads to active piezoelectric transducers for which we show experimental results with detailed bias condition studies up to 10 MHz. Associated with modeling of the transduction physics, this allows explaining how the 2-DEG properties lead to the transconductance effect in the electromechanical domain.**

I. INTRODUCTION

Gallium nitride is the second most used semiconductor material after silicon, with applications driven by optoelectronics, RF devices and DC power converters. All these devices benefits from GaN-on-silicon increasing material quality [1, 2]. Moreover, GaN combines piezoelectric properties to a high sound velocity, which could make possible the co-integration of SAW devices with HEMTs amplifiers [3, 4]. Among high frequency devices, SAWS [5, 6] and bulk acoustic filters [7] have already been studied. Quasi-static MEMS devices have been investigated for studying the gauge factor, which points out the importance of transport interaction with piezoelectric effects [8-10]. Moreover, GaN is able to withstand high electric fields and high temperatures, thus being a very interesting material for sensors operating in harsh environments [11, 12]. With the advances in RF power electronics, GaN heterostructures have reached high electron mobility in two dimensional electron gas (2-DEG) that makes them attractive for exploring active piezoelectric devices for high frequency resonant sensors [13]. The concept of small signal amplification coupled to transduction of motion in resonators dates back to Nathanson [14]. Recently, this approach has been revisited in silicon, with active devices exploiting the piezoresistive effect inside transistors [15]. Here we present an overview of a new class of electromechanical resonators based on GaN. The main idea is to use the 2-DEG at the AlGaN/GaN interface and the piezoelectric properties of the heterostructure in order to convert piezoelectric polarization generated in the bulk into an amplified current. Active piezoelectric transducers with experimental results are demonstrated with detailed bias condition studies up to 80 MHz. In particular, the importance of the 2-DEG properties is emphasized for the understanding of the electromechanical transconductance.

II. MATERIAL AND PROCESS

A. Development of MEMS compatible epilayers

Today, electronic grade GaN material is obtained by heteroepitaxy on silicon carbide, sapphire, or silicon substrates. Typical epilayers stacks consist in a nucleation layer, followed by a GaN buffer where the dislocation density lowers from its bottom to the top. Given the temperature growth range (800-1000°C) and the substrate properties, large thermal stresses develops at cooling, thus preventing to grow more than 2-3 µm of high quality GaN onto silicon or silicon carbide. As far as MEMS applications are concerned, ilicon is the prominent choice for the substrate material since it allows taking advantage of existing high aspect ratio machining techniques such as deep reactive ion etching (DRIE). In addition, silicon when patterned by fluorinated plasma or by wet etching techniques exhibits a high selectivity over GaN. Moreover, GaN-on-silicon has been demonstrated on 200 mm wafers using MOCVD technique [17]. In MBE growth on silicon substrates, 2µm thick GaN can be obtained provided that the large thermal mismatch (115%) effect is mitigated by using AlN/GaN/AlN layers, thus reducing induced stress. Here, buried cracks appear in this region and a compressive mean stress is obtained in the thick GaN layer [18]. In order

This work was supported by the French National Agency (ANR) in the frame of its programme in Nanosciences and Nanotechnologies (MEMSGAN project n°ANR-08-NANO-023).

to overcome the stress issues due to the use of standard HEMT wafers,1 µm thick AlN/GaN buffer epilayers have been engineered for MEMS applications. Using *in situ* wafer curvature measurement [19] as well a systematic processing of test structures, the growth process was optimized. Epilayers with a mean tensile stress in the range 250-400 MPa with no buried cracks have been obtained (Figure 1).

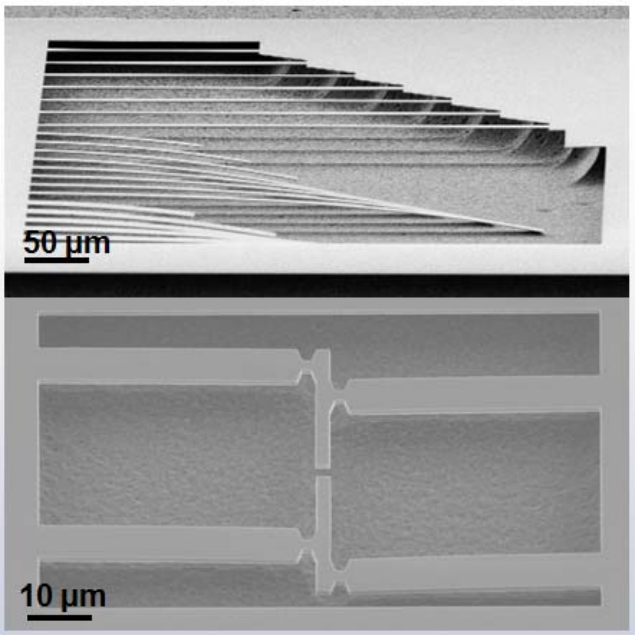


Figure 1. Micromechanical test structures of released AlGaN/GaN. The optimized growth conditions of the buffer allows to reach a tensile stress lower than 400 MPa.

B. *Simultaneous fabrication of HEMTs and R-HEMT transducers*

The fabrication of GaN resonators consists in processing HEMTs and R-HEMTs transducers first, and then defining mechanical parts. In our fabrication runs, we used two GaN epilayers: both are terminated by a 22 nm AlGaN barrier, the first one has a thickness 2.3 µm and a -2.3 V pinch-off voltage, the second is 1.8 µm thick and a -4.5 V pinch-off voltage. To make transistors and transducers ohmic contacts, we deposit a 350 nm Ti/Al/Ni/Au stack which is annealed during 30s at 850°C. Then a damage implant is used to define the 2-DEG electrical path for the R-HEMT channel extension and the actuator bottom electrode. A Mo/Au gate is formed by lift-off and 150 nm SiO_2 is deposited to passivate the devices. This layer is etched to make bridged connections to source and drain contacts as shown on Figure 2. This architecture allows using all the beam width as the R-HEMT drain extension, thus maximizing the signal. After the passivation step, a hard mask is deposited and patterned, followed by the etching of the entire GaN stack down to the silicon substrate. Then the resonators are released by isotropic etching of silicon using xenon difluoride process. Figure 3 shows a 50 µm long flexural resonator after release.

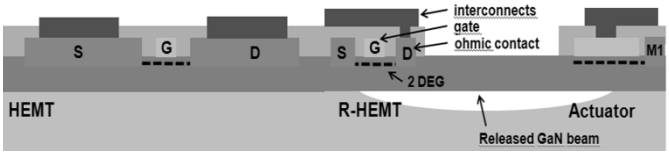


Figure 2. Architecture of the processed devices. Together with AlGaN/GaN/ HEMTs with large sizes (leftmost part), R-HEMTs transducers are fabricated. The Schottky gate and the actuator electrode are defined simultaneously. After passivation and metal pads deposition, micromechanical beams are etched in the GaN epilayer. Then, the resonators are released by Si substrate etching.

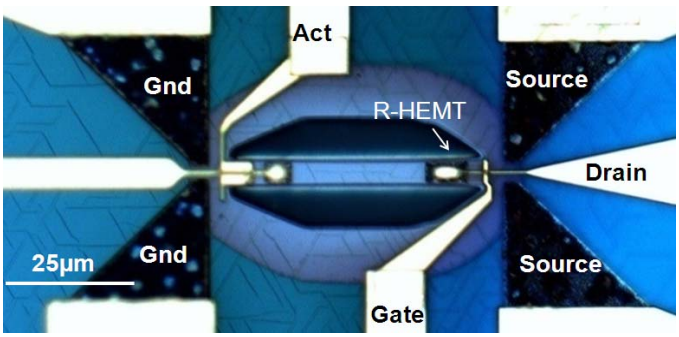


Figure 3. Optical image of a 50 µm long flexural GaN MEMS resonator with fully integrated transducers. The actuator is here the Schottky gate at left side of the beam. At right, a Resonant High Electron Mobility Transistor (R-HEMT) enables active mechanical detection.

III. ELECTRICAL AND MECHANICAL CHARACTERIZATION

A. *Transducers properties*

The testing of GaN MEMS devices involves optical, DC, and AC electrical measurements. The setup is based a vacuum testing chamber with coplanar probes enabling measurements from 77K to 330K. Three different bias tees are used at the actuator, drain, and source ports. The actuator is reverse biased in order to maintain the Schottky in the reverse regime, and the AC voltage is provided by port 1 of a network analyzer (Agilent 8753). The main effect observed in these resonators is shown on Figure 4. It consists of the increase of the motional current with applied drain bias, each maximum having first a linear dependence before 2V, and then followed by saturation (in practical, due to the high current density of 0.8A/mm enabled by the heterostructure, the small interconnecting metal bridge collecting the drain current withstands only voltages up to 4V).

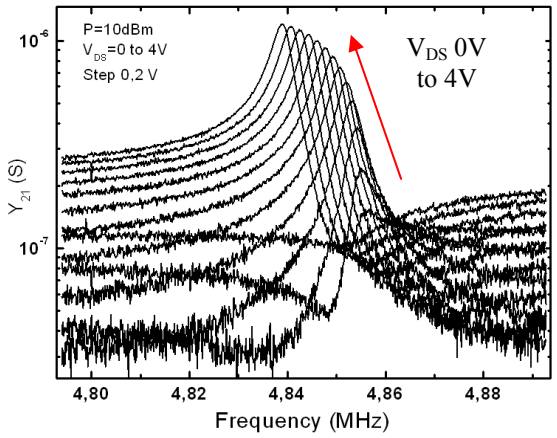


Figure 4. Optical GaN resonator response versus DC voltage applied on the drain at 0V on the gate. The amplification provided by the transducer enables motion detection with a ~20dB improvement as compared to a passive piezoelectric of similar size. The shift of the resonance frequency is attributed to the static piezoelectric strain due to the drain-source field, in addition to the portion of source lying inside released area.

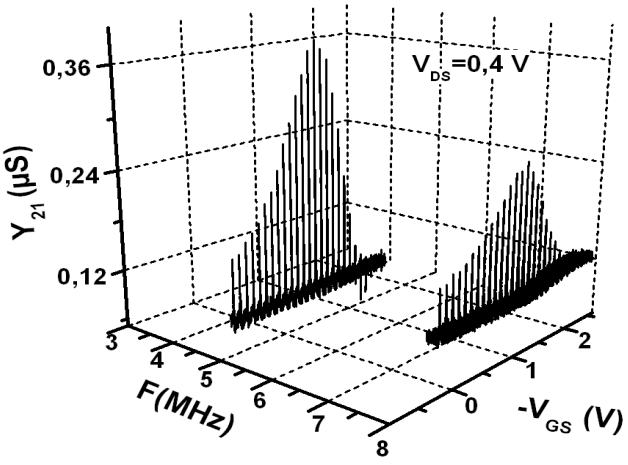


Figure 5. GaN resonator response versus the gate DC voltage (77K). The vibration modes are the 2nd and 3rd flexural of a 100 µm long doubly clamped beam. They evidence a variation of the resonator admittance that clearly denotes transconductance in the electromechanical domain.

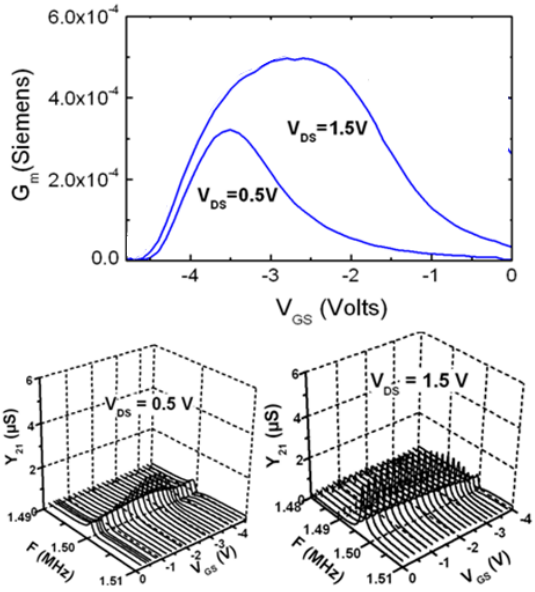


Figure 6. DC transconductance characteristics and electromechanical response versus gate in the linear and saturated regime (300K). A strong analogy is observed, evidencing the importance of 2-DEG mobility in the amplification-detection mechanism. The DC transconductance with other test structures allows to extract local and mean mobility, which are necessary to understand the response to the 2-DEG to additional carrier concentration originated by the piezoelectric buffer

The observation of amplification raises questions about the amplification mechanism and the physical effect occurring under the R-HEMT channel. A first calculation of the current in the linear regime shows that the corresponding sheet carrier density in the 2-DEG agrees with the order of magnitude of equivalent fixed piezoelectric charge generated in the GaN buffer and the heterostructure. A detailed study of the response, at cryogenic temperatures and on various resonators/vibration modes was performed. in Figure 5, we show at 77K the vibration spectra of a 100µm long resonator versus gate voltage. For both flexural modes of order 2 and 3, we observe that the R-HEMT response follows dependence very similar to the well known transconductance in AlGaN/GaN HEMTs. In particular, the gate completely switches-off the response near the pinch-off voltage (-2.4V), and a maximum is reached in the linear regime at a -1.8V value. This confirms that the electromechanical response occurs in the 2-DEG and not across parasitic conduction. On a different epilayer ($V_{pinch-off}$=-4.7V), the transconductance recorded under DC conditions confirms the analogy with the electromechanical admittance (Figure 6).

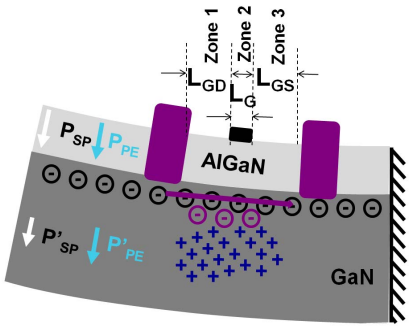


Figure 7. Extrinsic model of the R-HEMT. The active transducer includes 3 zones where the 2DEG can be influenced by the piezoelectric properties of the GaN buffer. Zones 1 and 3 are the access region with carriers having a mobility different from the 2-DEG located in the zone 2. With the gate bias, the carrier number and thus the local mobility can be tuned, resulting in a modulation of piezo-amplified transduction.

The HEMT DC transconductance (G_m) is known to display a pronounced maximum, since the main effect of the gate voltage is to control the mean carrier number in the channel. The mean carrier concentration turns in modifying the local mobility of the carriers, which is reflected in the extrinsic transconductance. The systematic observation of such maxima in the resonator response points towards the fact that the change of carrier concentration is the dominant mechanism in our case. We developed a quantitative approach based on 3 different zones in the R-HEMT showed in Figure 7. The idea is to take into account the response of the 2-DEG in the different zones, the zone 2 being the only one with mobilities controlled by the gate.

B. Minimum detectable amplitude with the R-HEMT transducer

Given the previous observation, we can bias the R-HEMT close to its maximal electromechanical transconductance and look for the minimum detectable signal in the saturation regime of the transistor. Figure 8. shows typical resolution reached with a 1.5 MHz resonator. Using laser Doppler vibrometry, the amplitude was calibrated so that the 2 mV excitation voltage corresponds to a floor of 1.2×10^{-13} m.Hz^{½}. Given that the stiffness constant is 15 N/m, this first result is promising for force measurements applications.

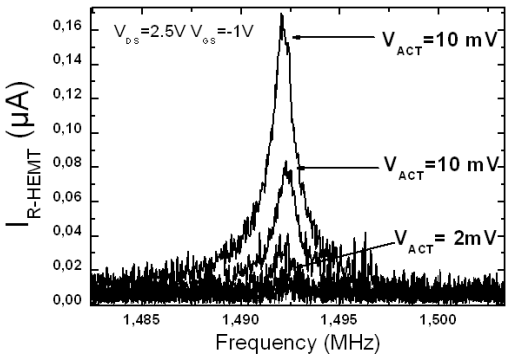


Figure 8. Performance of amplitude detection with the R-HEMT biased at saturation. The beam length is 100µm and a stiffness of 15 N/m. A resolution of 1.2×10^{-13} m. Hz ½ is reached.

IV. RESONATOR RESPONSE UP TO 100 MHZ

In order to evaluate the capability of R-HEMT transducers to detect motion at higher frequency, we acquired vibration spectra up to 100 MHz. First, a comparison between optical measurements and electrical response was performed (Figure 9). Together with FEM simulation, the conclusion is that R-HEMT senses all the flexural modes with out-of-plane motion.

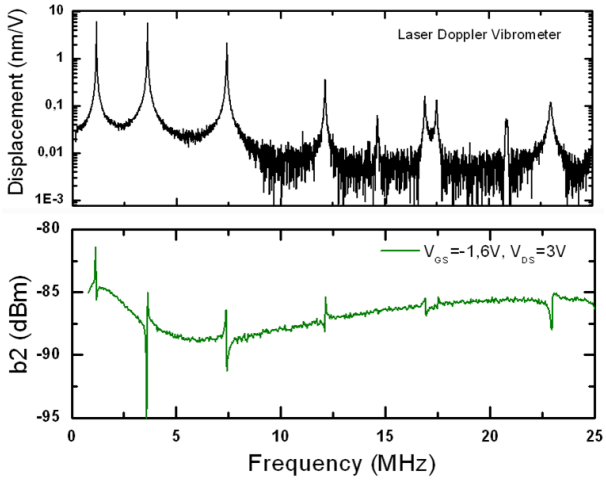


Figure 9. 1-25 MHz study of a 100µm long GaN resonator. The top curve represents the vibration amplitude recorded by a Laser Doppler Vibrometer, with the excitation provided by the piezoelectric 2-DEG actuator. The bottom curve is the power transmitted by the resonator measured at the R-HEMT drain. All the flexural out-of-plane modes are detected,.

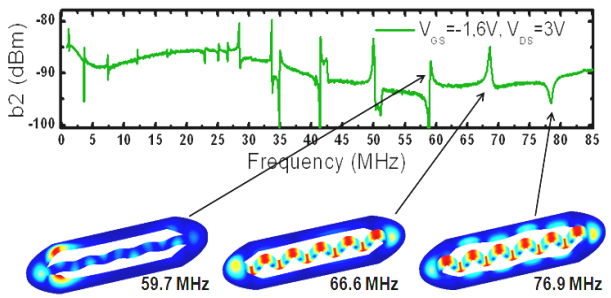


Figure 10. Electrical response up to 85 MHz. FEM simulations including the epitaxial stack and the exact dimensions of the released areas of the measured resonator allows to identify high order flexural modes, as well membrane-like modes involving peripheral motion.

At frequencies beyond the laser Doppler Vibrometer bandwidth, the transduction still evidences vibration modes. Finite Element Modeling allows identifying several high order flexural modes (Figure 10), and also confirms that in plane flexural modes are not detected by the R-HEMT. This is consistent with the interpretation that the 2-DEG screens the piezoelectric polarization, which leads to a charge density in the GaN buffer only when it has a non vanishing divergence.

V. Conclusion

We have presented the development of Gallium nitride MEMS resonators based on AlGaN/GaN heterostructures grown on silicon substrates. The use of piezoelectricity and 2-DEG in the R-HEMT device allows providing active transduction that evidences harmonic response up to 77 MHz on flexural beams.

Acknowledgment

We thank P. Tilmant, C. Boyaval, M. François, F. Vaurette for help with the e-beam processes. We also acknowledge S. Lepilliet, E. Delos, and D. Ducatteau for their help during measurements.

References

[1] J. Everts, J. Das, J. Van den Keybus, J. Genoe, M. Germain, and J. Driesen, "A high-efficiency, high-frequency boost converter using enhancement mode GaN DHFETs on silicon," Energy Conversion Congress and Exposition (ECCE), 2010 IEEE, 2010, pp. 3296-3302.

[2] F. Semond, P. Lorenzini, N. Grandjean, and J. Massies, "High-electron-mobility AlGaN/GaN heterostructures grown on Si(111) by molecular-beam epitaxy," Appl. Phys. Lett., vol. 78, pp. 335-337, 2001.

[3] [J. Pedros, R. Cuerdo, F. Calle, J. Grajal, J. L. Martinez-Chacon, and Z. Bougrioua, "6C-3 Field-Effect-Modulated SAW Devices on AlGaN/GaN Heterostructures," Ultrasonics Symposium, 2006. IEEE, 2006, pp. 273-276.

[4] T. Palacios, F. Calle, E. Monroy, J. Grajal, M. Eickhoff, O. Ambacher, and C. Prieto, "Nanotechnology for SAW devices on AlN epilayers," Materials Science and Engineering B, vol. 93, pp. 154-158, 2002.

[5] T. Pastureaud, A. Soufyane, S. Camou, S. Ballandras, D. Schenck, F. Semond, J. Desbois, and V. Laude, "AlN and GaN layers deposited on sapphire or silicon substrates: theory and experiment," Ultrasonics Symposium, 2000 IEEE, 2000, pp. 293-297 vol.1.

[6] K. Sun-Ki, P. Min-Jung, J. Cheol-Yeong, L. Jae-Hoon, C. Hyun-Chul, L. Jung-Hee, and L. Yong-Hyun, "Investigation of characteristics of SAW filter using undoped GaN epitaxial layer grown by MOCVD on sapphire substrate," Ultrasonics Symposium, 2001 IEEE, 2001, pp. 257-260 vol.1.

[7] V. J. Gokhale, Y. Shim, and M. Rais-Zadeh, "Observation of the acoustoelectric effect in gallium nitride micromechanical bulk acoustic filters," IEEE Frequency Cont. Symposium , 2010, pp. 524-529.

[8] T. Zimmermann, M. Neuburger, P. Benkart, F. J. Hernandez-Guillen, C. Pietzka, M. Kunze, I. Daumiller, A. Dadgar, A. Krost, and E. Kohn, "Piezoelectric GaN sensor structures," IEEE Electron Device Letters, vol. 27, pp. 309-312, 2006.

[9] A. D. Koehler, A. Gupta, C. Min, S. Parthasarathy, K. J. Linthicum, J. W. Johnson, T. Nishida, and S. E. Thompson, "Extraction of AlGaN/GaN HEMT Gauge Factor in the Presence of Traps," IEEE Electron Device Letters , vol. 31, pp. 665-667, 2010.

[10] J. Osvald, "Polarization effects and energy band diagram in AlGaN/GaN heterostructure," Applied Physics A: Materials Science & Processing, vol. 87, pp. 679-682, 2007.

[11] F. Medjdoub, D. Marcon, J. Das, J. Derluyn, K. Cheng, S. Degroote, N. Vellas, Gaquière, M. Germain, and S. Decoutere, "GaN-on-Si HEMTs above 10 W/mm at 2 GHz together with high thermal stability at 325°C," Microwave Integrated Circuits Conference (EuMIC), 2010, pp. 37-40.

[12] V. Cimalla, F. Niebelschütz, K. Tonisch, C. Foerster, K. Brueckner, I. Cimalla, T. Friedrich, J. Pezoldt, R. Stephan, M. Hein, and O. Ambacher, "Nanoelectromechanical devices for sensing applications," Sensors and Actuators B, vol. 126, pp. 24-34, 2007.

[13] A. M. Dabiran, A. M. Wowchak, A. Osinsky, J. Xie, B. Hertog, B. Cui, D. C. Look, and P. P. Chow, "Very high channel conductivity in low-defect AlN/GaN high electron mobility transistor structures," Appl. Phys. Lett., vol. 93, pp. 082111-3, 2008.

[14] H. C. Nathanson, W. E. Newell, R. A. Wickstrom, and J. R. Davis,"the resonant gate transistor" IEEE Trans. Electron Devices, vol. 14, p. 117, 1967.

[15] D. Weinstein and S. A. Bhave, "The Resonant Body Transistor," Nano Letters, vol. 10, pp. 1234-1237, 2010.

[16] A. R. Boyd, S. Degroote, M. Leys, F. Schulte, O. Rockenfeller, M. Luenenbuerger, M. Germain, J. Kaeppeler, and M. Heuken, "Growth of GaN/AlGaN on 200 mm diameter silicon (111) wafers by MOCVD," physica status solidi (c), vol. 6, pp. S1045-S1048, 2009.

[17] N.Baron, Y.Cordier, S.Chenot, P.Vennéguès, O.Tottereau, M.Leroux, F.Semond, and J.Massies, "The critical role of growth temperature on the structural and electrical properties of AlGaN/GaN high electron mobility transistor heterostructures grown on Si(111)", Journal of Appl. Phys. 105, 033701 (2009).

[18] Y.Cordier, N.Baron, F.Semond, J.Massies, M.Binetti, B.Henninger, M.Besendahl, T.Zettler, "In situ measurements of wafer bending curvature during growth of group-III-nitride layers on silicon by molecular beam epitaxy", J. Crystal. Growth (301-302) pp.71-74 (2007)]

[19] M. Faucher, B. Grimbert, Y. Cordier, N. Baron, A. Wilk, H. Lahreche, P. Bove, M. François, P. Tilmant, T. Gehin, C. Legrand, M. Werquin, L. Buchaillot, C. Gaquière, and D. Théron, Appl. Phys. Lett., vol. 94, p. 233506, 2009.

Hot Filament CVD Conductive Microcrystalline Diamond for High *Q*, High Acoustic Velocity Micromechanical Resonators

Mehmet Akgul, Robert Schneider, Zeying Ren, Gerry Chandler*, Victor Yeh, and Clark T.-C. Nguyen
Department of EECS, University of California, Berkeley, CA, USA
*sp3 Inc., Santa Clara, CA, USA

***Abstract*—A capacitively transduced micromechanical resonator constructed in hot filament CVD boron-doped microcrystalline diamond (MCD) structural material has posted a measured *Q* of 146,580 at 232.441 kHz, which is 3× higher than the previous high for *conductive* polydiamond. Moreover, radial-contour mode disk resonators fabricated in the same MCD film and using material mismatched stems, cf., Figure 1, exhibit a *Q* of 71,400 at 299.86 MHz, which is the highest *series-resonant Q* yet measured for any on-chip resonator at this frequency. The material used here further exhibits an acoustic velocity of 18,516 m/s, which is now the highest to date among available surface micromachinable materials. For many potential applications, the hot filament CVD method demonstrated in this work is quite enabling, since it provides a much less expensive method than microwave CVD based alternatives for depositing doped CVD diamond over large wafers (e.g., 8") for batch fabrication.**

I. Introduction

Frequency-selective resonators with on-chip *Q*'s >30,000 at GHz frequencies, if possible, would offer a paradigm shift in transceiver design by enabling narrow band low insertion loss RF channel-select filters capable of suppressing adjacent-channel interferers directly after the antenna in the receive chain, thereby significantly improving robustness and power consumption. Unfortunately, no existing room temperature on-chip resonator can yet achieve such *Q*'s at GHz frequencies.

Nevertheless, recent advances in capacitively transduced micromechanical resonator technology have yielded devices that draw ever closer to this goal, with *Q*'s reaching past 200,000 at HF [1], 160,000 at VHF [2] and 14,600 at 1.2 GHz [3]. Among structural material options for capacitive resonators, polydiamond is among the most compelling, since its acoustic velocity is considerably higher than that of other popular surface-micromachinable materials, surpassing polysilicon by 2.3× and SiC by 1.6× and allowing polydiamond resonators to achieve GHz frequency with larger, more easily definable dimensions. Furthermore, phonon-phonon interaction based loss theory predicts polydiamond to achieve the highest $f \times Q$ product at GHz frequencies among alternative materials, such as silicon, quartz, SiC and AlN [4][5]. Indeed, contour mode micromechanical disk resonators fabricated in polydiamond with (material-mismatched) polysilicon stems have already demonstrated very impressive *Q's* of 11,555 at 1.14 GHz and 55,300 at 498 MHz [6]. However, most previous work on high *Q* diamond resonators [6][7] employ microwave plasma CVD (MPCVD) tools [8] that are often constrained to deposit films over small areas, due to the high cost of generating a uniform plasma over large wafers. Hot filament CVD (HFCVD) [9] presents an alternative deposition technique that can uniformly deposit polydiamond over large area wafers at much lower cost. Unfortunately, previous attempts to employ HFCVD diamond material for micromechanical resonators have so far yielded vastly inferior performance, causing some to dismiss HFCVD as a viable option for MEMS-based resonators.

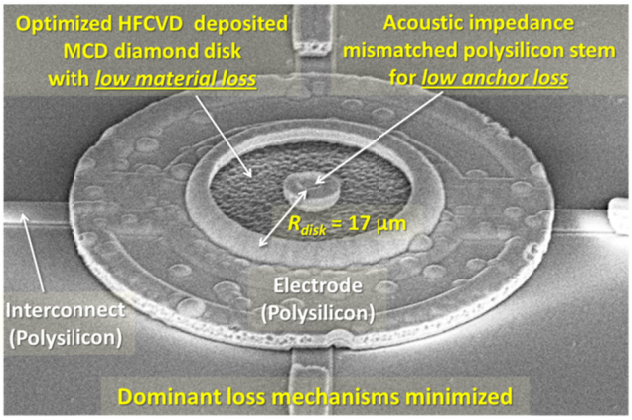


Figure 1: SEM image of a 300-MHz contour mode disk resonator with fully surrounding electrode used for high frequency HFCVD polydiamond resonator evaluation.

This work dispels such doubts by demonstrating HFCVD diamond MEMS resonators with unprecedented performance. Specifically, it presents a capacitive-comb transduced micromechanical resonator constructed in HFCVD boron-doped microcrystalline diamond (MCD) structural material with a measured *Q* of 146,580 at 232.441 kHz, which is 3× higher than the previous mark for MPCVD-based devices [7]. Furthermore, contour mode disk resonators fabricated in the same HFCVD film and using material-mismatched stems demonstrate a *Q* of 71,400 at 299.8 MHz, which is the highest series resonant *Q* measured at this frequency for an on-chip room temperature micromechanical resonator, recording an $f \times Q$ product of 2.13×10^{13} Hz that is on par with similar resonators fabricated using MPCVD diamond [6]. The material used here further exhibits an acoustic velocity of 18,516 m/s, which is now the highest to date among available surface micromachinable materials. It is very possible that this new polydiamond recipe can eventually enable an increase beyond the *Q*>30,000 at GHz frequencies required for next-generation

The authors would like to thank DARPA for supporting this work.

TABLE I. COMPARISON OF ACOUSTIC VELOCITIES AND THEORETICALLY PREDICTED QUALITY FACTORS AT 1GHZ AND 3GHZ

Material	v_a (m/s)	Q_{Theory} @ 1GHz[4]	Q_{Theory} @ 3GHz[4]
Diamond	18,500	380,000	380,000
SiC	13,300	320,000	116,000
Quartz	5,720	39,000	13,000
Si	8,500	36,000	36,000
AlN	10,970	8,700	3,400

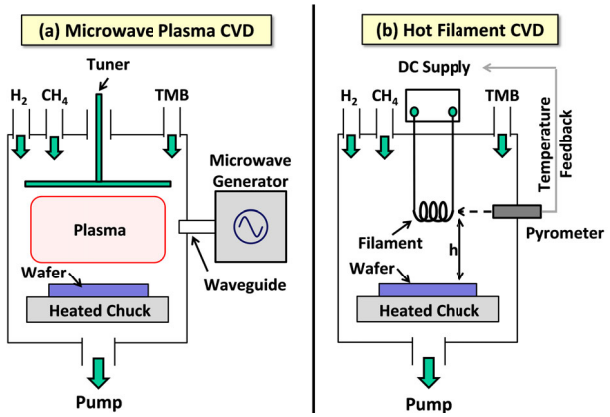


Figure 2: Schematic descriptions of (a) a microwave plasma CVD (MPCVD) system; and (b) a hot filament CVD (HFCVD) polydiamond deposition system.

RF channel-selecting transceiver front-ends [10].

For this and other potentially high volume applications, the hot filament CVD method demonstrated in this work is quite enabling, since it provides a much less expensive method than microwave CVD based alternatives for depositing doped CVD diamond over large wafers (e.g., 8").

II. THE DIAMOND FREQUENCY & QUALITY FACTOR ADVANTAGE

Anchor loss and intrinsic material loss due to phonon-phonon interactions are the two major Q limiting mechanisms for room temperature GHz frequency micromechanical resonators operated under vacuum. To alleviate the former, various anchor design practices, from the use of Bragg reflectors to quarter-wavelength support design, are commonly used to minimize energy lost to the substrate via anchors. Once anchor losses are sufficiently removed (if ever), intrinsic material losses remain to set the Q. While a complete understanding of intrinsic dissipation mechanisms in micromechanical resonators remains elusive, recent studies predict an $f \times Q$ product that is linear with frequency [4][5]—a prediction that deviates significantly from earlier belief in a frequency independent upper limit. Table I presents a comparison of predicted material Q's based on phonon-phonon energy dissipation calculations [4] at 1-3 GHz for common MEMS resonator materials, together with measured acoustic velocities. Polydiamond clearly stands out among materials in this table, having more than 3× the predicted Q and 1.4× the acoustic velocity of its closest competitors, perhaps making it the material of choice for high Q, high frequency micromechanical resonators.

III. POLYDIAMOND DEPOSITION TECHNIQUES

Figure 2 describes the two most common methods for thin-film deposition of polydiamond: microwave plasma CVD (MPCVD) [8] and hot filament CVD (HFCVD) [9]. As shown in the figure, an MPCVD tool supplies energy for CVD via microwave plasma, which must be uniform over the receiving wafer to ensure diamond film uniformity. Despite their previously mentioned prowess in attaining high Q at high frequency, the high cost of the technology needed to generate uniform microwave plasmas over large substrates often constrains MPCVD diamond tools to small area film depositions, making them less suitable for high volume batch production.

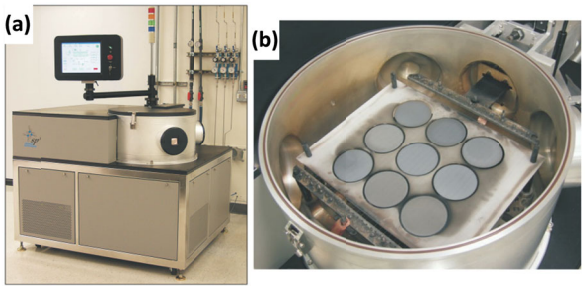


Figure 3: (a) Photo of the sp3 Model 650 HFCVD diamond deposition reactor used in this work. (b) Uniform film deposition over nine 4" wafers in a single batch.

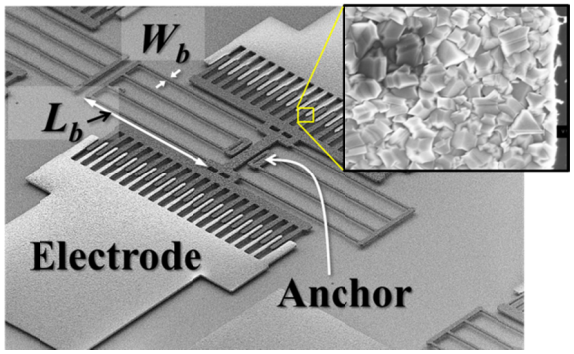


Figure 4: SEM of a fabricated comb-driven folded-beam resonator used to evaluate different HFCVD diamond deposition recipes.

In contrast, a hot filament CVD tool supplies energy for reaction via direct heat delivered by filaments long enough to extend over several wafer-sized substrates. The filaments are simple and easy to replace, and the tool is much simpler to build and operate, hence its much lower cost. Before this work, HFCVD had not been able to produce diamond micromechanical resonators with Q and acoustic velocity performance equaling that of the more expensive MPCVD, causing many would-be-adopters to shy away from the former.

This work uses an sp3 Model 650 HFCVD diamond deposition reactor, pictured in Figure 3, to explore more fully the efficacy of the hot filament approach and perhaps dispel prevailing biases against it. Since vibrating resonators for RF applications are of main interest, Q and acoustic velocity are the main metrics used to gauge performance in this work. Since the mechanisms for Q-limiting loss are often frequency and device geometry dependent, evaluations are done at both low and high frequencies, as is now described.

IV. HOT FILAMENT CVD POLYDIAMOND FOR HIGH Q LOW FREQUENCY MICROMECHANICAL RESONATORS

To evaluate the low frequency material Q as a function of HFCVD polydiamond deposition recipe, folded-beam electrostatic comb-driven resonators [11], such as depicted in Figure 4, were used. As amply covered in the literature, these devices

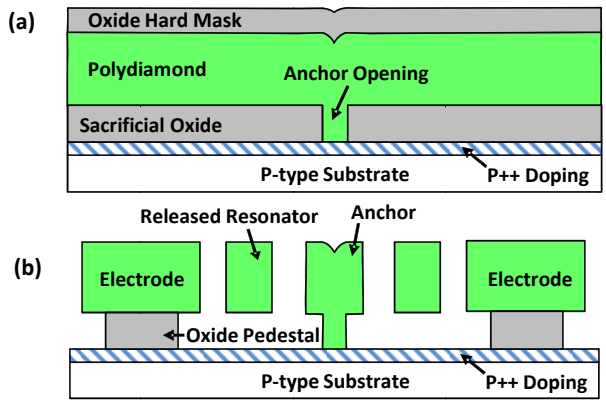


Figure 5: Two-mask comb-driven folded-beam resonator process flow cross-sections after: (a) diamond and hard mask deposition; and (b) diamond etch and HF release.

employ capacitive-comb transducers to linearize the ac voltage-to-force transfer function, as well as compliant folded-beam suspensions that alleviate the effects of residual stress. Such devices make for excellent gauges of intrinsic material Q at low frequencies, since their stiffnesses on the order of 1 N/m are quite low, making energy losses to the substrate through anchors quite small. If anchor losses can be made significantly smaller than the intrinsic material Q, then the measured Q more accurately gauges the intrinsic material Q.

A. Two-Mask HFCVD Polydiamond Fabrication Process

Figure 5 summarizes the quick two-mask fabrication process used to achieve folded-beam comb-driven resonators. The process starts with p++ doping of substrates via boron solid-source (to form a blanket ground plane), followed by a 2 µm LPCVD deposition of sacrificial oxide at 450°C. Anchors to the substrate are then lithographically defined and etched into the oxide, followed by HFCVD deposition of boron-doped polydiamond via recipes (one for each wafer) summarized in Table II. Here, an array of recipes is used with the intent of finding optimal conditions that maximize the material Q. For each recipe, the diamond deposition process consists of:

1. Seeding of nano-diamond particles by spin coating onto the wafers a liquid loaded with hydrocarbon nano-particles, prepared via ultrasonic agitation of fine diamond powder in DI water [12].
2. Flowing CH_4, which provides hydrocarbon radicals as the source of carbon for diamond deposition; and H_2, which dissociates into atomic hydrogen via thermal energy from the heating filaments and promotes the formation of sp3 diamond bonds over the undesired sp2 graphite phase.

After diamond deposition, a blanket LPCVD oxide is deposited to serve as a hard mask during etching of diamond. The diamond resonator structure and electrodes are defined by a second lithography step, then etched in an RIE chamber with 50 sccm O_2 and 2 sccm CF_4 flow at 50 mTorr pressure and 750W power. This recipe, with the help of the aforementioned oxide hard mask, creates vertical sidewalls. Finally, the wafer is diced, and individual dies are released in 49 wt. % HF via a timed etch that releases the resonator structures, but leaves their electrodes on oxide pedestals, firmly attached to the substrate. Figure 4 already presented the SEM image of a fabricated and released HFCVD diamond resonator, with a zoom in showing the 700 nm grain size achieved by recipe MCD^2.

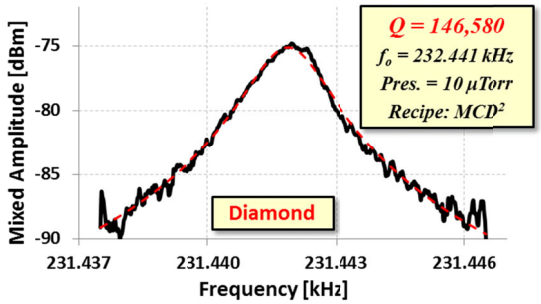


Figure 6: Measured frequency response for the polydiamond resonator using the best HFCVD recipe, showing a measured Q of 146,580 at 232.441 kHz.

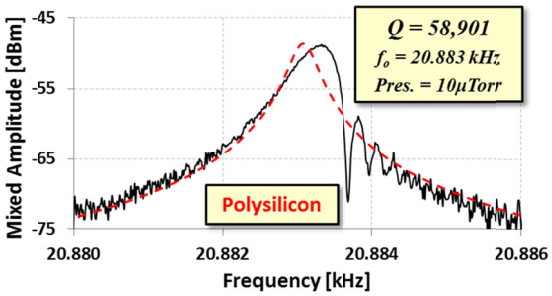


Figure 7: Measured frequency response (and Lorentzian fit with dashed lines) for a polysilicon resonator with identical geometry to that of Figure 6, showing a 3× lower Q of 58,901.

B. Measurement Results

Fabricated folded-beam resonators were measured using a two-port mixing measurement setup, as described in [13], at room temperature and under 10 µTorr vacuum to eliminate air damping. Figure 6 presents the measured frequency response for the resonator fabricated using the best HFCVD polydiamond recipe, achieving a measured Q of 146,580 at 232.441 kHz, which surpasses the previous high for conductive polydiamond [7] by 3×.

For comparison, polysilicon folded-beam resonators with geometries similar to those of the polydiamond versions were also fabricated using phosphorus-doped polysilicon structural material deposited via LPCVD at 615°C. Figure 7 presents the measurement for one such polysilicon resonator in a design nearly identical to that of Figure 4, exhibiting a measured Q of 58,901 at 20.883 kHz. The diamond version not only posts an 11.1× higher resonance frequency than the polysilicon one, but also bests its Q by 2.4×. Since anchor losses are higher for the stiffer diamond resonator, and if one assumes that other loss mechanisms (e.g., surface contamination) are comparatively negligible, these measurements suggest that the intrinsic material Q of HFCVD polydiamond structural material is at least 2.4× that of polysilicon in the kHz frequency range. Given that the previous mark for Q in conductive MPCVD diamond material is 47,900 [7], this measurement further indicates that it is possible to deposit lower loss materials using HFCVD than MPCVD, with the additional benefits of higher throughput and lower tool cost.

C. HFCVD Diamond Recipe Optimization

To determine the optimum HFCVD polydiamond deposi-

TABLE II. HOT FILAMENT DEPOSITED CVD DIAMOND PROPERTIES VS. DEPOSITION CONDITIONS

Parameter	NCD^1	NCD^2	MCD^1	MCD^2	Unit
CH_4 Concentration	2.7	1.5	1.5	1.0	%
Dopant TMB:CH_4	7500	1350	2250	675	ppm
Wafer Temp.	720	720	710	730	°C
Filament Temp.	2285	2200	1975	2010	°C
Quality Factor, Q	12,171	18,246	25,580	146,580	–
Acoustic Velocity	16,767	14,442	16,886	18,516	m/s
Resonance Freq., f_o	41.840	36.035	210.650	232.441	kHz
Young's Modulus	984	730	998	1198	GPa
Density, ρ	3500	3500	3500	3500	kg/m^3

tion conditions that maximize Q, resonators using variants of the deposition recipe were measured and compared. Table II compares the recipes and resulting measurements, showing a 12× improvement in Q between the initial (NCD^1) and final (MCD^2) optimized recipes. Each recipe shown in the table minimizes residual stress via proper adjustment of the filament and wafer temperature during deposition, and this makes possible the non-curled folded resonator beams shown in Figure 4. The column denoted as MCD^2 indicates the best deposition conditions for microcrystalline polydiamond, marked by the highest measured Q of 146,580 and acoustic velocity of 18,516 m/s for conductive polydiamond.

From the table, three important strategies for raising Q can be inferred:

1. Reduce the TMB:CH_4 ratio, which reduces the boron doping level.
2. Reduce the CH_4 flow, which reduces non-diamond carbon formation, thereby yielding a lower loss film.
3. Target microcrystalline polydiamond (MCD) films instead of nanocrystalline diamond (NCD), since the former seems to be able to attain higher Q, despite its surface roughness.

V. HOT FILAMENT CVD POLYDIAMOND FOR HIGH Q HIGH FREQUENCY MICROMECHANICAL RESONATORS

Because they have historically achieved the highest Q's among resonator designs operating past VHF, this work employs radial-contour mode disk resonators, such as shown in Figure 1, to evaluate the high frequency Q achievable by HFCVD diamond material. As shown in Figure 1 and in the cross section of Figure 9-(b), this device consists of a polydiamond disk suspended 700 nm above a polysilicon ground plane and anchored to the substrate by a polysilicon stem at its very center. Polysilicon electrodes surround the disk with an electrode-to-disk gap spacing of only 80 nm. All structural and interconnect materials are boron-doped, so are conductive.

The disk device operates similarly to the previous folded-beam resonators, where a combination of a dc-bias and ac excitation voltages are applied across the resonator-to-input electrode gap to excite resonance vibration; and currents flowing through the resulting time-varying dc-biased resonator-to-output electrode capacitive gap are sensed as a measure of the resonance displacement against frequency. As detailed in [6], the resonance frequency of a radial-contour mode disk is inversely proportional to radius and directly proportional to acoustic velocity, the latter of course being one of the major advantages of diamond.

The prospects of attaining Q's at high frequencies as high as seen for the low frequency folded-beam devices of the previous section are mediocre, at best. Indeed, the results of the previous section, although very impressive, only apply to low frequency resonators for which anchor losses are small and the intrinsic material Q is set primarily by TED [14]. At the much higher frequencies needed for RF applications, Landau-Rumer regime phonon-phonon interactions supplant TED as the dominant intrinsic material loss mechanism [4], and this changes the value of the theoretical maximum attainable Q. Perhaps more importantly, the higher stiffnesses needed to attain high frequency greatly accentuate the role of supports and anchors as conduits for energy loss, to the point where anchors generally dominate over other loss mechanisms when a high frequency resonator is operated under vacuum. Indeed, anchor losses have historically prevented measurement of intrinsic material Q at UHF frequencies.

Needless to say, the key to attaining a best estimate of the Q set by intrinsic HFCVD polydiamond material loss mechanisms at high frequency is to eliminate anchor losses.

A. Material-Mismatched Stem forMinimal Anchor Loss

Anchor loss has long been recognized as an important Q limiter on the micro-scale, from the very first MEMS-based resonators to reach VHF frequencies with high Q [15], to more recent rings [3], disks [6], and lamb wave structures [16] operating at GHz frequencies. Indeed, VHF free-free beam micromechanical resonators [15] were the first to employ quarter-wavelength supports attached at nodal points in the resonator mode shape to minimize support-derived energy losses—a technique still very much used in the latest generation of GHz devices [3], and still among the most effective at reducing anchor loss. To be sure, other strategies for reducing support losses have also been explored, including the use of Bragg reflectors [17] or photonic bandgap structures [18] between the resonator and substrate, but these have so far not been as effective as quarter-wavelength supports attached at nodal locations, at least from the perspective of maximum Q.

Ultimately, the above strategies to suppress anchor loss all seek to create large acoustic impedance mismatches at the resonator-anchor boundaries in an attempt to confine the acoustic energy within the resonator structure during resonance vibration, thereby preventing energy loss to the surroundings. Knowing this, and further recognizing the very large difference in characteristic acoustic impedance between polydiamond and other common micromachinable materials, the work of [6] used MPCVD polydiamond for the disk structure, but polysilicon for the stem, to effect a material mismatch between the disk and stem that reflects energy back into the disk structure, preventing the energy leakage that would otherwise occur. The degree of reflection can be modeled and designed analogously to electrical transmission lines to generate a non-zero reflection coefficient for acoustic waves at the resonator-anchor boundary, preventing energy from flowing into the stem anchor towards the substrate, as depicted in Figure 8. Figure 8 specifically illustrates how a stem made in the same material as the disk does little to impede energy flow; whereas a mismatched stem made in a material different from that of the disk suppresses energy loss to the substrate. This

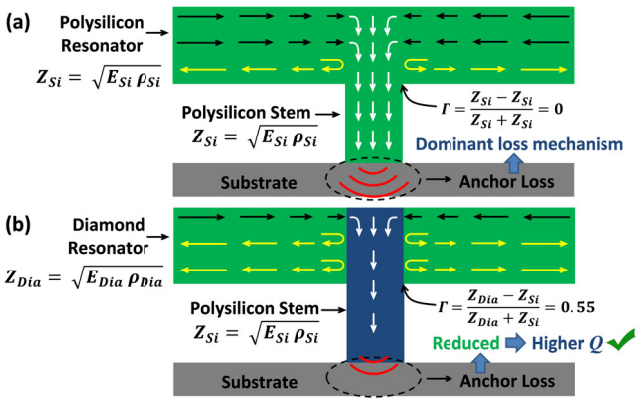


Figure 8: Schematics comparing energy losses to the substrate (indicated by arrows) for (a) a polysilicon disk with an impedance matched polysilicon stem; and (b) a polydiamond disk with a material-mismatched (so impedance-mismatched) polysilicon stem, where (b) loses much less energy.

strategy allowed the device of [6] to achieve measured Q's as high as 55,300 at 498 MHz, which corresponds to a whopping frequency-Q product of 2.75×10^{13}, setting a record for on-chip room temperature resonators at the time of its publication.

To best minimize competing loss mechanisms that might mask the intrinsic structural material Q, the current work borrows from that of [6] in its use of polydiamond for the disk structure, but polysilicon for the stem material.

B. Five-Mask Fabrication Process for High Frequency Contour Mode Disk Resonators

The fabrication process for disk resonators uses five masks and is nearly identical to that of [6], except for the use of an HFCVD diamond tool that allows conformal film deposition over multiple large area wafers at low cost per run, instead of MPCVD diamond, which in [6] was limited to single small area depositions. Figure 9 presents a very brief process flow. The process starts with p++ doping of wafers via boron solid-source, followed by subsequent LPCVD depositions of 2-µm LTO and 500-nm nitride insulation layers. 500 nm of polysilicon is then deposited and boron-doped via solid-source, then patterned and etched to define interconnects and ground planes. A 700-nm LTO layer is deposited as the bottom sacrificial layer, followed by diamond seeding to nucleate a subsequent HFCVD growth of 2-µm structural boron-doped microcrystalline (MCD) polydiamond using recipe MCD^2 from Table II. A 1.2-µm oxide layer is deposited via LPCVD using flow rates of 40 sccm DCS and 100 sccm N_2O at 835°C to serve as a hard mask during the subsequent etch step, and as a vertical spacer for the ensuing polysilicon electrode deposition. The disk resonator structure is then defined and etched to yield the cross-section shown in Figure 9-(a), after which 80 nm of HTO is LPCVD'ed to serve as a sidewall sacrificial oxide that defines the electrode-to-resonator gap spacing. The stem opening and electrode anchors are then lithographically defined and etched, each in separate sequences to allow separate optimization of tiny stems and considerably larger electrode contact vias. A 3-µm-thick boron-doped polysilicon deposition then follows that conformally fills the stem and electrode contact openings to create the aforementioned material-mismatched stems, plus electrodes over the resonator sidewalls. The final lithography and etch step defines the electrodes, and a subsequent liquid HF release frees the resonators,

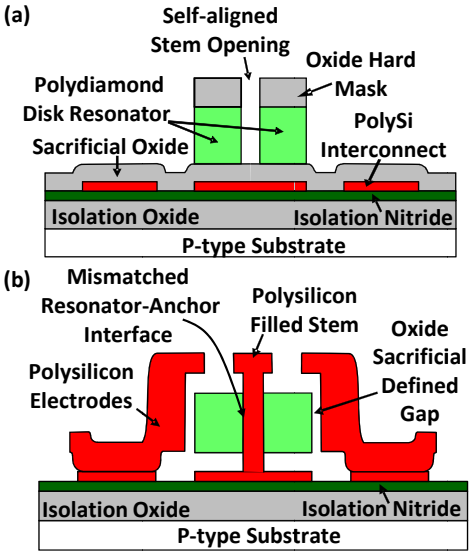


Figure 9: Five-mask disk resonator process flow cross-sections after: (a) diamond deposition and etch (b) electrode etch and HF release.

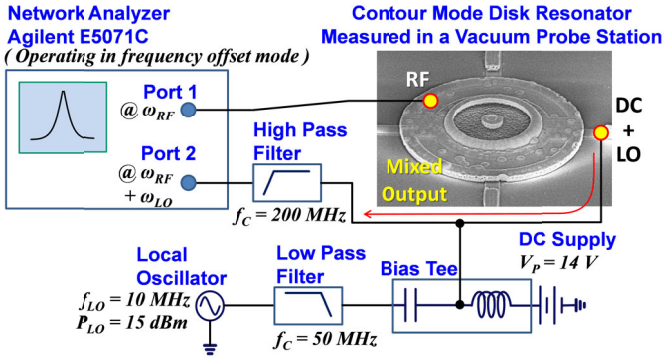


Figure 10: Schematic of the one-port mixing measurement setup used for contour mode disk resonator response measurements.

yielding cross-sections like that of Figure 9-(b).

Figure 1 already presented an SEM image of the resonator measured in this work, which has a disk radius R_{disk} = 17 µm, a 2 µm-diameter acoustic impedance mismatched polysilicon stem, and a fully surrounding electrode with an 80 nm electrode-to-resonator capacitive gap.

VI. Experimental Results

A Lakeshore model FWPX probe station provided 50 µTorr vacuum and probing to allow room temperature frequency characteristic measurements using the mixing measurement setup [13] depicted in Figure 10. Here, a bias tee combines a 10-MHz local oscillator signal with a DC bias voltage to generate the resonator voltage required for mixing measurement; and an Agilent E5071C Network Analyzer sources out of its port 1 an input RF signal that sweeps over a 100 kHz span centered at 10 MHz below the expected resonance frequency. As detailed in [13], the RF and LO signals mix via the square-law voltage-to-force transfer function of the input capacitive transducer, generating a force at their sum frequency, which equals the resonance frequency of the resonator, allowing port 2 of the network analyzer to sense the resonant peak without interference from feedthrough.

Figures 11 and 12 present measured frequency characteristics for HFCVD diamond and polysilicon disks, identically

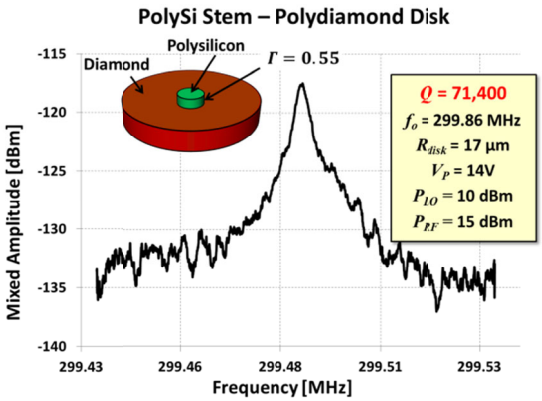


Figure 11: Measured frequency response (using mixing) for a contour mode disk resonator fabricated in HFCVD polydiamond and employing a material-mismatched polysilicon stem.

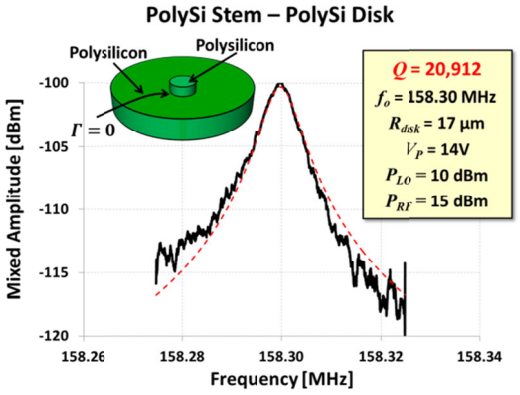


Figure 12: Measured frequency response for a contour mode disk resonator fabricated entirely in polysilicon, with identical geometry as the diamond disk of Figure 11.

dimensioned, both with polysilicon stems. The polydiamond disk with material-mismatched polysilicon stem exhibits a very high Q of 71,400 at 299.86 MHz, which is the highest series resonant Q measured at this frequency for an on-chip micromechanical resonator at room temperature. The polysilicon disk, on the other hand, not only resonates at a much lower frequency of 158.3 MHz, but since it lacks a material-mismatched stem, also posts a much lower Q of only 20,912, which is 3.4× lower than that of the acoustic impedance mismatched diamond device. The measured results clearly confirm the efficacy of stem-to-disk material-mismatching to suppress energy loss to the substrate, and provide resounding testament to the utility of HFCVD polydiamond material for high frequency micromechanical resonators.

VII. Conclusions

The very impressive Q's of 146,580 at 232.4 kHz and 71,400 at 299.86 MHz measured for the folded-beam and disk resonators of this work, respectively, are considerably higher than previously achieved on similar devices constructed using microwave CVD polydiamond (or any other material, for that matter) and now elevate hot filament CVD polydiamond as one of the (if not *the*) most compelling material to use for high frequency on-chip micromechanical resonators. The material used here further exhibits an acoustic velocity of 18,516 m/s, which is now the highest to date among available surface micromachinable materials. It is very possible that this new polydiamond recipe can eventually enable Q >30,000 at GHz frequencies, and work to demonstrate this continues. If achievable with adequately small motional impedance (which will take some work), a paradigm-shift in the design of next generation RF communication transceivers might be possible, such as described in [10], where channel-selecting RF filter banks enabled by Q's this high remove all interferers from the signal directed to demodulation electronics, allowing realization of a fast, low power frequency gating spectrum analyzer that in turn enables true software-defined cognitive radio.

References

[1] T. Mattila et al., "A 12 MHz micromechanical bulk acoustic mode oscillator", *J. of Sensors and Actuator A, vol. 101, no. 1-2, pp.1-9,2002.*

[2] Y. W. Lin, S. S. Li, Z. Ren, and C. T. C. Nguyen, "Low phase noise array-composite micromechanical wine-glass disk oscillator," in Electron Devices Meeting, 2005. IEDM Technical Digest, p. 4–281.

[3] Sheng-Shian Li, Yu-Wei Lin, Yuan Xie, Zeying Ren, and C. T.-C. Nguyen, "Micromechanical 'hollow-disk' ring resonators," in 17th *IEEE Int.Conf. on MEMS, 2004, Maastricht, N*etherlands, pp. 821-824.

[4] R. Tabrizian, M. Rais-Zadeh, and F. Ayazi, "Effect of phonon interactions on limiting the fQ product of micromechanical resonators," in TRANSDUCERS 2009, p. 2131–2134.

[5] V. B. Braginsky, et al., Systems with Small Dissipation, The University of Chicago Press, 1985.

[6] J. Wang, J. E. Butler, T. Feygelson, and C. T. C. Nguyen, "1.51-GHz nanocrystalline diamond micromechanical disk resonator with material-mismatched isolating support," 2004, p. 641–644.

[7] N. Sepulveda, D. M. Aslam, and J. P. Sullivan, "Polycrystalline Diamond RFMEMS Resonators with the Highest Quality Factors," *19th IEEE Int.Conf. on MEMS, 2006, Istanbul, p. 238–241.*

[8] S. Yugo, T. Kanai, T. Kimura, and T. Muto, "Generation of diamond nuclei by electric field in plasma chemical vapor deposition," Applied Physics Letters, vol. 58, no. 10, p. 1036, 1991.

[9] S. Matsumoto, Y. Sato, M. Tsutsumi, and N. Setaka, "Growth of diamond particles from methane-hydrogen gas," Journal of Materials Science, vol. 17, no. 11, p. 3106–3112, 1982.

[10] C. T.-C. Nguyen, "Integrated micromechanical RF circuits for software-defined cognitive radio", 26th Symposium on Sensors, Micromachines & App. Sys., Tokyo, Japan, Oct. 15-16, 2009, pp. 1-5.

[11] W.C. Tang, T.-C.H. Nguyen, M.W. Judy, and R.T. Howe, "Electrostatic-comb drive of lateral polysilicon resonators," Sensors and Actuators A: Physical, vol. 21, Feb. 1990, pp. 328-331.

[12] M.P. Everson, "Studies of nucleation and growth morphology of boron-doped diamond microcrystals by scanning tunneling microscopy," J. of Vacuum Sci. & Tech. B, vol. 9, May. 1991, p. 1570.

[13] A.-C. Wong and C. T.-C. Nguyen, "Micromechanical Mixer-Filters," Journal of Microelectromechanical Systems, vol. 13, no. 1, pp. 100-112, Feb. 2004.

[14] R.N. Candler, A. Duwel, T.W. Kenny, et al., "Impact of geometry on thermoelastic dissipation in micromechanical resonant beams", J. of Microelectromechanical Systems, vol. 15, Aug. 2006, pp. 927-934.

[15] K. Wang, Y. Yu, A.-C. Wong, and C. T.-C. Nguyen, "VHF free-free beam high-Q micromechanical resona-tors", 12th International IEEE MEMS Conference, Orlando, Florida, Jan. 17-21, 1999, pp. 453-458.

[16] T.-T. Yen, A.P. Pisano,et al. ,"Characterization of aluminum nitride lamb wave resonators operating at 600°C for harsh environment RF applications", 23rd IEEE International Conference on MEMS, 2010

[17] C. Chung, Y. Chen, C. Cheng, C. Wei, K. Kao, "Influence of surface roughness of Bragg reflectors on resonance characteristics of solidly-mounted resonators," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 54, Apr. 2007, pp. 802-808.

[18] Vlasov Yu.A., Norris D.J., Bo X.Z., Sturm J.C., "On-chip assembly of silicon photonic band gap crystals," Quantum Electronics and Laser Science Conference, 2002, vol., no., pp. 116- 117, 2002

Very Low-Loss High Frequency Lateral-Mode Resonators on Polished Ultrananocrystalline Diamond

Hediyeh Fatemi and Reza Abdolvand
School of Electrical and Computer Engineering
Oklahoma State University
Tulsa, Oklahoma, USA
hediyeh.fatemi@okstate.edu

Hongjun Zeng and John Carlisle
Advanced Diamond Technologies Inc.
Romeoville, Illinois, USA
zeng@thindiamond.com

***Abstract*— For the first time, high frequency (~900MHz) lateral-mode thin-film piezoelectric-on-diamond resonators are reported with insertion loss values as low as 2.6dB when terminated by 50Ω impedance. In this work, the surface of the as-deposited ultrananocrystalline diamond (UNCD) is polished to promote the direct growth of highly oriented c-plane AlN which is critical in enhancing the electromechanical coupling and therefore reducing the motional impedance of the resonator. The frequency of resonators fabricated on 3-4um UNCD with a 930GPa Young's modulus is measured to be about 2x higher than similar devices fabricated on 5um of silicon. These resonators exhibit a low temperature coefficient of frequency (TCF) of -9.6 ppm/°C and have a great potential for being utilized in channel-select high frequency filters [1].**

I. INTRODUCTION

There is a growing demand for miniature and low-cost resonators for several applications such as stable and low-power oscillators, multi-channel filters, and sensors. Lateral-extensional piezoelectric resonators exhibit relatively low motional impedance and can operate at atmospheric pressure as opposed to capacitive resonators [2], and they could enable single-chip multi-standard radio frequency front ends and offer smaller footprint compared to FBARs.

One type of such piezoelectric resonators is the class of thin-film piezoelectric-on-substrate (TPoS) devices which are comprised of a piezoelectric layer sandwiched between two metal layers through which an electric field is applied to excite and sense the resonance. These layers are stacked on top of a substrate (Fig. 1). The resonance frequency of a lateral-extensional resonator is defined by its lateral dimension or in the case of higher harmonics by the finger-pitch which is the center-to-center distance of two adjacent electrodes as denoted in Fig. 1. Such resonators usually offer higher quality factors and better power handling relative to purely piezoelectric resonators [3, 4]. Another advantage of having a substrate is the ability to extend the frequency without compromising the ease of fabrication by choosing a substrate with high Young's modulus (E).

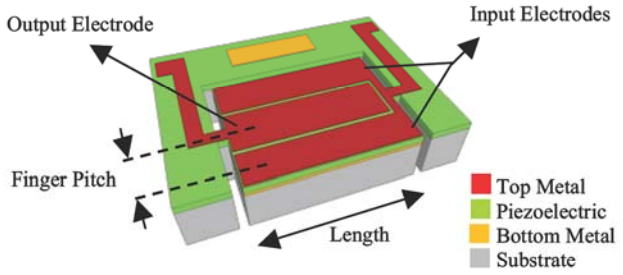


Figure 1. 3D structure of a thin-film piezoelectric-on-substrate resonator.

Diamond has gained much popularity in MEMS community essentially due to its high Young's modulus and low dissipation properties [5]. Thin-film piezoelectric-on-diamond (TPoD) bulk-lateral-mode resonators with 80% higher frequency than the same devices fabricated on silicon have shown much promise in scaling the resonance frequency beyond the limits achievable by silicon [6]. However, the orientation of the piezoelectric film is critical to retain the low insertion loss performance of the resonators. In the earlier work, a polished oxide layer had to be deposited on the nanocrystalline diamond substrate to promote the growth of a highly-oriented piezoelectric film on the rough surface of the substrate. This was the bottleneck for realization of very high frequency resonators since the presence of the oxide would significantly alter the mode shape and lower the quality factor and the frequency [6]. In this work the surface roughness of diamond substrate is substantially reduced by depositing ultrananocrystalline diamond (UNCD) followed by standard chemical mechanical planarization (CMP). Such smooth surfaces allow for direct sputtering of metal/piezoelectric/metal heterostructures with a large coupling coefficient that enable monolithic integration of dispersed-frequency resonators.

This paper describes the work to optimize the properties of UNCD films used to make the TPoD devices as well as the integration of high-quality AlN films on smooth (~1 nm rms roughness) UNCD films achieved via wafer-scale polishing. Next, the design procedure to optimize the performance of the TPoD resonators is discussed and measurement results are presented.

This work was funded by national science foundation under grant number **#0930676.**

II. FABRICATION

A. Development of UNCD Films

The acoustic velocity is a function of Young's modulus (E) and mass density (ρ) as follows (1):

$$\alpha = \sqrt{E/\rho} \qquad (1)$$

where α symbolizes the acoustic velocity. Since the frequency of a resonator built on a diamond substrate is directly proportional to α of the diamond, it is critical that the Young's modulus of the UNCD be consistently high and uniform across the wafer.

Standard UNCD films have Young's modulus from 650 GPa to 750 GPa, yielding up to 1.6 times increase in frequency compared to silicon-based devices designed with the same lateral dimensions. The challenge in increasing the modulus closer to the largest theoretical value possible (~1200 GPa) is maintaining as small of a grain size as possible to retain the as-deposited smoothness of the material so that CMP-processing is still capable of reducing the rms roughness to about 1 nm.

The UNCD films used for fabricating TPoD devices in this work were produced using hot filament chemical vapor deposition (HFCVD) technique. To increase E, the deposition temperature was increased from about 650 °C to 810 °C by increasing the temperature of the filaments during the deposition step. As depicted in Fig. 2, increasing the deposition temperature yields a clear increase in Young's modulus from 640 GPa (660 °C) to 1028 GPa (810 °C). On the other hand, the residual stress of the as-deposited film is reduced from about 300-400 MPa compressive to about 100 MPa tensile even for films as thick as 5-6 microns. However, increasing the growth temperature also increases the average grain size and the surface roughness. For higher deposition temperatures, films had rms roughness values around 60 nm rms as presented by AFM data in Fig. 3. Such high rms roughness values make the integration of high quality AlN films on the diamond very challenging due to the crystallographic orientation disorder that occurs at the AlN/diamond interface. A highly disordered AlN film results in very poor average piezoelectric coefficient and therefore an inferior insertion loss (IL) for the resonators. For these rougher films a two-stage polishing process was developed in which a more aggressive slurry and greater applied force was used initially to bring the roughness down to about 15 nm,

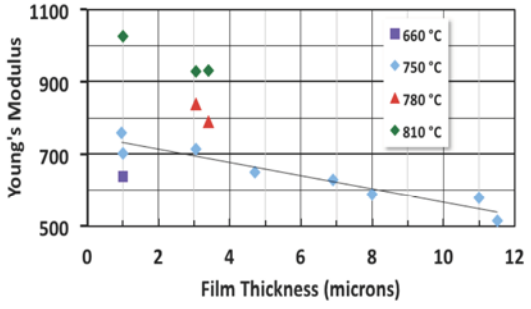


Figure 2. Dependence of UNCD Young's modulus on the deposition temperature.

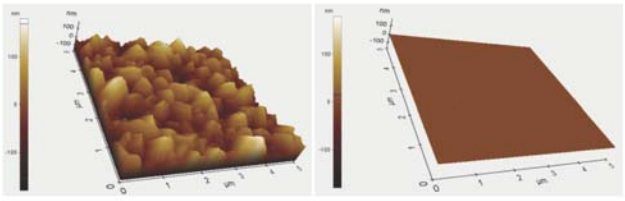


Figure 3. AFM image of high Young's modulus (950 GPa) diamond film surface (a) before and (b) after CMP showing less than 1nm final rms roughness.

after which the CMP process was performed to further reduce the roughness to less than 1nm (Fig. 3).

After UNCD deposition and polishing, the stack of Mo (100 nm)/ AlN (500 nm)/ Mo (100 nm) was sputtered (Fig. 4). AlN is used as the piezoelectric layer due to superior process compatibility, relatively high acoustic velocity and high electrical resistivity. Molybdenum is chosen as the metal layer because of the small lattice and thermal expansion coefficient mismatch with AlN in addition to low acoustic-loss. Fig. 5 represents the X-ray rocking curve of the deposited AlN film with a full-width half maximum (FWHM) of about 3 degrees. The quality of the film was very close to that of the c-axis aligned AlN deposited on polished single crystal silicon substrates for which the typical FWHM is in the range of 2.0-2.5°.

B. Fabrication Process

The schematic flow diagram of the fabrication process is shown in Fig. 6. After depositing UNCD film and the bottom metal layer on the polished surface of the silicon wafer, the bottom metal is patterned by dry-etching to remove the areas underneath the pads and tracks in order to reduce the parasitic capacitances. Then the AlN and the top metal layer are sputtered, and the top metal is patterned. Then the AlN is wet-

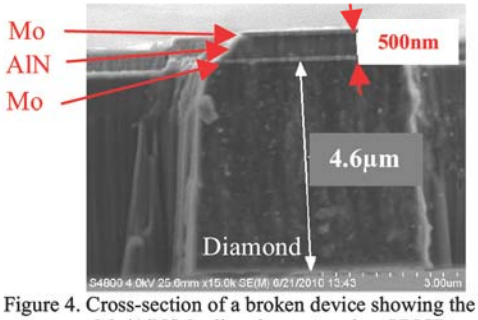


Figure 4. Cross-section of a broken device showing the stack of Mo/AlN/Mo directly sputtered on UNCD.

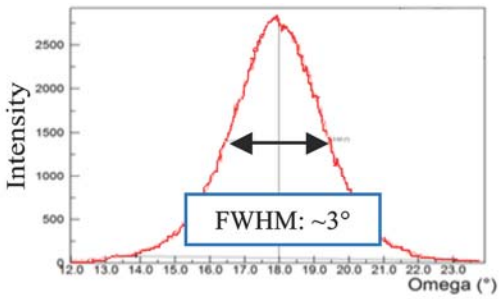


Figure 5. Rocking curve of the AlN film deposited on UNCD film with FWHM of 3°.

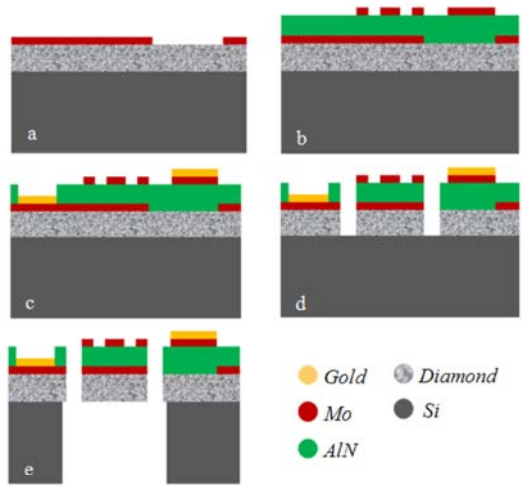


Figure 6. The schematic of the fabrication process flow.

etched in a Tetramethylammonium hydroxide (TMAH) solution to create access to the bottom metal followed by sputtering Au on the electrode areas to reduce ohmic losses. The AlN/Diamond stack is then etched down to the Si substrate in an inductively coupled plasma etcher (AlN in Cl_2 and Diamond in O_2/CF_4 plasma). Finally, devices are released by dry etching silicon from the backside in a deep reactive ion etching chamber.

III. Design Concept

In order to reduce the motional impedance of the resonator or particularly the insertion loss, the actuation area on the resonator should be increased (i.e. resonator's lateral dimensions). This can be obtained either by increasing the number of electrode fingers on the resonator or in other words designing higher order harmonics, or by increasing the length of the resonator. But the drawback in designing a larger resonant structure is that the structure supports a higher number of strong spurious modes that are close to the target frequency. In the simulated frequency response of Fig. 7, examples of such spurs and their corresponding mode-shapes for a 5th order harmonic TPoS resonator are depicted. As established by our group earlier, by adding multiple support tethers to the resonator structure the spurious modes can be effectively suppressed [7]. Here, the same technique was employed to improve the resonator performance.

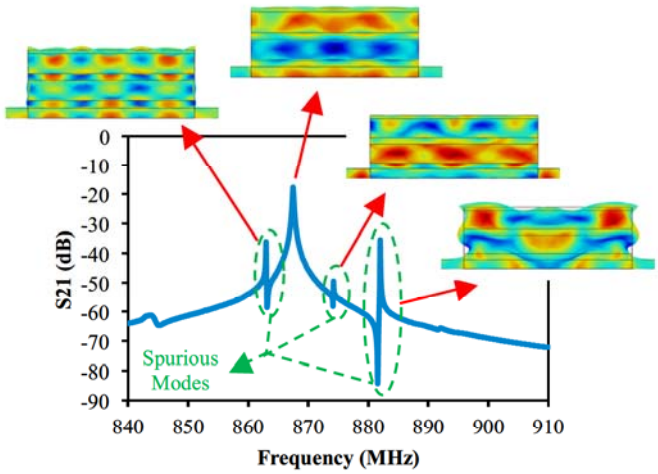


Figure 7. Simulated frequency response for a TPoD resonator designed for excitation of the 5th harmonic mode. The insets show the mode-shapes of the target mode as well as the neighboring spurious modes.

IV. Experimental Results

Frequency response measurements were performed using an Agilent E8358A PNA Network Analyzer and a pair of GSG probes. Prior to measurements, a short-open-load-thru (SOLT) calibration was carried out on a reference substrate.

To evaluate the performance of TPoD resonators, a multi-tethered 21st harmonic resonator was fabricated on two diamond substrates with different Young's modulus (650 and 930GPa) and also on a 5µm thick silicon-on-insulator (SOI) substrate. The typical measured frequency responses are plotted in Fig. 8. The measured frequency of the TPoD resonators is 1.6x (for 650GPa UNCD) and 2x (for 930GPa UNCD) the frequency of the same design on 5µm thick silicon substrate.

Following the discussion in the previous section, a

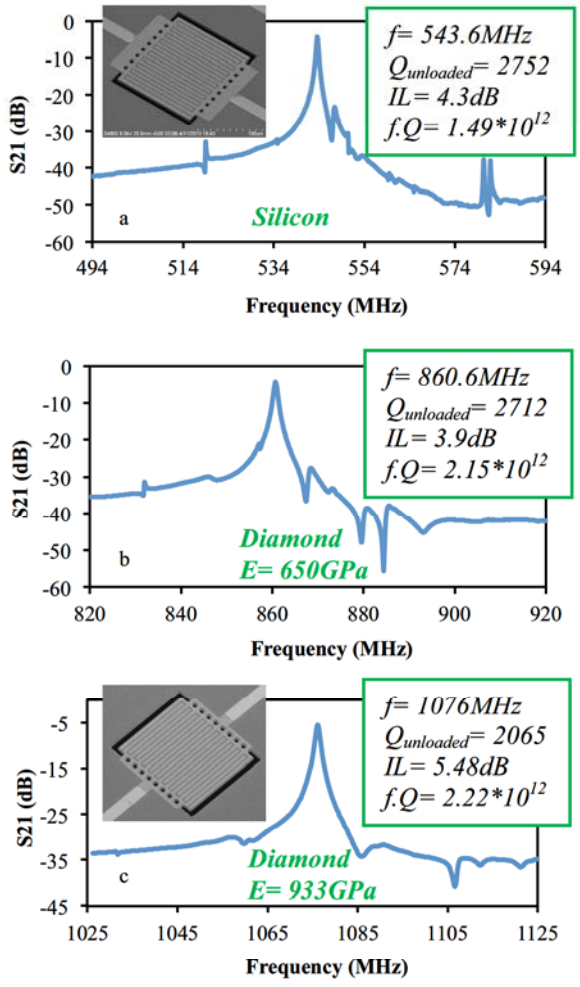


Figure 8. Measured frequency response for a 21st harmonic resonator fabricated on (a) Silicon, (b) Diamond with E=650GPa, (c) Diamond with E=933GPa.

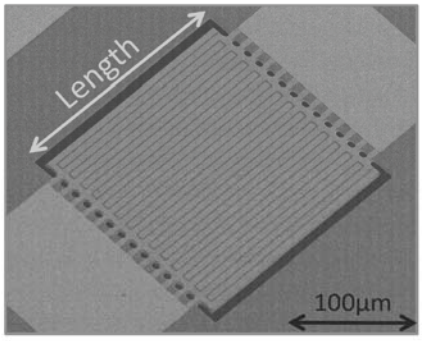


Figure 9. The SEM of the TPoD resonator optimized for excitation of the 29th harmonic.

~205µm long resonator with 7.2µm finger-pitch was designed to excite the 29th harmonic in the lateral dimension. The resonator structure is supported by 13 pairs of support tethers. The SEM of the fabricated device is shown in Fig. 9. This resonator exhibits an unloaded quality factor of ~3000 at 888.6MHz with a **2.6dB** insertion loss (IL) when terminated by standard **50Ω** impedance (Fig. 10). This record low IL corresponds to a 22Ω motional impedance as shown in the equivalent electrical model of the resonator in Fig. 10. The effective electromechanical coupling coefficient ($k^2_{t,eff}$) of this resonator is approximately %0.43 extracted from the equivalent electrical circuit [8]. The lowest IL measured from this device fabricated on three diamond wafers with three different Young's modulus are compared in Table I.

The temperature coefficient of frequency (TCF) for a typical 29th harmonic design described above, was measured in a vacuum chamber to be -9.6 ppm/°C (Fig. 11) which is much lower compared to devices fabricated on silicon for

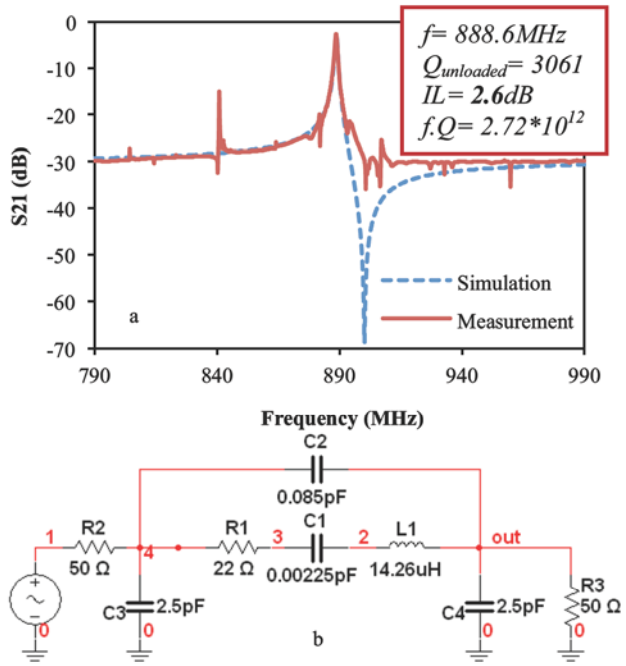


Figure 10. (a) Measured and simulated frequency responses of the 29th harmonic resonator, (b) Equivalent electrical model of the resonator.

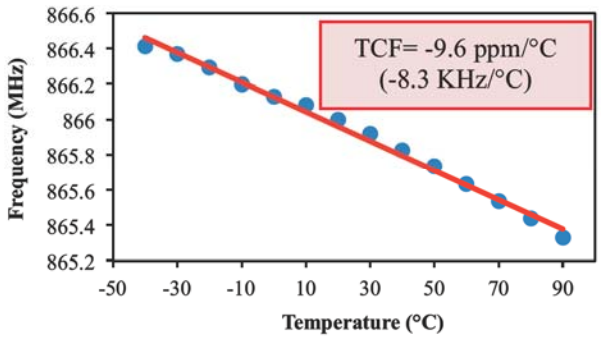


Figure 11. A typical frequency vs. temperature plot measured from a 29th harmonic TPoD resonator.

which the TCF is in the range of -20 to -30 ppm/°C [4, 9]. And also the nonlinearity of the device was examined by varying the input power from -20dBm to 15dBm. The resulted frequency responses are plotted in Fig. 12. The nonlinear behavior is only noticeable from 12dBm input power.

V. CONCLUSIONS

Ultrananocrystalline diamond was deposited using hot filament chemical vapor deposition (HFCVD) technique and further polished to reduce the surface roughness to less than 1nm, therefore promoting the growth of highly oriented c-plane piezoelectric AlN. High frequency and low motional impedance lateral-extensional resonators are fabricated integrating ultrananocrystalline diamond (UNCD) thin films with AlN thin-films. A multi-tethered 29th harmonic TPoD resonator is reported with a motional impedance of 22Ω (insertion loss of 2.6dB) at 888.6MHz. This resonator exhibits

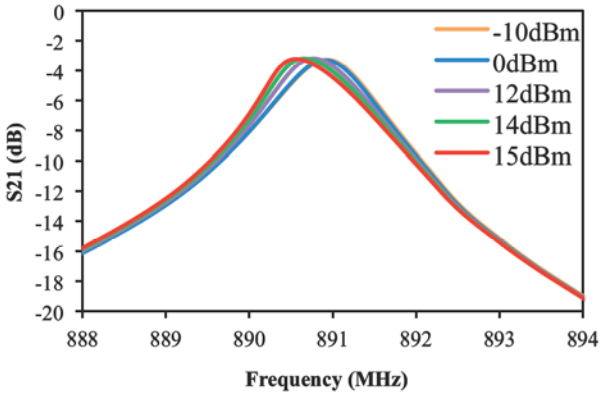


Figure 12. Overlapped frequency responses of the 29th harmonic TPoD resonator measured at several input power values.

TABLE I. RESULTS FROM DIFFERENT DIAMOND WAFERS WITH DIFFERENT YOUNG'S MODULUS.

Young's Modulus (GPa)	Surface Roughness (nm)	AlN FWHM	Frequency (MHz)	Quality Factor	Minimum Insertion Loss (dB)
491	0.4-0.5	4.1°	797	2050	3.2
650	0.4-0.5	3.1°	888.6	3061	2.6
933	1.0-1.8	3.2°	1074	2631	3.9

an unloaded quality factor of ~3000 and has a low TCF value of about -9.6ppm/°C.

ACKNOWLEDGMENT

Authors wish to thank Jonathan Gonzales for his contribution in fabrication of the wafers.

REFERENCES

[1] H. Fatemi, B. P. Harrington, H. Zeng, J. Carlisle, and R. Abdolvand, "50Ω-Terminated 900MHz Monolithic Lateral-Extensional Piezoelectric Filters on Ultrananocrystalline Diamond, "Micro Electro Mechanical Systems, pp.744-747, Jan. 2011.

[2] S. Humad, R. Abdolvand, G. K. Ho, G. Piazza, and F. Ayazi, "High Frequency Micromechanical Piezo-on-Silicon Block Resonators," IEEE Electron Devices Meeting, Technical Digest., pp. 39.3.1- 39.3.4, Dec. 2003.

[3] B. P. Harrington, and R. Abdolvand, "Q-enhancement Through Minimization of Acoustic Energy Radiation in Micromachined Lateral-Mode Resonators," Solid-State Sensors, Actuators and Microsystems Conference, TRANSDUCERS, pp.700-703, June 2009.

[4] G. Piazza, P. J. Stephanou, and A. P. Pisano, "Piezoelectric Aluminum Nitride Vibrating Contour-Mode MEMS Resonators," Journal of Microelectromechanical Systems, vol.15, no.6, pp.1406-1418, Dec. 2006.

[5] O. Auciello, S. Pacheco, A. V. Sumant, C. Gudeman, S. Sampath, A. Datta, R. W. Carpick, V. P. Adiga, P. Zurcher, M. Zhenqiang, H. Yuan; J. Carlisle, B. Kabius, J. Hiller, and S. Srinivasan, "Are Diamonds a MEMS' Best Friend?," Microwave Magazine, IEEE , vol.8, no.6, pp.61-75, Dec. 2007.

[6] R. Abdolvand, G. K. Ho; J. Butler; and F. Ayazi; "ZnO-on-nanocrystalline diamond lateral bulk acoustic resonators", Micro Electro Mechanical Systems, pp.795-798, Jan. 2007.

[7] B. P. Harrington, M. Shahmohammadi, and R. Abdolvand, "Toward ultimate performance in GHZ MEMS resonators: Low impedance and high Q," Micro Electro Mechanical Systems, pp.707-710, Jan. 2010.

[8] M. Rinaldi, C. Zuniga,; Z. Chengjie, G. Piazza, "AlN contour-mode resonators for narrow-band filters above 3 GHz," Frequency Control Symposium Joint with European Frequency and Time forum, pp.70-74, April 2009.

[9] M. Shahmohammadi, B. P. Harrington, and R. Abdolvand, "Concurrent Enhancement of Q and Power Handling in Multi-Tether High-Order Extensional Resonators," Microwave Symposium Digest, pp.1452-1455, May 2010.

$LiNbO_3$ thin single crystal layer for RF applications

B.Imbert, A. Reinhardt, T. Ricart, C.Billard, M. Pijolat,
E. Defaÿ, E. Augendre, T. Signamarcheix, C. Deguet
CEA - LETI
17 rue des Martyrs, 38054, Grenoble, France
imbert.bruno@cea.fr

Abstract—**For the first time, fabrication of sub-micron thick single crystal lithium niobate ($LiNbO_3$) membranes on which Film Bulk Acoustic Resonator (FBAR) is reported. The transfer of thin single crystal piezoelectric layers on specific base substrate has been physically and electrically characterized showing to be an innovative and efficient substrate solution for FBAR applications: resonances close to 2.35 GHz with an electromechanical coupling factor of 3.8 % were obtained. Characterized membranes were obtained using the Smart Cut™ technology developed for piezoelectric material such as $LiNbO_3$.**

I. INTRODUCTION

RF filters are a key component in modern front-end circuits, especially when multiband/multimode operations are needed. The use of BAW (Bulk Acoustic Wave) or SAW (Surface Acoustic Wave) technologies enables synthesis of low-loss and compact RF filters. The bandwidth of these filters is primarily defined by the electromechanical coupling factor of each resonators used to build the filter (ladder or lattice topologies). The effective required bandwidth of filter matches at least what is defined by the targeted standard (usually 60 MHz close to 2 GHz). It is however usually increased to account for post-process frequency dispersions and for frequency shifts induced by temperature. Depending on how tightly frequency can be controlled during process and on the thermal sensitivity, these margins can range from a few MHz only up to 15 MHz, increasing the effective bandwidth requirement. As far as BAW filters are concerned, usual designs tend to sacrifice effective electromechanical coupling factor to reach the quality factors required to ensure low losses, pushing even further requirements in terms of piezoelectric material properties. Aluminium Nitride (AlN) is usually sufficient to meet the requirements for most standards (for example PCS and WCDMA) [1], but is hardly sufficient for the most stringent applications such as DCS for example [2]. Thus, the use of materials with larger piezoelectric properties is a possible approach to overcome these limitations.

In order to improve piezoelectric properties, the transfer of thin single crystal piezoelectric layers onto specific base substrates has shown to be an innovative solution [3]. Incorporated into the engineered substrate, buried metallic electrodes enabled the fabrication of HBAR (High Overtone Bulk Acoustic Resonators). Using this approach, we fabricated recently HBARs based on a $LiNbO_3$ thin film transferred onto a $LiNbO_3$ substrate using the Smart Cut™ technology, which allows the transfer of a sub-micron thick monocrystalline layer [4]. To go one step further, we also fabricated Film Bulk Acoustic Resonators (FBAR) using the direct transfer of a $LiNbO_3$ onto a silicon substrate, the grinding and polishing of the piezoelectric substrate to decrease the thickness of the piezoelectric layer, and the forming of free standing membranes through bulk silicon micromachining [5]. Such an approach can only provide films with thicknesses of a few tens of micrometers. Therefore, fabricated resonators exhibited frequencies in the range of 250 MHz. Moreover, the parallelism of the two surfaces of the grinded film proved to be a critical issue with regard to quality factors.

This motivated us to investigate the use of the Smart Cut™ technology to transfer a sub-micron thick $LiNbO_3$ layer to form an FBAR (Film Bulk Acoustic Resonator). In section II we describe the technique used to obtain sub-micron thick single crystal lithium niobate films used in FBAR resonators. We then report on the electrical characterization of the devices in section III.

II. SUBSTRATE REALISATION AND CHARACTERIZATIONS

The Smart Cut™ process is a well known method successfully used by SOITEC for the mass production of SOI (Silicon on Insulator) wafers [6]. This technology has the ability to the transfer of piezoelectric materials (Figure 1), like $LiNbO_3$ [4]. Several articles [5,7] have demonstrated examples of thin $LiNbO_3$ single crystal layer transferred onto silicon substrates. This approach suffers from the large thermal expansion mismatch between silicon and $LiNbO_3$. It causes large strains, which could be detrimental for the self-standing membrane. For this reason, we chose the option of transferring a thin $LiNbO_3$ layer onto a $LiNbO_3$ substrate.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

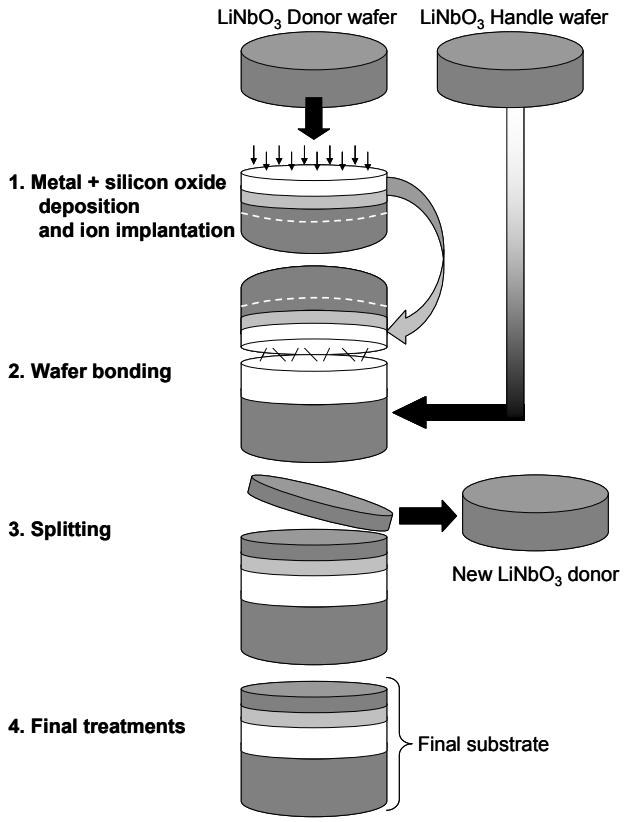


Figure 1. Smart Cut™ technology applied to $LiNbO_3$ thin film fabrication.

A cleaning process has been developed, to obtain a particle free surface without any degradation of the initial roughness, leaving substrates compatible with the bonding technology employed. As required by the final structure, a metallic (low resistivity metal) electrode and PECVD silicon oxide are deposited on the X-cut $LiNbO_3$ bulk donor/handle substrates. The donor structure is then subjected to ion implantation (usually He ions) [8,9]. Top surfaces are then prepared to be bonded at room temperature to a base X-cut, single side polished $LiNbO_3$ substrate thanks to a hydrophilic bonding. Then, a thermal anneal is performed to initiate the splitting of the donor wafer in order to transfer a 1 µm-thick film of $LiNbO_3$. Finishing treatments (annealing and polishing) make $LiNbO_3$ surfaces compatible for device processing. CMP treatment removes the damaged surface layer resulting from implantation and from the splitting. It is also used as a way to control the $LiNbO_3$ layer thickness. After CMP, surface roughness is as smooth as the initial bulk substrate surface and compatible with device fabrication requirements (below 0.6 nm RMS) [10,11]. The donor wafer, from which a thin film has been transferred, can be re-used.

Physical characterizations realized on different samples were used to determine the crystalline quality of the transferred $LiNbO_3$ layer. X-Ray Diffraction (XRD) measurements performed on the as-split transferred layer confirm its single-crystal nature and quality. The $LiNbO_3$ (110) X-ray spectra shown in Figure 2 correspond to measurements of the handle substrate and of the transferred layer performed right after the splitting step. On one hand, the widening of the peak corresponding to the transferred layer can be attributed to implantation sub-surface damages localised in the first 50 nm. On the other hand, the shift observed is the expression of a residual layer strain (about 0.18 %) after transfer. During the final process steps, the remaining damage are completely removed, leaving a bulk-like single crystal material. Furthermore, no extended defects have been observed in the transferred layer by High Resolution TEM (Figure 3).

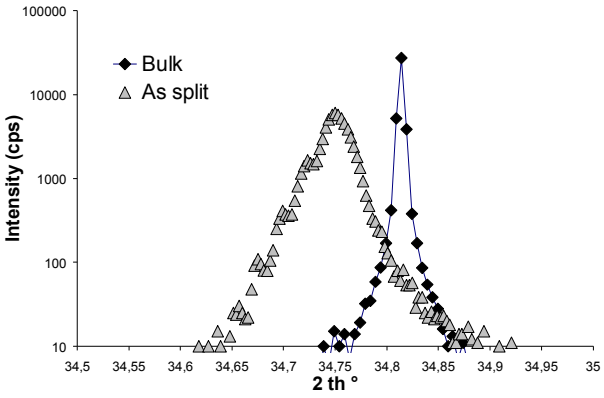


Figure 2. $LiNbO_3$ (110) XRD signals measured right after the Smart Cut™ splitting step for the $LiNbO_3$ layer and for the handle substrate.

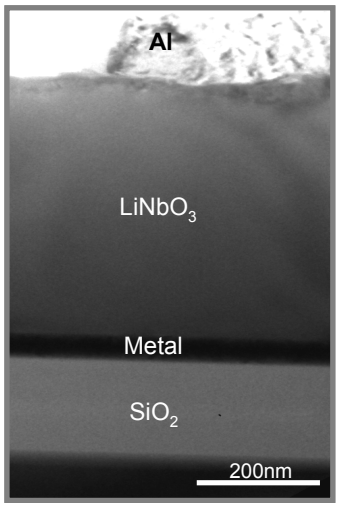


Figure 3. HRTEM cross section of $LiNbO_3$ thin monocrystalline layer with metallic electrode for FBAR devices.

After transfer, the $LiNbO_3$ film is isolated from the bulk substrate (Figure 3) by forming an air-gap between the transferred layer and the handle substrate so as to realize a piezoelectric membrane. Finally, a top aluminium electrode is deposited by evaporation and patterned using wet etch chemistry.

III. SINGLE CRYSTAL FBAR ELECTRICAL CHARACTERIZATION

So as to pattern the only top electrode, we use a set of two resonators connected in series at the level of the bottom electrode. These resonators have been probed on wafer using a vector network analyser. To avoid destruction of the

membrane, probes were positioned on contact pads located outside of the membrane. For this reason, however, we measure both the electrical response of the FBAR located on the free standing membrane and the response of the resonator formed by the contact pads located above the continuous bottom electrode and lying directly in contact with the substrate, as shown in Figure 4.

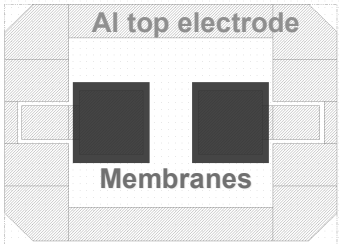


Figure 4. Layout of the resonator indicating the position of the cavity and of contact pads.

As expected from this layout configuration, electrical response of resonators, shown in Figure 5, reveals features associated with both the part of the resonator located over the cavity and the part located at the level of the ground ring and of contact pads. Indeed, two relatively sharp resonances are located at 2.35 GHz for the most coupled mode (fundamental) and at 2.48 GHz for a less coupled one. Their effective electromechanical coupling factors respectively reach 3.8 and 0.7 %, while their quality factors exceed 60. They could correspond to the resonances of the two shear waves piezoelectrically coupled in X-cut $LiNbO_3$. A resonance with a lower quality factor is also visible close to 3 GHz and is likely to correspond to the response of the electrical pads and of the ground ring, where the vibrations in the transferred film radiate in the substrate.

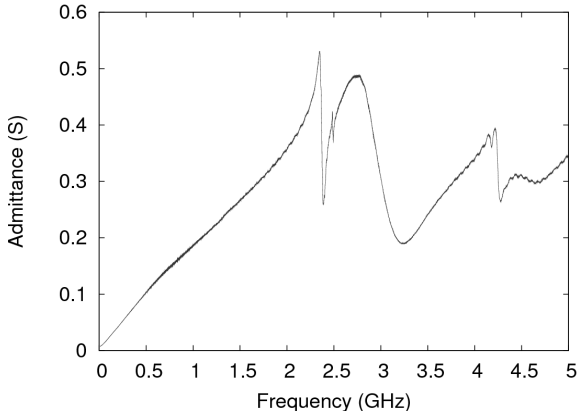


Figure 5. Electrical response of a $LiNbO_3$ FBAR resonator.

To confirm these assumptions, we simulated both parts of the material stack using the scattering matrix method [12] and an estimation of the thicknesses of the various layers obtained from TEM cross sections as shown in Figure 3. The membrane part consists in a SiO_2 / metal / X-cut $LiNbO_3$ 500 nm / Al stack. Its lossless theoretical response (scaled to match the impedance of the complete device) is superimposed with measurements in Figure 6. An excellent agreement is obtained for the position of the fast shear wave resonance, both for the fundamental and for the second order resonance above 4 GHz. The effective electromechanical coupling factor is also well predicted for the second harmonic, suggesting that the discrepancy between the relatively small effective coupling factor of the fundamental (3.8 %) compared to theoretical values (above 45 %) is not related to insufficient piezoelectric properties of the transferred film, but to perturbations brought by the proximity with the low quality factor resonance located at 2.7 GHz.

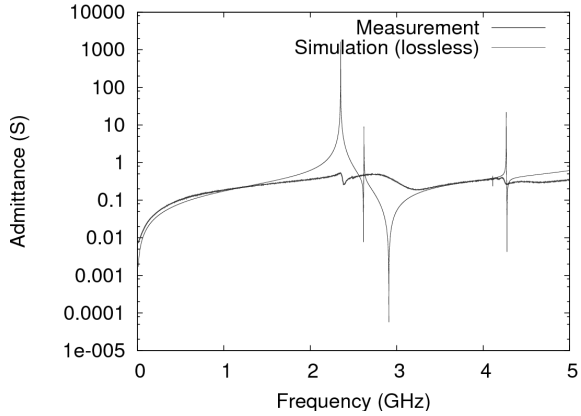


Figure 6. Simulation (light grey curve) of the FBAR part of the device.

We also simulated the same material stack radiating in a semi-infinite X-cut $LiNbO_3$ substrate, representing the behaviour of the single-side polished substrate. As shown in Figure 7, the position of the resonance and antiresonance caused by the piezoelectric film radiating in the semi-infinite substrate corresponds to theoretical predictions. It is to be noted that due to the wide spectrum of these resonances, those caused by the fast shear and by the slow shear wave merge together.

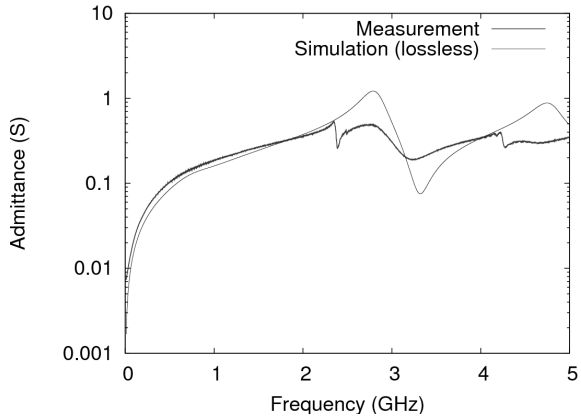


Figure 7. Simulation (light grey curve) of the part of the device located directly on the substrate.

Finally, we merge the two parts of the device by adding their admittances, ponderated by their relative areas. We also add an additional 2 Ω series resistance corresponding to the resistivity of the electrodes and to the length of the electrical path. Response shown in Figure 8, shows that the parallel parasite capacitance brought by the contribution of the part of

the resonator not located above the air-gap dramatically decreases the effective electromechanical coupling factor by a factor close to 10. Therefore, we believe that to the first order, transferred film has kept most of its piezoelectric properties. Quality factors at antiresonance seem higher in simulations than in practice, but it is to be noted that the only sources of losses included in the model are radiation within the substrate and the small series resistance. So, material losses in the stack may account for this phenomenon. Especially, PECVD silicon dioxide used as a bonding layer could explain this result.

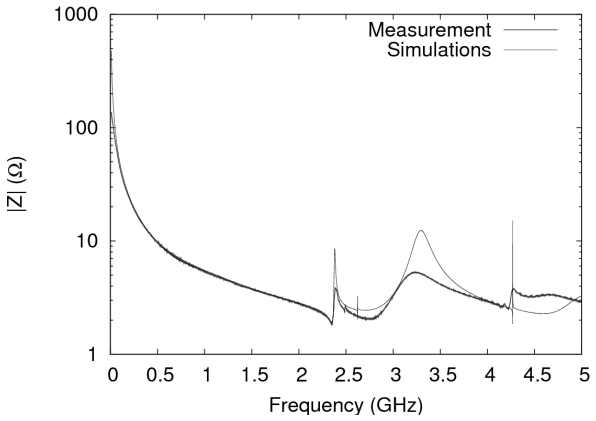


Figure 8. Comparison between measurement (dark grey) and simulation (ligh grey) of the full device.

IV. CONCLUSION

We have reported the first fabrication of sub-micron thick single crystal $LiNbO_3$ membranes in which FBAR resonators were fabricated. Unlike previous works, these membranes were obtained using the Smart Cut™ process and are submicron thick. Furthermore, resonances close to 2.35 GHz were obtained. Although the electromechanical coupling factor is only of 3.8 %, simulations reveal that this relatively low electromechanical coupling factor is largely due to the parallel parasitic capacitance brought by the large part of the resonator located outside of the membrane which is also responsible for the additional resonance damped by radiation within the substrate. Beyond this first study, future work will focus on improving the design of these devices, principally by patterning the bottom electrode to have only the resonators located on the membrane, and substrate optimization to reduce material losses.

ACKNOWLEDGMENT

This work has been performed with the help of the "Plateforme Technologique Amont de Grenoble" with the financial support of the French ANR under "POEM" project.

REFERENCES

[1] R. Aigner, "SAW and BAW technologies for RF filter applications: a review of the relative strengths and weaknesses", Proceedings of the 2008 International Ultrasonics Symposium, p. 582, 2008.

[2] M. Clement, E. Iborra, J. Olivares, N. Rimmer, S. Giraud, S. Bila and A. Reinhardt, "DCS Tx filters using AlN resonators with Iridium electrodes", IEEE Trans. Ultrason., Ferroelectr. and Freq. Contr., 57 (3), pp. 518-523, 2010.

[3] C. Deguet, et al., "$LiNbO_3$ single crystal layer transfer techniques for high performances RF filters", ECS Meeting vol. 33, p. 225, 2010.

[4] M. Pijolat, A. Reinhardt, E. Defaÿ, C. Deguet, D. Mercier, M. Aïd, J.S. Moulet, B. Ghyselen, "Large Q.f product for HBAR using Smart Cut transfer of $LiNbO_3$ thin layers onto $LiNbO_3$ substrate", International Ultrasonics Symposium, p. 201, 2008.

[5] M. Pijolat, S. Loubriat, S. Queste, D. Mercier, A. Reinhardt, E. Defaÿ, C. Deguet, L. Clavelier, H. Moriceau, M. Aïd and S. Ballandras, "Large electromechanical coupling factor film bulk acoustic resonator with X-cut $LiNbO_3$ layer transfer", Appl. Phys. Lett. 95, 182106, 2009.

[6] M. Bruel, "Silicon on insulator material technology", Electron. Lett.,vol. 31, p. 1201, 1995.

[7] T. Pastureaud, M. Solal, B. Biasse, B. Aspar, J.-B. Briot, W. Daniau, W. Steichen, R. Lardat, V. Laude, A. Laëns, J.-M. Friedt, and S. Ballandras, "High-Frequency Surface Acoustic Waves Excited on Thin-Oriented $LiNbO_3$ Single-Crystal Layers Transferred Onto Silicon", IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol. 54, 2007.

[8] Y.B. Park et al, "Integration of Single-Crystal $LiNbO_3$ Thin Film on Silicon by Laser Irradiation and Ion Implantation–Induced Layer Transfer", Advanced Materials, vol 18, p. 1533, 2006.

[9] I. Radu et al, Mat. Res. Soc. Symp. Proc., vol 748, 2003.

[10] K. A. Reinhardt and R. F. Reidy, Handbook of cleaning for semiconductor manufacturing, Fundamentals and Applications, pp.509-510, 2011

[11] H. Moriceau, O. Rayssac, B. Aspar, B. Ghyselen, Semiconductor Wafer Bonding : Science, Technology and Applications VII, (H. Baumgart, S. Bengtsson, C.E. Hunt, and T. Suga, eds.) 2003-19:49, The Electrochemical Society, Pennington, New Jersey, 2003.

[12] A. Reinhardt, T. Pastureaud, S. Ballandras and V. Laude, "Scattering matrix method for modelling acoustic waves in piezoelectric, fluid, and metallic multilayers", Journ. Appl. Phys. Vol 94, n°10, pp. 6923-6931, 2003.

The time stability of PPP links for TAI

Gérard Petit, Aurélie Harmegnies
Bureau Internatonal des Poids et Mesures
92312 Sèvres France
gpetit@bipm.org

Flavien Mercier, Félix Perosanz
Centre National d'Etudes Spatiales
31401 Toulouse France

Sylvain Loyer
CLS
31520 Ramonville Saint-Agne France

***Abstract*— In the last few years the BIPM has started using GNSS phase and code observations and the Precise Point Positioning technique to compute time links for the generation of TAI. The estimated instability of such links for averaging time up to 1 month has been taken as 0.3 ns. In this paper, we investigate methods to estimate this instability, following several approaches.**

I. INTRODUCTION

In the last few years the BIPM has started using GNSS phase and code observations to compute time links. The Precise Point Positioning (PPP) technique has been chosen for its relative simplicity of operation while delivering state of the art results, thanks to the precise satellite orbits and clock products generated by the International GNSS Service. Following earlier studies [1], the uncertainty u_A for PPP links used in TAI is taken as 0.3 ns. As specified in Circular T, this value accounts for measurement noise (which is negligible for PPP due to the low noise of the GNSS phase measurements) but should also account for all effects with typical duration up to one month. In this paper we use different techniques to estimate the time instability of PPP links in the aim of validating the chosen u_A value.

In Section II, some background is given on the use of PPP links in TAI. In Section III, the instability of these operational links is studied through the discontinuities between monthly solutions. Section IV presents results of tests using different processing software and options, including link comparisons to Two Way time transfer. Some implications on the present and future use of this technique in the TAI computation are discussed in Section V.

II. BACKGROUND ON PPP LINKS IN TAI

Among GPS processing techniques, Precise Point Positioning (PPP) appeared as a natural choice for TAI computation needs because it is particularly adapted to a global, but sparse, network of stations and because such processing is flexible and easy to implement [2]. Therefore, to answer the requests from the Consultative Committee for Time and Frequency, the BIPM initiated in 2008 a pilot experiment, named TAIPPP, where time laboratories contribute GPS phase and code data and where the BIPM uses the PPP technique to generate monthly solutions, at first in slightly deferred time after the regular TAI computation. After an initial buid-up, some 30 laboratories have regularly contributed data. Since Circular T 261 (September 2009) PPP has been used for TAI time links and some twenty such links are now regularly used.

All results of the PPP computation are posted on the BIPM web site at ftp://tai.bipm.org/TimeLink/TAIPPP/. In addition, all available link comparisons between PPP and other techniques are included in the link comparisons web page at ftp://tai.bipm.org/TimeLink/LkC.

In March 2011 (the last month in this study), 33 stations sent Rinex data to the BIPM for PPP computation, see the map in Fig. 1. Seven of these stations also participate to the network of the International GNSS Service (IGS) [3] and clock solutions for these stations are also generally available from the IGS.

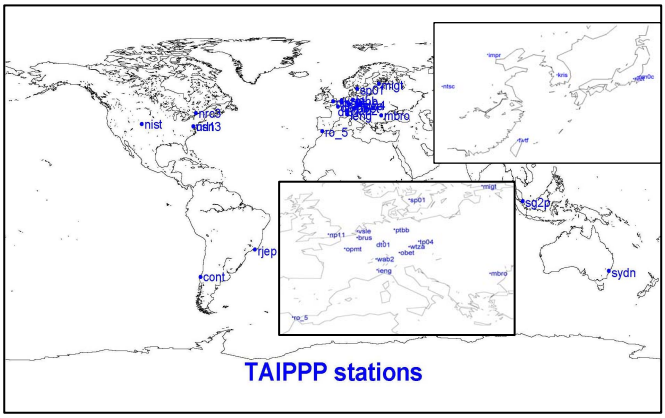


Figure 1. The network of the TAI stations for PPP analysis

The software used in the BIPM operational analysis for TAI is the GPSPPP software developed by Natural Resources Canada [4], presently version 1.04/1087, released in May

2007 and upgraded in July 2008. The software directly uses IGS files for satellite ephemerides and clocks, as well as for the absolute antenna phase center offsets [5]. The BIPM uses the IGS Rapid products, which are available with less than 3 days of delay so are available for the whole month at the date of TAI computation, therefore these same products will be used in this study.

The main operational parameters of the GPSPPP software for the computation of 35-day (or 40-day) batches for TAI are the following: use IGS Rapid SP3 orbits and 5-min satellite clocks referred to the rapid version IGRT of the IGS time scale [6]; ionosphere-free linear combination of code and phase measurements are taken with *a priori* weights of 1 m for code and 1 cm for phase; the elevation cut-off is set to 10°; the observation sampling and clock solution interval are both 5 minutes; the tropospheric zenith delay is modeled as 3mm/√hr random walk, with the GMF [7] mapping function used; ocean loading coefficients are from [8]; station coordinates are estimated on each 1-month batch.

III. Analysis of Discontinuities between Monthly PPP Results

A well-known feature of GPS phase and code solutions based on successive (e.g. daily) batches is the presence of so-called "boundary discontinuities" [9]. The discontinuities are due to the fact that phases are aligned on codes in different manners over the two intervals. They may thus originate either from noise of the code measurements, or from build-up of errors e.g. in the phase ambiguity determination. The effect of code noise would decrease, and thus the discontinuities would also decrease, over long batches only in case of a white noise processes. However, it has been shown that other noise processes are present and sometimes dominant [10], so that discontinuities do not necessarily decrease, and may eventually increase, as the batch duration increases.

We here examine the 1-month discontinuities observed over 35 months of PPP analysis for TAI, i.e. from an ensemble of (at most) 34 discontinuities for 27 stations. Because each monthly batch has a 5-day common period with the preceding, we estimate an average value of the discontinuity over the 5-day period and we also estimate an average rate of the difference between the two results over this period. The determination of discontinuities therefore does not depend on the stability of either the local reference or the IGS reference time scale. Table I shows the statistical information computed from the discontinuity results. We see that the typical discontinuity (median value) is about 170 ps in phase and about 4×10^{-16} in rate. Note that a few of the values may be affected by cases when one of the monthly processing had an undetected error, however the large database and the use of the median should mitigate such effects. Not shown in the Table, the average value of the discontinuity is nearly always statistically consistent with zero, with one notable exception: The equipment designated by nist, operated by the NIST in Boulder (USA) has a very specific behaviour and the average value of the discontinuity (-0.32 ns) is very significantly biased. This equipment is the only one which produces C1/P2 (not P1/P2) data and is also the only one of its brand. However no physical process has yet been proposed to explain the significant bias in discontinuity.

TABLE I. Statistical Data on the Monthly Discontinuities in the TAI PPP Processing

Station	# disc	Average value of		
		# points	***Phase disc. / ns***	***Rate disc. /ns/d***
brus	32	1408	0.221	0.044
cont	30	1424	0.22	0.137
dt01	14	1110	0.188	0.084
ieng	33	1318	0.132	0.035
impr	30	1398	0.231	0.299
kris	29	1390	0.175	0.044
mbro	26	1344	0.271	0.092
migt	16	1313	0.1	0.03
nict	34	1431	0.159	0.025
nist	30	1400	0.385	0.035
nm0c	34	1403	0.15	0.03
np11	18	1356	0.079	0.027
nrc3	26	1429	0.288	0.053
nrl1	32	1334	0.138	0.038
obet	32	1344	0.225	0.038
opmt	34	1360	0.229	0.079
ptbb	34	1407	0.128	0.028
ro_5	19	1423	0.101	0.028
sg2p	21	1405	0.173	0.047
sp01	31	1440	0.18	0.034
sydn	19	1366	0.437	0.141
tp04	32	1411	0.084	0.023
twtf	30	1422	0.168	0.151
usn3	32	1440	0.139	0.029
vsle	20	1420	0.165	0.068
wab2	32	1408	0.148	0.094
wtza	31	1400	0.125	0.032

IV. Tests with Different Processing Software and Options

In this section, we first describe tests carried out using two releases of the GPSPPP software. In subsection A, we study the influence on the GPSPPP solutions of the a priori uncertainty assigned to the phase and code measurements, and we compare the two releases in this respect. In subsection B, we form links with the PPP solutions considered in subsection A, and compare these links to Two-Way time transfer (TW).

Finally in subsection C we provide a first comparison of GPSPPP to using the GINS software and the Integer PPP (IPPP) approach which is based on an integer ambiguity resolution scheme.

A. *Tests with the GPSPPP software*

Operational TAI PPP computations are carried out with the NRCan GPSPPP release 1087, hereafter referred to as GPSPPP(2007). Recently NRCan has produced the release 05211 hereafter referred to as GPSPPP(2011). Major changes in release 05211 are (from NRCan release information):

- combined and stand-alone GPS and/or GLONASS processing;
- improved models for GPS yaw/attitude and for station antenna orientation tracking;
- improved processing techniques, such as an adaptive cycle-slip detection and repair, 1ms jumps handling.

One important processing option is the a priori uncertainty for the two types of measurements: Usual default values are u_p = 1 cm for phase and u_c = 1 m for code, representative of the actual measurement uncertainties (such a solution is referred to as "standard"). Here we test the impact of choosing significantly different values for u_c, from 0.5 to 3 m, where low values of u_c tend towards a "code-only" solution while high values tend towards a "code-free" solution. Such expressions are hereafter used to design the two types of solutions.

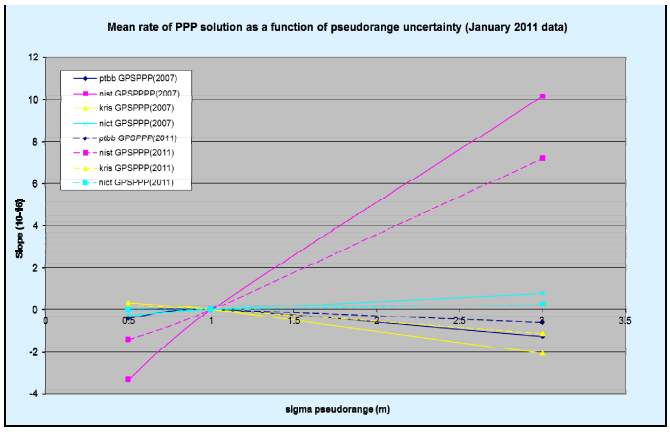


Figure 2. Rate variation of different GPSPPP solutions with respect to the code a priori uncertainty.

We observe that solutions with different u_c values differ essentially by a rate of order a few parts in 10^{16}. This causes somewhat significant differences between the PPP solutions, e.g. the RMS of the difference between two solutions over one month is of order 150 ps for a $2x10^{-16}$ rate difference. As seen in Figure 2, the effect is about similar for all stations, except for nist for which it is much larger. This effect for nist is in relation with the significant bias in the monthly discontinuities, where the net effect of a rate which repeats month after month is to produce a saw-tooth behavior with a discontinuity always of the same sign. Also we note that the effect is significantly reduced using GPSPPP(2011) in all cases, although it is not clear what is the main reason for this change in behavior between the two releases.

B. *Comparisons of GPSPPP links with TW*

Each monthly computation provides the station clock results in the form [Reference – IGRT] with a 5-min interval. Station clock results provide, by simple difference between two stations, link results which can be compared to the TW link results when these are available. Because the quality of several TW links has been somewhat degraded since the change of satellite and set-up in August 2009, all tests are carried out with several stations over July 2009 (noted 0907), and with two specific links (NIST-PTB and KRIS-NICT) over January 2011 (1101). Note that stations are identified either by the name of the TAI laboratory (when upper case in used) or by the name given to the GNSS receiver (lower case).

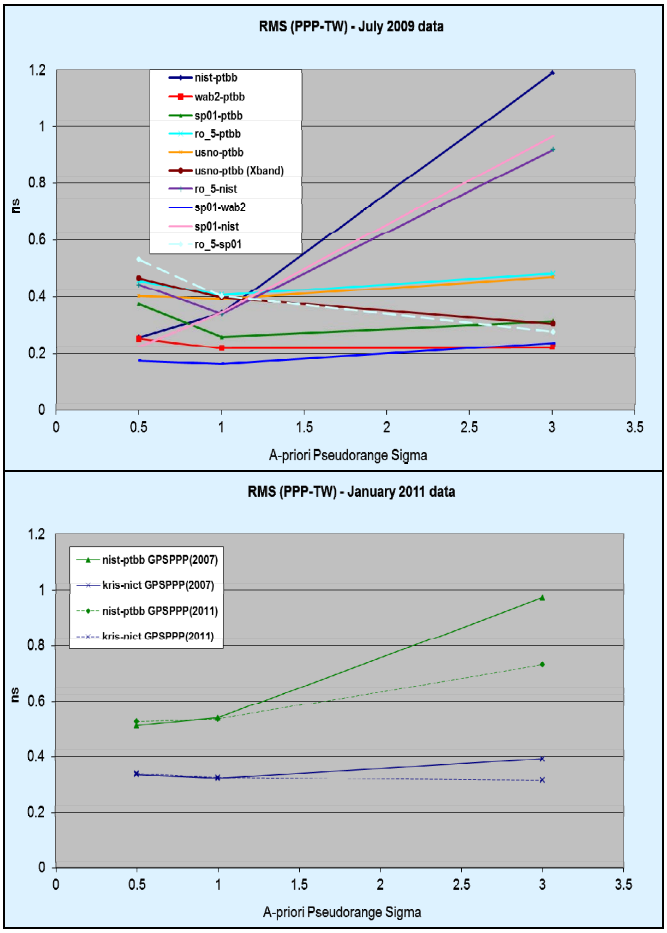


Figure 3. RMS of the difference to TW with respect to the code a priori uncertainty. Top plot: for several links with GPSPPP(2007) for July 2009; Bottom plot: for two links and the two GPSPPP releases for January 2011.

As the magnitude of the effect observed in the previous section is typically of order 150 ps for a given station, the effect on a given link will typically be of order 200 ps or larger, so that it might be possible to discriminate between the different PPP solutions by comparing them to the independently measured TW links. In this aim, for all pairs of stations for which TW is available, PPP links are formed for all types of PPP solution cited in the preceding section and the

RMS of the difference with TW is computed. We observe (Figure 3) that for nist "code-free" solutions (solution with a large code a priori uncertainty) are obviously worse than "code-only" solutions so that the effect must be affecting the phase. We note that, for GPSPPP(2007), there seems to be an "optimal value" of u_c, in the sense that for the best links a minimum can be found in the RMS of the difference to TW. We also observe (case of kris-nict) that GPSPPP(2007) "code-free" solutions seem sub-optimal compared to the "standard" solution which has a better agreement with TW. Using GPSPPP(2011), this feature seems to disappear (all solutions seem equivalent compared to TW), which may be linked to the improved physical model in the GPSPPP(2011) release.

C. *Tests with the GINS-IPPP software*

GINS is a multi-technique geodetic software which is being developed by CNES-GRGS for more than 40 years [11]. In addition to the classical double-difference and PPP GPS data processing strategy, the Integer-PPP (IPPP) technique has recently been implemented into GINS. IPPP is based on the resolution of the phase data integer ambiguities at the un-differenced level, a feature that is expected to greatly improve the solution. This new and alternative approach needs an accurate model of GPS measurements, precise GPS satellite orbit and clock products and the a priori knowledge of individual satellite fractional-cycle biases. The underlying theory is described by several authors [12, 13, 14, 15]. IPPP results presented hereafter are derived from GPS satellite products that are routinely processed by the CNES-CLS IGS Analysis Center (freely available under the acronym GRG). As station clock solutions rely on integer phase observations only, two successive batch series will differ by an integer number of phase cycles. Continuous (phase) receivers clock solutions can easily be generated.

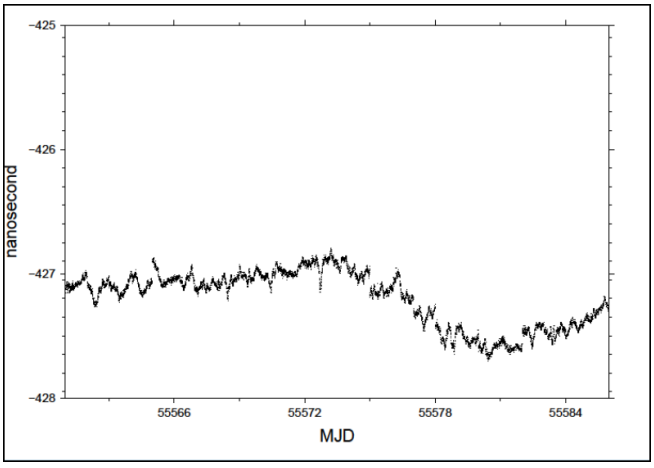


Figure 4. Difference between the IPPP solution and the GPSPPP(2011) "standard" solution for the link nist-ptbb over January 2011 (RMS is 221 ps).

We compute IPPP clock solution difference between NICT and PTBB stations at a 30 second sampling using daily batches and then resolve the integer ambiguities between consecutive days by simple matching. We note that the IPPP and GPSPPP(2011) "standard" solutions differ somewhat (RMS = 221 ps) but do not show significant systematic drift like that observed for the GPSPPP "code free" solution. We infer that IPPP, like TW in the preceding subsection, seems to determine that the GPSPPP "code-free" solution is not optimal. However, with the available data, it is not possible to use TW to determine the preferred solution among them as it seems that the comparison to TW is dominated by noise inherent to TW itself, see Table II.

TABLE II. RMS OF DIFFERENT COMPARISONS FOR THE LINK NIST-PTB OVER JANUARY 2011

Comparison	RMS / ns
GPSPPP (2011) – (2007)	**0.129**
GPSPPP (2011) – IPPP	**0.221**
GPSPPP (2007) – TW	**0.539**
GPSPPP (2011) – TW	**0.534**
IPPP - TW	**0.518**

V. DISCUSSION AND CONCLUSIONS

Time links using PPP solutions have been implemented for TAI computation and now number more than 20, some of them in combination with TW links [16]. The standard uncertainty assigned to these links is 0.3 ns and covers all instability sources with typical duration up to one month. Several studies presented in this paper have shown the following:

- Discontinuities between consecutive monthly PPP solutions have a median value of order 170 ps. Assuming that the results for the two stations of a given link are independent, this indicates that the one-month time instability of a given link is a the level of 250 ps;
- One station shows a significant systematic behavior in discontinuities. Investigations are under way to determine the causes of this effect.
- In the best cases, the comparisons to TW are at the level of 0.2 ns RMS over one month. In other cases, there is no indication that a larger RMS difference is due to the PPP processing.
- Comparisons of PPP links computed with independent software also show a difference of order 0.2 ns RMS. More studies are needed to expand the basis of this comparison.

We therefore conclude that in all typical cases, the value u_A representing the time instability of the link over one month can be taken as 0.3 ns. We expect that further studies involving comparison of different processing software will allow improving on this estimation.

In the present situation, this instability limits the accuracy of frequency comparisons, e.g. for comparing Cs fountains to

TAI, to the level of 1-2x10^{-16} depending on the duration of the comparison (of order 15-30 days). As the best fountains claim an accuracy of order 3x10^{-16} and new frequency standards have already been shown to have an accuracy budget well below 1x10^{-16}, it will shortly be necessary to improve the performances of frequency comparisons.

ACKNOWLEDGMENT

We thank Natural Resources Canada for providing the GPSPPP software, and François Lahaye for his help in maintaining the software and solving run-time problems. The support of all TAI laboratories participating to the TAIPPP experiment is gratefully acknowledged.

REFERENCES

[1] G. Petit, "The TAIPPP pilot experiment," Proc. EFTF-IFCS 2009 Joint Conference, pp. 116–119, 2009.

[2] G. Petit and Z. Jiang, "GPS Precise point positioning for TAI computation," IJNO Article ID 562878, doi:10.1155/2008/562878 2008.

[3] J.M. Dow, R.E. Neilan, G. Gendt, "The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade," Adv. Space Res. 36 vol. 36, no. 3, pp. 320-326, 2005. doi:10.1016/j.asr.2005.05.125

[4] J. Kouba, P. Héroux, "Precise Point Positioning using IGS orbits and clock products," GPS Solutions, 4, p.31, 2001.

[5] G. Gendt, Th. Nischan, "First validation of new IGS products generated with absolute antenna models," IGS Workshop, 2006; see http://igscb.jpl.nasa.gov/mail/igsmail/2006/msg00184.html.

[6] K. Senior, P. Koppang, J. Ray, "Developing an IGS time scale," Proc. IEEE Trans. UFFC 50, 585-593, 2003.

[7] Boehm, J., Niell, A. E., Tregoning, P., and Schuh, H., "Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data," Geophys. Res. Lett., 33, L07304, 2006 doi:10.1029/2005GL025546.

[8] http://www.oso.chalmers.se/~loading/.

[9] J. Ray and K. Senior, "Geodetic Techniques for Time and Frequency Comparisons Using GPS Phase and Code Measurements," Metrologia, vol. 42, pp. 215-232, 2005.

[10] P. Defraigne, C. Bruyninx, "On the link between GPS pseudorange noise and day-boundary discontinuities in geodetic time transfer solutions," GPS Solutions, vol. 11, no. 4, pp. 239–249, 2007.

[11] J.C. Marty (eds), Documentation algorithmique du programme GINS, Version 5 juillet 2009, www.igsac-cnes.cls.fr/documents/gins/GINS_Doc_Algo.html

[12] D. Laurichesse, F. Mercier,"Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP," ION GNSS 2007 20th International Technical Meeting of the Satellite Division, 25-28, September 2007, Fort Worth, TX, 2007.

[13] J. Delporte, F.Mercier, D. Laurichesse, O. Galy, "GPS Carrier Phase Time transfer Using Single Difference Integer Ambiguity Resolution," IJNO, Article ID 273785, 2008.

[14] Ge, M., Gendt, G., Rothacher, M., Shi, C., Liu, J. Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. J. Geod. 82(7), 389-399, 2008.

[15] Geng, J., Meng, X., Dodson, A., and Teferle, F.N., "Integer ambiguity resolution in precise point positioning: method comparison," J Geod 84:569–581, 2010.

[16] Z. Jiang and G. Petit, "Combination of TWSTFT and GNSS for accurate time transfer, Metrologia 46, 305-314, 2009.

Precision and Accuracy of USNO GPS Carrier Phase Time Transfer: Further Studies

Christine Hackman and Demetrios Matsakis
United States Naval Observatory
Washington, DC USA
christine.hackman@usno.navy.mil

Abstract—**The United States Naval Observatory (USNO) produces GPS carrier-phase time-transfer (GPSCPTT) estimates for approximately 100 receiver clocks with 16-hr latency daily in its so-called "rapid" processing. We investigate the uncertainty of these estimates by comparing one year's worth of USNO rapid GPSCPTT estimates to those obtained from two-way satellite time transfer (TWSTT) and from the IGS rapid combined solutions. The time stability of USNO rapid GPSCPTT estimates appears to lie between tens and hundreds of ps for averaging times of 300 s and larger, remaining approximately constant as averaging time increases. The frequency stability of USNO GPSCPTT estimates lies between 10^{-15} and 10^{-14} at an averaging time of one day, decreases as averaging time increases, and could potentially achieve 10^{-16} in 10 d. More work is needed to establish credible frequency-uncertainty estimates at the averaging times (5-20 d) characteristic of the overlap between primary frequency standard evaluations.**

I. INTRODUCTION

The GPS Analysis Division, part of the Earth Orientation Department at the United States Naval Observatory (USNO), produces GPS carrier-phase time-transfer (GPSCPTT) estimates for approximately 100 GPS-receiver clocks daily in its service as an associate analysis center of the International GNSS Service (IGS) [1]. Clock estimates are produced in both "rapid" and "ultra-rapid" operations, each of which creates time-transfer estimates spaced at 5-minute intervals. The processing is conducted using *Bernese 5.0 GPS Analysis Software* [2] in tandem with automation routines developed in-house, with each product set containing estimates of GPS satellite orbits and earth-orientation parameters in addition to GPSCPTT values. While the ultra-rapid clock products are useful for near-real-time applications (they are available with 3-hr latency and provide 24 h of predictions), an initial assessment of their precision was performed in [3] and they are not discussed further here.

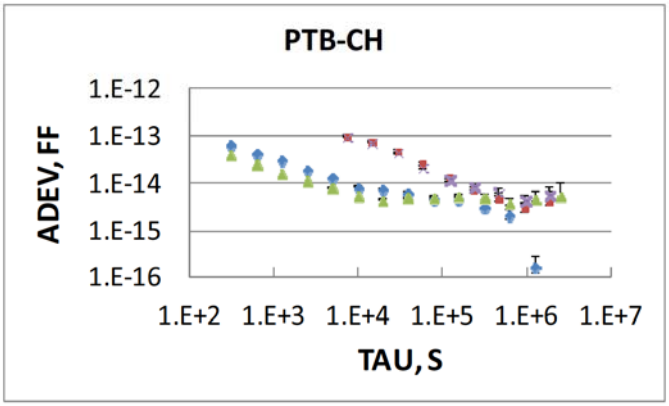
"Rapid" processing is conducted once/day using 24 h of measurements collected the previous UTC day. Solutions for approximately 100 clocks are obtained using either network or precise point positioning (PPP) [4] algorithms and are available with 16-h latency. We evaluate the uncertainty of these estimates by comparing them to those obtained from two-way satellite time transfer (TWSTT) and from the IGS rapid combined solutions [1].

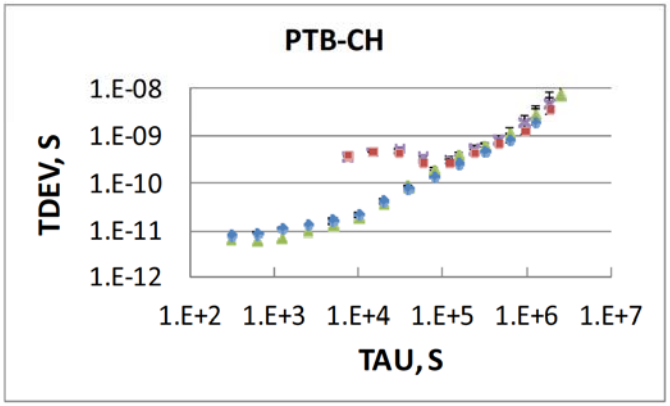
II. METHOD

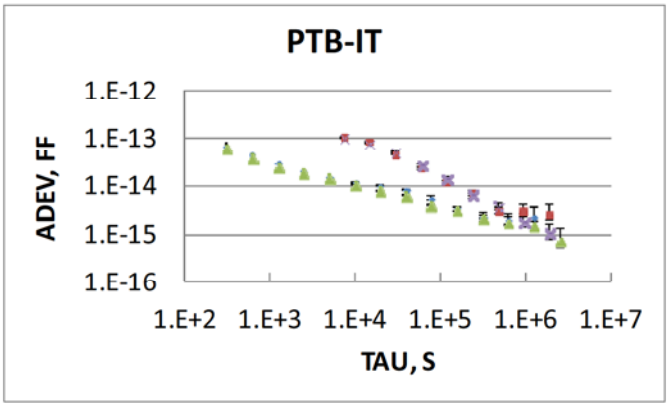
Clock estimates from time-transfer links between three pairs of BIPM timing laboratories were compared. The laboratories are Physikalisch-Technische Bundesanstalt (PTB; Braunschweig, Germany), Istituto Nazionale di Recerca Metrologica (IT; Torino, Italy), (Swiss) Federal Office of Metrology (CH; Wabern, Switzerland) and National Institute of Standards and Technology (NIST; Boulder, Colorado, USA). The links studied were PTB-CH (635 km), PTB-IT (835 km), and PTB-NIST (7532 km). Data from 1 April 2010 – 31 March 2011 (MJDs 55287 – 55651) were analyzed. Both GPSCPTT techniques yield clock values every five minutes, whereas TWSTT yields one value every two hours. Linear interpolation of adjacent GPSCPTT values was used to perform the GPSCPTT – TWSTT subtraction.

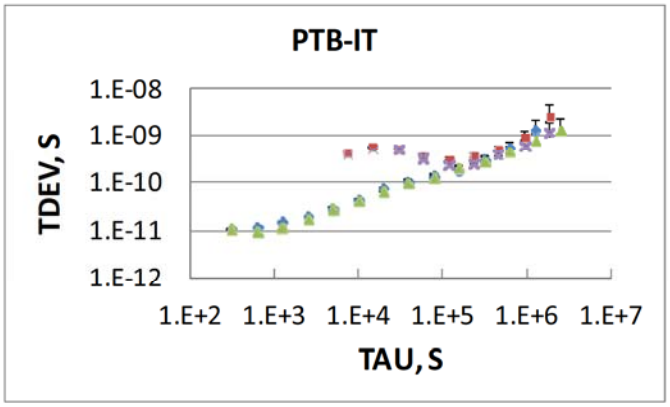
Comparing USNO GPSCPTT values to TWSTT values yields an estimate of the combined uncertainty of the GPSCPTT and TWSTT methods. Equipment and data-analysis uncertainties will be included, as well as uncertainties in models applied to only one of the methods. We use the GPSCPTT-TWSTT values to form an upper bound on our GPSCPTT uncertainties, for reasons that will become apparent in "Results."

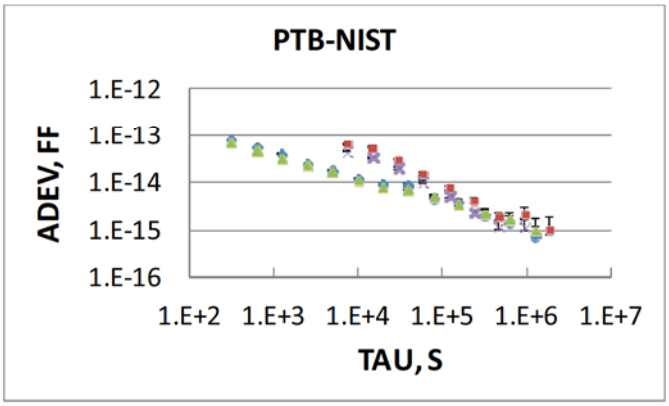
The USNO rapid estimates are not included in the IGS rapid combination (at present). Although USNO GPSCPTT results are thus independent of those obtained by the IGS, both estimate sets are GPSCPTT-based. Therefore, subtracting USNO and IGS GPSCPTT estimates will remove most hardware instabilities, as well as the hardware-and-software-based day-boundary discontinuities, because both estimates will be created using by-and-large the same raw data in which these hardware instabilities are manifest. Therefore, a comparison of these methods yields insight primarily into the uncertainty of processing models and methods: the IGS solution is a combination of analysis-center solutions, and the











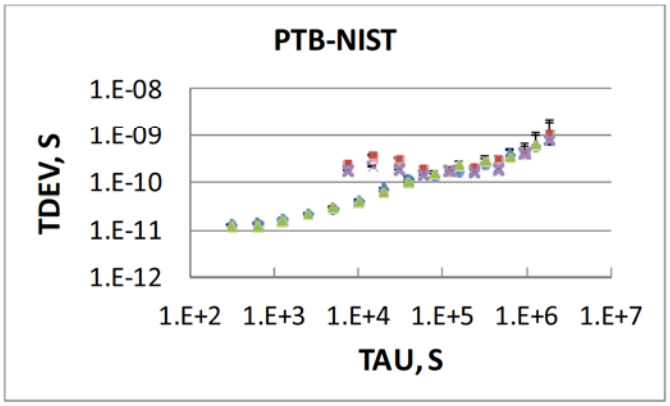


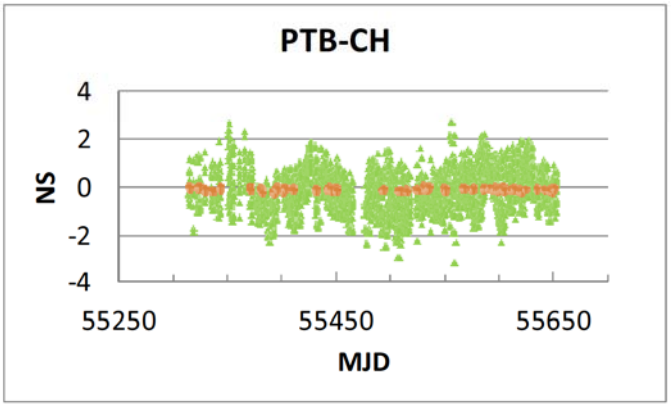
Figure 2. Frequency stability (Allan deviation; ADEV) of the Fig. 1 clock differences. Values shown are for GPSCPTT, Segment 1 and Segment 2, and TWSTT Segment 1 and Segment 2. "FF" = fractional frequency; "TAU" = averaging time. PTB-CH outlier is likely caused by small sample size at tau = $1.2 \cdot 10^6$ s.

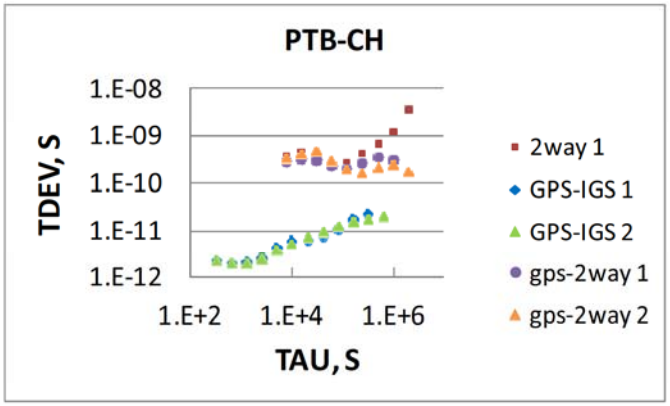
Figure 3. Time stability (time deviation; TDEV) of the Fig. 1 clock differences. Values shown are for GPSCPTT, Segment 1 and Segment 2, and TWSTT Segment 1 and Segment 2.

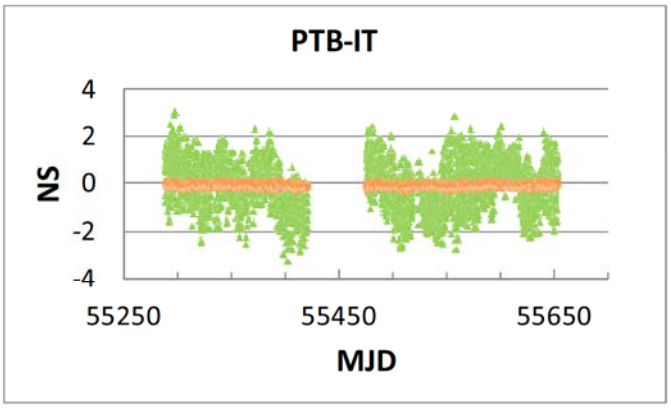
IV. Comparison of Frequency Estimates

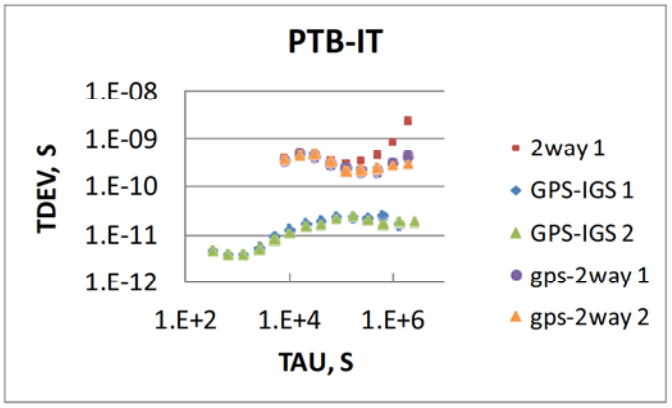
The time stability of USNO GPSCPTT estimates lies between tens and hundreds of ps, as evidenced by the GPSCPTT-IGS and GPSCPTT-TWSTT differences respectively. It remains nearly constant as averaging time increases, implying that longer data averaging yields better frequency stability. The frequency stability of USNO GPSCPTT lies between $1 \cdot 10^{-15}$(GPSCPTT-IGS) and $1 \cdot 10^{-14}$ (GPSCPTT-TWSTT) at tau = 1 d, improves with averaging time, and may reach $1 \cdot 10^{-16}$ in 10 d, based on the comparison with IGS values.

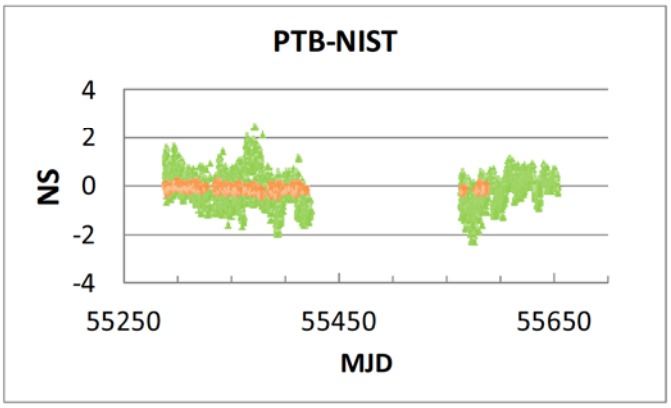
PTB, NIST and IT have cesium fountain primary frequency standards (CFPFSs) whose evaluation values and uncertainties are published in *BIPM Circular T* [5]. Table 1 shows the periods during MJD 55287 – 55651 in which a PTB CFPFS evaluation coincided with a CFPFS evaluation occurring at NIST or IT, along with the uncertainties of each











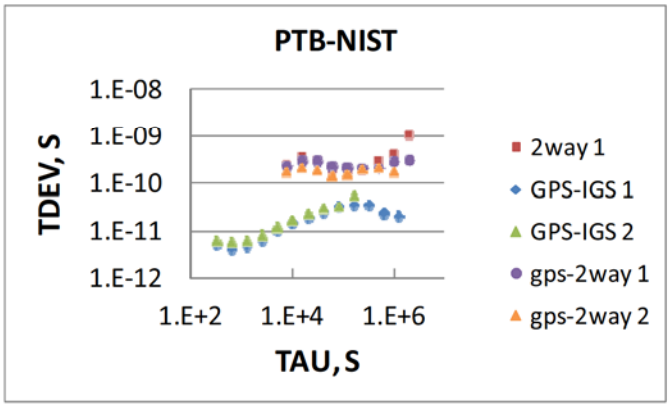


Figure 4. Time differences between solutions for clock pairs: USNO GPSCPTT – TWSTT, USNO GPSCPTT – IGS.

Figure 5. Time stability of the subtraction of solutions (Fig. 4): USNO GPSCPTT - IGS Segment 1, Segment 2; USNO GPSCPTT - TWSTT Segment 1 and Segment 2. Time stability of clock pair obtained from non-subtracted TWSTT measurements shown in red.

evaluation and the RSS combination of the two uncertainties. The uncertainties listed pertain to the entire period over which each CFPFS was evaluated – not just to the overlap time shown in Table 1 – and include the uncertainties of transfers into TAI. Thus, the RSS values of Table 1 show the least-stringent frequency uncertainty requirements GPSCPTT would have to meet to contribute new information to a joint CFPFS evaluation. A detailed treatment of frequency standard evaluation uncertainties is given in [6].

We computed frequencies from GPSCPTT, TWSTT and *BIPM Circular T* for the clock pairs and time periods shown in Table 1, assigning error bars based on the Allan deviation (ADEV) values shown in Fig. 6 or from values given in *Circular T*. We then checked if the frequency values from the three sources agreed within their uncertainties. Frequency values were computed using endpoints, i.e.,

$$y_{A\text{-}B}(t_2 - t_1) = [x_{A\text{-}B}(t_2) - x_{A\text{-}B}(t_1)]/(t_2 - t_1). \quad (1)$$

This test highlighted the need for more-reliable frequency

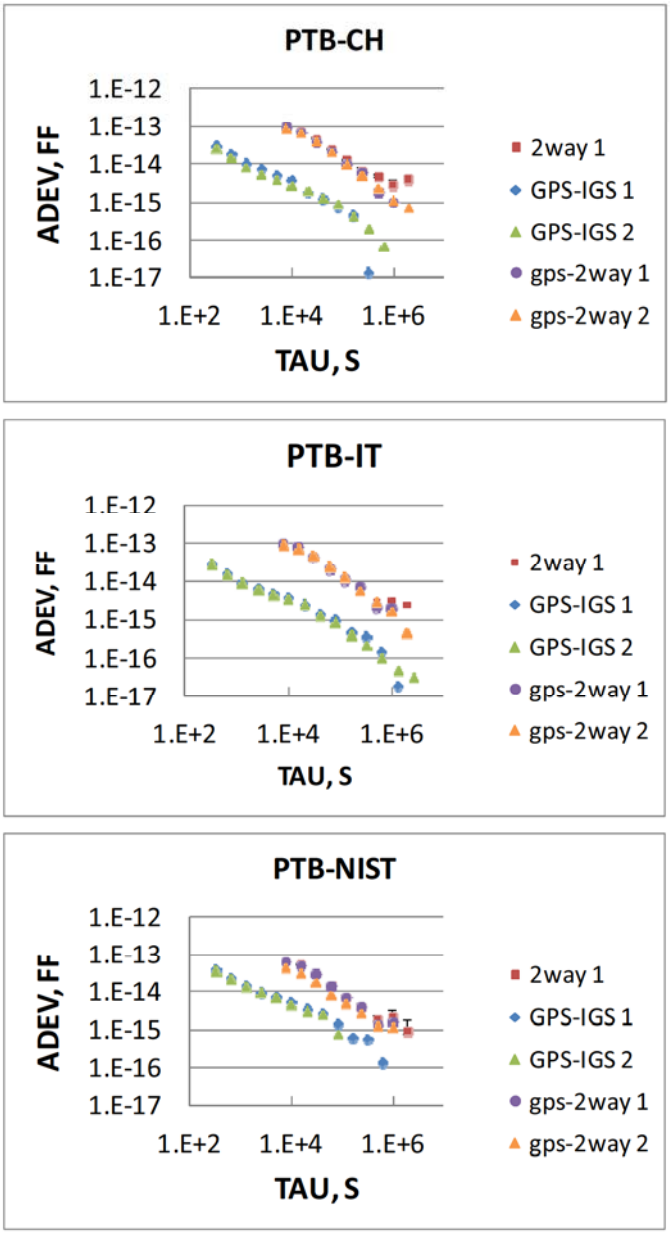


Figure 6. Frequency stability of the Fig. 4 subtraction of solutions: USNO GPSCPTT - IGS Segment 1, Segment 2; USNO GPSCPTT - TWSTT Segment 1 and Segment 2. Time stability of clock pair obtained from non-subtracted TWSTT measurements shown in red. GPS-IGS 1 outliers are likely caused by small sample sizes at large averaging times.

uncertainty values at averaging times of 10-20 d. As Fig. 6 shows, these averaging times meet or exceed the maximum for which Allan deviation values can be computed from the measurements under study. Therefore, in assigning error bars for a given averaging time, we generally used an ADEV from a smaller averaging time. Because the ADEV values (Fig. 6) decrease with averaging time, this led to conservative uncertainty estimates. Table 2 shows the uncertainties chosen

TABLE I. PRIMARY FREQUENCY STANDARD EVALUATIONS AT PTB, NIST AND IT WITH OVERLAPPING EVALUATION PERIODS, 55287 - 55651

Labs	Overlap Dates (MJD)	*Circ T* No.	Evaluation Uncertainty (10^{-15}) (evaluation time, d, in braces)		
			PTB	***IT or NIST***	***RSS***
PTB-IT	55399 409	272	0.80 (30)	0.76 (35)	1.1
	55569-89	277	0.85 (20)	0.98 (25)	1.3
PTB-NIST	55354-64	270	0.88 (15)	0.73 (20)	1.1
	55404-09	272	0.80 (30)	0.95 (15)	1.2
	55574-89	277	0.85 (20)	0.71 (15)	1.1
	55634-44	279	0.63[a] (15)	0.72 (15)	1.0

[a]PTB CSF2 evaluation. PTB CSF1 evaluated in all other cases.

TABLE II. CHARACTERISTICS OF FREQUENCY COMPARISONS COMPUTED USING USNO GPSCPTT, TWSTT AND BIPM CIRCULAR T

Link	Dates (MJD)	Averaging Time (d)	Frequency Uncertainty (10^{-15}) (ADEV tau, d, in braces)		
			TWSTT	***GPSCPTT***	***Circ T***
PTB-IT	55399-409	10	2.4 (5)	N/A	3.5
	55569-89	20	1.8 (11)	0.054 (14)	2.3
PTB-NIST	55354-64	10	1.5 (5)	1	8.3
	55404-09	5	1.5 (5)	1	16
	55574-89	15	1.4 (5)	1	5.5
	55634-44	10	1.4 (5)	1	5.8

for each frequency comparison.

Fig. 7 shows the computed frequencies and error bars. As Fig. 7a shows, the *Circular T* values have large uncertainties, and therefore both GPSCPTT and TWSTT values agree with them. Fig. 7b shows only the GPSCPTT and TWSTT values. In all cases, the two values agree, but in some cases, a good fraction of the error bars is needed to reach agreement. This indicates that the uncertainty estimates were appropriate, that (1) is not an optimal frequency computation technique given the noise types present in the data, or some combination of both.

V. DISCUSSION

[7] indicates that frequency values obtained from TWSTT using the endpoint method (1) may be improved by using end points created by averaging 24 h of TWSTT measurements. More-detailed inspection of the ADEV, Modified Allan

deviation (MDEV), and time deviation (TDEV) values of our TWSTT measurements would confirm this true or false in our case. It may also be more appropriate to use MDEV rather than ADEV as a frequency uncertainty estimator if flicker PM noise is present [8].

Day-boundary discontinuities were not removed or otherwise mitigated when computing frequencies from the GPSCPTT values. The frequency uncertainty caused by these is thought to average out after approximately 8 d [9]; it is therefore possible that these degraded some of the GPSCPTT frequency computations. While the additional uncertainty caused by day-boundary discontinuities is not reflected in the USNO GPS – IGS difference, it is reflected in the USNO GPS – TWSTT difference, and therefore would have been incorrectly assigned to the TWSTT uncertainty. Nonetheless, the oversized uncertainty of the TWSTT values would have compensated the undersized uncertainty of the GPSCPTT values, allowing the error bars to meet in Fig. 7 regardless.

The computational methods of [9, 10] circumvent the added frequency uncertainty associated with day-boundary discontinuities. However, [9] would require modification of processing algorithms, and [10] yields biased frequency values when days of data are missing (as was true in this case). Therefore, neither was attempted here.

VI. Conclusions

The time stability of USNO GPSCPTT estimates lies between tens and a few hundreds of ps, based on comparisons with TWSTT and IGS rapid combined values. This time stability remains nearly constant as averaging time increases, indicating that better frequency-estimation uncertainty can be achieved by averaging over longer periods.

The frequency stability of USNO GPSCPTT estimates is between 10^{-15} and 10^{-14} at an averaging time of 1 d based on comparisons made as above. Subtraction of USNO and IGS GPSCPTT estimates along the PTB-CH and PTB-IT links indicates that 10^{-16} stability may be achieved in 10 d.

An attempt was made to compute frequency values from TWSTT and GPSCPTT data in a manner similar to that which might be used in a primary frequency standard comparison. This indicated that (a) more study is needed to obtain reliable frequency uncertainty estimates at averaging times typical of fountain evaluations (15-25 d), (b) refinement is needed in the frequency computation techniques used (by us) and (c) a reduction of the uncertainty associated with the processing of GPSCPTT data would likely be needed as well.

Acknowledgment

The IGS provided the raw GPS measurements and a priori orbit/clock values used in USNO GPSCPTT data analysis. Constructive comments were contributed by Mr. Nick Stamatakos and the USNO editorial board.

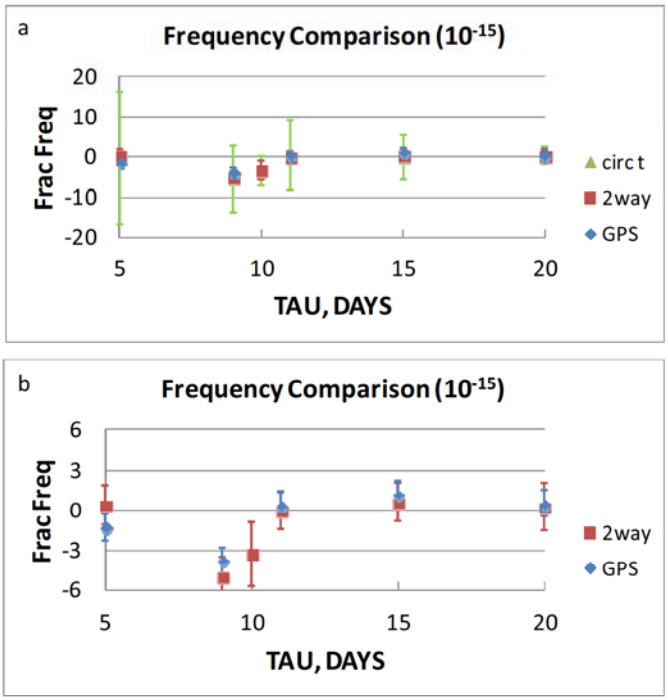


Figure 7. Fractional frequency of PTB-IT and PTB-NIST clock values as estimated using *BIPM Circular T*, TWSTT, and USNO GPSCPTT. Values with 10-day averaging times were plotted at tau = 9, 10 and 11 days for better visibility.

References

[1] http://www.igs.org

[2] http://www.bernese.unibe.ch/index.html

[3] C. Hackman and D. Matsakis, "Accuracy and Precision of USNO GPS Carrier Phase Time Transfer," *Proc. 42nd Ann. Precise Time and Time Interval Systems and Applications Meeting*, in press, 16 pp., 2010.

[4] J.F. Zumberge, M.B. Heflin, D.C. Jefferson, M.M. Watkins, and F.H. Webb, "Precise Point Positioning for the Efficient and Robust Analysis of GPS Data from Large Networks," *J. Geophys. Res., 102 (B3)*, 5005-17, 1997.

[5] http://www.bipm.org/en/publications/scientif/tfg.html

[6] T.E. Parker, "Long-Term Comparison of Caesium Fountain Primary Frequency Standards," *Metrologia, 47,* 1-10, 2010.

[7] T.E. Parker, personal communication, 2011.

[8] D.W. Allan, M.A. Weiss and J.L. Jesperson, "A Frequency-Domain View of Time-Domain Characterization of Clocks and Time and Frequency Distribution Systems," *Proc. 45th Ann Symposium on Frequency Control, IEEE*, 667-78, 1991.

[9] R. Dach, T. Schildknecht, U. Hugentobler, L-G. Bernier and G. Dudle, "Continuous Geodetic Time-Transfer Analysis Methods," *IEEE Trans. UFFC, 53(7)*, 1250-9, 2006.

[10] C. Hackman, J. Levine, T.E. Parker, D. Piester and J. Becker, "A Straightforward Frequency-Estimation Technique for GPS Carrier-Phase Time Transfer," *IEEE Trans. UFFC, 53(9)*, 1570-83, 2006.

analysis centers use a variety of software and models in creating their results. We use the difference between USNO and IGS GPSCPTT values to create a lower bound for USNO GPSCPTT uncertainty estimates.

III. RESULTS

Fig. 1 shows the time differences for the three clock pairs as measured using TWSTT and GPSCPTT. Each set of measurements was divided into two segments, 55287 – 55470 (Segment 1) and 55474 – 55651 (Segment 2), due to hardware outages/discontinuities. A mean value was removed from each segment to compensate uncalibrated hardware delays. Despite the fact that GPSCPTT provides 288 measurements per day and TWSTT only 12, the GPSCPTT results have lower short-term noise. The GPSCPTT and TWSTT values agree well long-term.

Figs. 2-3 show the frequency and time stability of the clock differences shown in Fig. 1. The stability values include both clock and measurement noise. Stability values were computed independently for Segments 1 and 2.

Figs. 2-3 show that for each clock pair, the two GPS segments have similar stability as do the two TWSTT segments. Fig. 2 shows that the frequency stability of the TWSTT measurements is worse than that of GPSCPTT in short-term, with the frequency stabilities of the two measurement sets coming together at about 3-10 d. Fig. 3 shows a similar result for the time stability of the two techniques, with the time-stability values of the two measurement sets becoming similar at 1-3 d. The per-measurement noise of the GPSCPTT measurements is around 10 ps, whereas that for TWSTT is a few hundred picoseconds.

Fig. 4 shows the subtracted values GPSCPTT-TWSTT and GPSCPTT-IGS. The GPSCPTT-TWSTT differences have a point-to-point noise of +/- 2 ns along with some long-term variations. The GPSCPTT-IGS values of Fig. 4 are miniscule in comparison, with a jitter level well under 100 ps.

Fig. 5 shows the time stability of the subtracted values, along with Segment 1 TWSTT time stability of the non-subtracted values for reference. The GPSCPTT-IGS time stability remains below 10 ps out to averaging time (tau) =4800 s, then rises to tens of picoseconds as tau increases, and then remains at that level (with some variations) as tau further increases. The GPSCPTT-TWSTT differences have a time stability of a few hundred ps which generally remains constant (minus a few variations) for all averaging times. Fig. 5 shows that the TWSTT values must be averaged for 32 – 64 hours to access the time stability of the clock pairs.

Fig. 6 shows the frequency stability of the subtracted values, which in both subtraction pairs trends downward with increased averaging time (a consequence of time stability that remains constant as averaging time increases). The GPSCPTT-IGS frequency stability values remain nearly an order of magnitude better than the GPSCPTT-TWSTT values

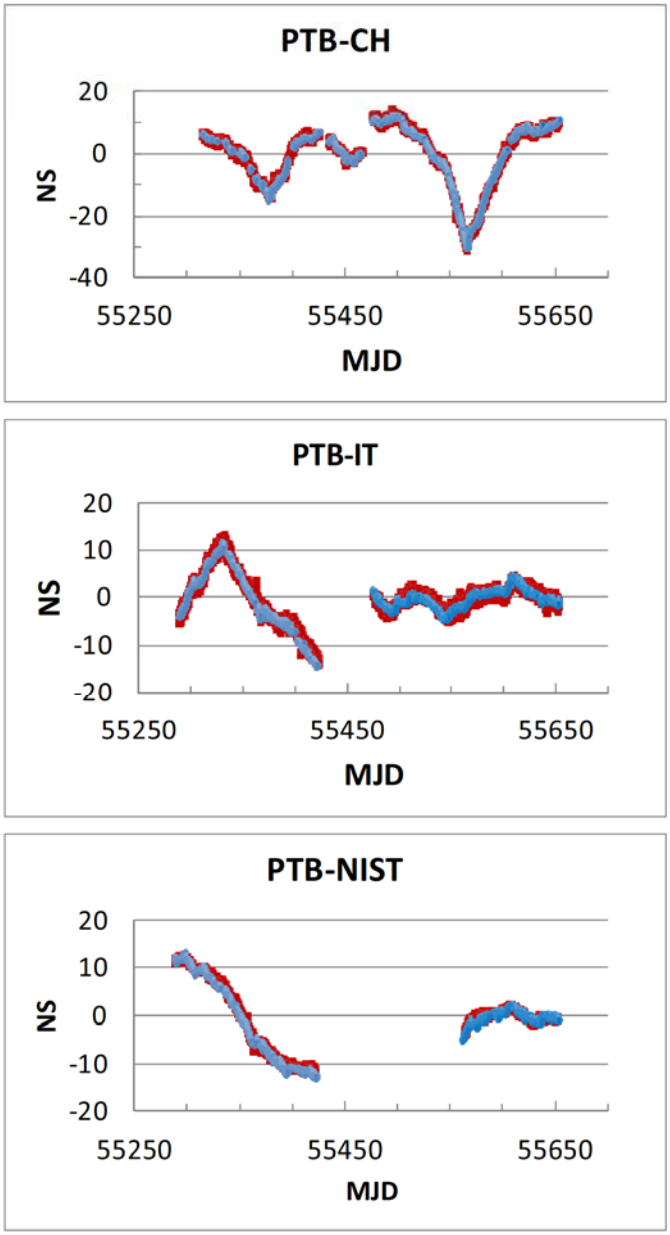


Figure 1. Time differences between clock pairs measured using TWSTT and USNO GPSCPTT. A (different) mean was removed from the two TWSTT and two GPSCPTT segments (55287 – 55470; 55474 – 55651) of each clock pair. Day-boundary discontinuities were not removed.

for averaging times past 10 d. The GPSCPTT-IGS values have a stability of $1 \cdot 10^{-15}$ or better at tau = 1 d, which decreases to $1 \cdot 10^{-16}$ at 8-10 d, except possibly along the the PTB-NIST link. The GPSCPTT-TWSTT differences have a frequency stability of about $1 \cdot 10^{-14}$ at 1 d. Along the PTB-NIST link, it is unclear whether the combined TWSTT-GPSCPTT link noise ever averages down enough to access the frequency stability of the PTB-NIST clock pair.

PPP using NRCan Ultra Rapid products (EMU): near real-time comparison and monitoring of time scales generated in time and frequency laboratories

Giancarlo Cerretto, Patrizia Tavella
Optics Division
Istituto Nazionale di Ricerca Metrologica (INRIM)
Torino, Italy
g.cerretto@inrim.it

François Lahaye, Yves Mireault
Geodetic Survey Division
Natural Resources Canada
Ottawa, Canada
flahaye@nrcan.gc.ca

Daniele Rovera
LNE-SYRTE
Observatoire de Paris
Paris, France
daniele.rovera@obspm.fr

***Abstract*— The Precise Point Positioning (PPP) time-transfer technique requires the availability of precise estimates of GNSS satellite orbits and clock offsets. Several such products are available from the International GNSS Service (IGS), each having their own characteristics: robustness, update rate, latency and satellite clock offset time interval. The most frequently updated IGS products are the Ultra Rapid products, which are generated four times a day with a latency of three hours. Natural Resources Canada (NRCan) contributes its own Ultra Rapid GPS product to the IGS for combination. However, the underlying processes running at NRCan generate products much more frequently - 24 times a day - with a latency of 90 minutes, offering an opportunity for more timely time-transfer results when used in PPP. INRIM and NRCan hereby assess the potential of using the PPP with the NRCan Ultra Rapid GPS products to serve as a short latency time-transfer tool. A specific experiment has been set up, where the NRCan Ultra Rapid GPS products, as well as all currently available IGS products, are used in PPP time transfer between selected IGS stations collocated in timing laboratories. Results and relative merits are compared in light of their respective delivery and frequency stability characteristic, in view of designing an automated near real-time monitoring system to assist timing laboratories in operational maintenance of frequency standards and time scales dissemination to external users.**

I. INTRODUCTION

The PPP time-transfer technique [1] requires the availability of precise estimates of GNSS satellite orbits and clock offsets. Several such products are available from the IGS [2], each having their own characteristics in terms of robustness, update rate, latency and satellite clock offset time interval. The most frequently updated IGS products are the Ultra Rapid products, which are generated four times a day with a latency of three hours. NRCan contributes its own Ultra Rapid product to the IGS for combination. Moreover, the underlying processes running at NRCan generate these products much more frequently, 24 times a day, with a latency of 90 minutes. When used in PPP processing, the reduced latency of Ultra Rapid products and their quality, as will be shown later, offer an opportunity for more timely time-transfer results.

Hence two objectives were sought undertaking this work: 1) assess the potential of using Ultra Rapid products in PPP to serve as a short latency time-transfer tool; and 2) prototype an automated near real-time monitoring system.

II. EXPERIMENT

In order to evaluate the time transfer capabilities of the NRCan PPP and NRCan Ultra Rapid products as a monitoring tool to assist timing laboratories in their operational maintenance of frequency standards and time scales dissemination to external users, a specific experiment has been set up.

A. Precise satellite orbit and clock products

As indicated in Table I, a list of external products have been taken into account, namely the NRCan Ultra Rapid (EMU), the IGS Ultra Rapid (IGU) and the IGS Rapid (IGR) products.

978-1-61284-110-6/11/$26.00 ©2011 IEEE

1) NRCan Ultra Rapid products (EMU)

The EMU products are generated hourly with a delay of 90 minutes after the last observation. They consist of a SP3 format orbit/clock file at 15 minutes interval for all satellites in the GPS constellation, covering a total of 48 hours, the first 24 hours being estimated from observations and the last 24 hours being an extrapolation of the orbital elements and clocks. Also included is a RINEX clock format file at 30 seconds interval for all satellites in the GPS constellation for the 24 hours estimated portion only. The EMU products generation process was described in [3].

2) IGS Ultra Rapid products (IGU)

The IGU products are generated 4 times a day with a delay of 3 hours after the last observation as a combination of up to seven individual Analysis Center solutions, including the NRCan EMU products. The IGU products also consist of a SP3 format orbit/clock file at 15 minutes interval for all satellites in the GPS constellation, covering a total of 48 hours, the first 24 hours being estimated from observations and the last 24 hours being an extrapolation of the orbital elements and clocks. However there are no associated RINEX clock format files, so the satellite clock interval is restricted to 15 minutes.

TABLE I. PRECISE SATELLITE PRODUCTS

Label	Description	Update Interval	Delay[a]	Clock Interval	Time Reference
IGR	IGS Rapid	24 hr	17 hr	300s	Timescale IGRT
IGU	IGS Ultra Rapid	6 hr	3 hr	900s	Piecewise alignment
EMU	NRCan Ultra Rapid	1 hr	1.5 hr	30s	

a. after last observation

TABLE II. IGS STATIONS SELECTED

Name	Country	Receiver Type	Clock[a]
AMC2	Colorado Springs, USA	ASHTECH Z-XII3T	HM
BRUS	Brussels, Belgium	ASHTECH Z-XII3T	HM
IENG	Torino, Italy	ASHTECH Z-XII3T	HM
NRC1	Ottawa, Canada	ASHTECH Z-XII3T	HM
NRL1	Washington, USA	ASHTECH Z-XII3T	HM
ONSA	Onsala, Sweden	JPS E_GGD	HM
PTBB	Braunschweig, Germany	ASHTECH Z-XII3T	HM
SFER	San Fernando, Spain	ASHTECH Z-XII3T	HM
SPT0	Boras, Sweden	JPS E_GGD	HM
SYDN	Sydney, Australia	JPS E_GGD	Cs
USN3	Washington, USA	ASHTECH Z-XII3T	HM
WSRT	Westerbork, Netherlands	AOA SNR-12 ACT	HM

a. HM: Hydrogen Maser; Cs: Cesium Frequency standard

3) IGS Rapid product (IGR)

The IGR products are generated daily with a delay of 17-18 hours after the last observation as a combination of up to eight individual Analysis Center solutions. The IGR products consist of a SP3 format orbit/clock file at 15 minutes interval for all satellites in the GPS constellation and a RINEX clock format file at 300 seconds interval for all satellites in the GPS constellation. All clock estimates are referred to the IGRT timescale [4].

Fig. 1 shows the median satellite orbit and clock RMS differences of EMU and IGU with respect to IGS Rapid products. It can be seen that the orbit and clock quality are comparable at the levels of 2-4cm and 80-100ps respectively. It can also be seen how the prediction degrades, especially for the clocks.

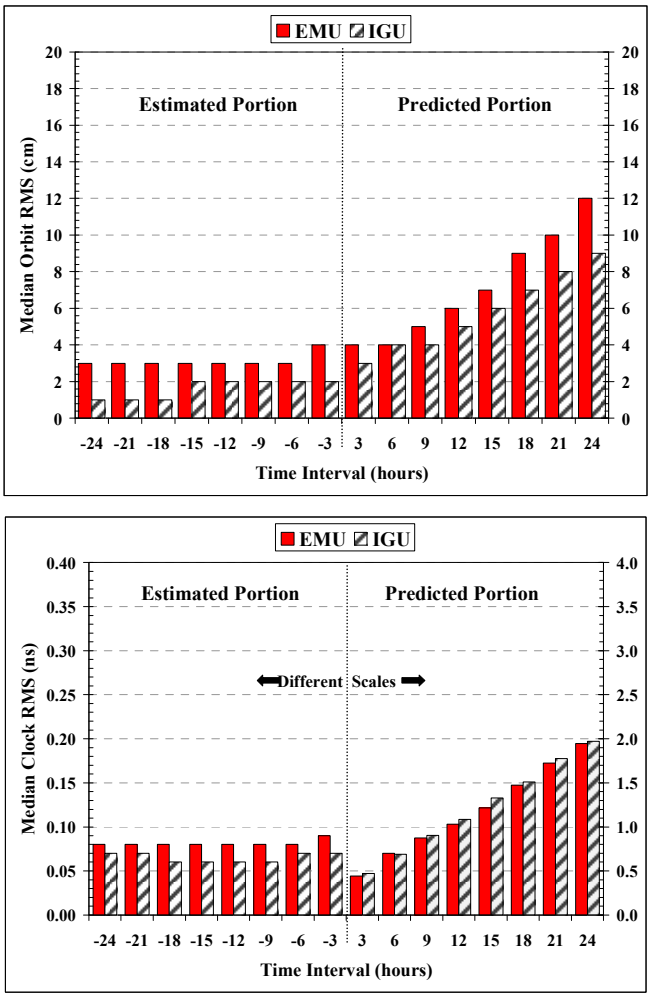


Figure 1. Median satellite 3-D position (top) and clock(bottom) RMS of EMU and IGU vs IGR.

B. IGS stations selected

As indicated in Table II, a set of 12 IGS stations, mostly co-located in time and frequency laboratories, have been selected and considered. Their daily and hourly RINEX 30 seconds observation files have been automatically retrieved and processed by means of the NRCan PPP algorithm.

C. Automated processing

In order to have an automated operational process, procedures able to periodically retrieve the required PPP input, namely the RINEX station observation files and the products described above, were developed and operated on an NRCan computer. Three independent automated processes are regularly scheduled:

- a process gathering the hourly/daily RINEX observation files for the selected 12 stations, running three times per hour, namely 10, 30 and 50 minutes after the hour;
- a process gathering the orbit and clock products (EMU, IGU and IGR), running every 10 minutes, namely 0, 10, 20, 30, 40 and 50 minutes after the hour;
- a process checking the presence of specific instances of each product within their availability window, running every 10 minutes, namely 5, 15, 25, 35, 45 and 55 minutes after the hour; this process triggers PPP processing whenever a specific product instance is detected.

As soon as a specific product instance is available, the PPP processing for the 12 selected stations is scheduled using the available data and specified products. For each of the product types, the multiday product selection is as follows:

- the EMU products for each of the initial 5 days, plus the latest EMU product, covering the hours of the additional day (1, 2, 3, … , 24 hours);
- the IGU products for each of the initial 5 days, plus the latest IGU product covering the hours of the additional day (6, 12, 18 and 24 hours);
- only those data files containing the full 6 days are processed with the corresponding IGR products.

The station RINEX data files are concatenated to span the previous 5 days and the first n hours of current day corresponding to the current product instance time span (estimated portion only in the case of Ultra Rapid products). This process provides multiple RINEX data file for each station per day, each file n hours longer than the previous and the last file covering the full 6 days.

The NRCan PPP version considered for this experiment is Version 1.5, Release 06010, characterized by the following new features:

- 2003 IERS tide model;
- Global Pressure & Temperature (GPT) model [5] for hydrostatic troposphere delay;
- Kouba's simplified yaw/attitude model [6];
- adaptive widelane & narrowlane thresholds for cycle clip detection;
- kinematic antenna orientation tracking;
- station clock constraints;
- GLONASS-only and GPS+GLONASS processing.

The multi-day solutions were generated using dual-frequency pseudorange and carrier-phase with a 10° elevation mask and exercising the adaptive widelane and narrowlane cycle slip detection. One static position was estimated for the whole multiday period with earth tides and ocean loading applied. Hydrostatic tropospheric delays were modeled with Hopfield using pressure and temperature from the GPT model and 50% relative humidity. The wet troposphere delay component was estimated with 0.05m/√hr process noise as well as North and East hydrostatic delay gradients. Finally the station clocks were estimated at the interval of the satellite clock products, without constraints. Note the EMU 30 second RINEX satellite clock files were decimated to 300 seconds to speed-up the processing. No GLONASS observations or products were used herein.

The NRCan PPP station clock estimates related to each product are respectively labeled PPP(IGU) at 900 seconds interval, and PPP(IGR) and PPP(EMU), at 300 seconds interval.

III. RESULTS

The three automated processes were operated continuously over 6 weeks, namely for the MJD 55399-55441 period (2010 July 22 – 2010 September 2; GPS Week 1593(day 4) – 1599(day 4)), yielding 8844 PPP(EMU) solutions, 1483 PPP(IGU) solutions and 450 PPP(IGR) solutions. All solutions were automatically graphed and posted on a private WWW page.

Over the experiment period there were many missed hourly RINEX observations files but no product outages. The delay in obtaining monitoring results was 1h40-2h00 for PPP(EMU) solutions, 3h20 to 3h40 for PPP(IGU) solutions and 17hours to 18 hours for PPP(IGR) solutions. There were two problems in PPP solutions where receiver tracking malfunctions were misidentified for 1ms jumps in data, causing the resulting clock phase graphs to be unreadable.

A. Effect of clock product time reference

The top two panels of Fig. 2 shows PPP(EMU) results for two single stations, USN3 and AMC2. The rough alignment of the time reference of the Ultra Rapid clock products to GPS

time is quite clear, the clock phase and frequency jumps being readily apparent. For a meaningful clock monitoring application the single station PPP clock results will need to be differenced between stations to remove the discontinuous time reference, as shown in the bottom panel of Fig. 2.

B. Monitoring results

For visual appreciation of the consistency of the monitoring results, Fig 3 shows examples of all 6-day PPP(EMU) solutions ending on Sundays and Wednesdays at 0h00 for all baselines between USN3, AMC2 and BRUS stations. Each solution thus overlaps by 3 or 4 days with the previous one. The PPP(IGU) solutions are similar but at 900 seconds interval rather than 300 seconds.

The AMC2-BRUS and USN3-BRUS baselines show a slow frequency drift followed by a correction in the later part of August 2010.

TABLE III. IGR-EMU DOUBLE DIFFERENCE STATISTICS

Baseline	Bias (ns)	SDev[a] (ns)	RMS (ns)	Max(|Δφ|) (ns)	Max(|Δf|) 10^{-15}
USN3-AMC2	0.037	0.063	0.073	0.285	1.66
USN3-BRUS	-0.075	0.109	0.132	0.522	1.87
USN3-IENG	0.017	0.111	0.112	0.331	3.32
USN3-NRC1	0.018	0.066	0.068	0.321	1.88
USN3-NRL1	0.021	0.059	0.063	0.311	0.55
USN3-PTBB	-0.096	0.107	0.144	0.443	1.92
USN3-SFER	-0.152	0.145	0.210	1.232	2.15
USN3-SPT0	-0.011	0.112	0.113	0.398	2.10
USN3-SYDN	-0.083	0.350	0.360	1.472	5.84

a. standard deviation about mean bias

TABLE IV. IGR-IGU DOUBLE DIFFERENCE STATISTICS

Baseline	Bias (ns)	SDev[a] (ns)	RMS (ns)	Max(|Δφ|) (ns)	Max(|Δf|) 10^{-15}
USN3-AMC2	-0.017	0.075	0.077	0.240	1.16
USN3-BRUS	-0.074	0.159	0.175	0.503	2.33
USN3-IENG	-0.040	0.136	0.142	0.441	2.06
USN3-NRC1	-0.005	0.066	0.066	0.442	1.52
USN3-NRL1	0.004	0.092	0.092	0.520	2.71
USN3-PTBB	-0.135	0.138	0.193	0.509	1.82
USN3-SFER	-0.133	0.179	0.223	1.542	2.70
USN3-SPT0	-0.028	0.141	0.144	1.037	2.83
USN3-SYDN	-0.140	0.200	0.244	3.035	4.68

a. standard deviation about mean bias

C. Quality of PPP(EMU) and PPP(IGU)

The PPP(EMU) and PPP(IGU) baseline solutions involving USN3 were differenced with the corresponding PPP(IGR) solutions for all 6-day solutions ending on Sundays and Wednesdays 0h00 during the experiment campaign. The mean bias, standard deviation about mean bias, RMS and maximum absolute value of the double differences are reported in Table III and IV and examples are given in Fig. 4.

Most baselines show biases of less than 100ps and standard deviations of less than 150ps. In Fig. 4 some baselines show signs of frequency biases. The frequency bias was estimated for each solution by differencing the first and last double difference clock phases and their maximum absolute value is reported in Table III and IV. In most cases the frequency bias does not exceed few parts in 10^{15}.

IV. CONCLUSION

NRCan PPP estimates generated using the NRCan and IGS Ultra Rapid products show a good agreement with respect to those obtained using IGS Rapid products. The level of performance is very promising, especially taking into account the reduced latency of the station clock product of about 2 hours when using NRCan products. The monitoring process can detect nanosecond level clock phase jumps and few parts in 10^{15} clock frequency changes. Although the processing is demanding in terms of computation time and robustness, it was possible to effectively automate the whole process, from data collection to plotting of the stations clock differences. Use of several products is recommended, as was done herein, to cover for potential product outages. The station clock differences obtained using Ultra Rapid and Rapid products could be posted in an dedicated webpage to provide the timing community timely information, helping in maintaining an objective level of "quality of service" for the time and frequency laboratories.

ACKNOWLEDGMENT

The authors thank the many institutions and individuals, including NMI, for their efforts in voluntarily contributing data and products to the IGS [2]. Precise single station GNSS application such as the Precise Point Positioning would not exist without them.

REFERENCES

[1] J. Kouba, P. Héroux, "Precise point positioning using IGS orbit and clock products," GPS Solutions, vol. 5, no. 2, pp. 12–28, 2001.

[2] J.M. Dow, R. E. Neilan, C. Rizos, "The International GNSS Service in a changing landscape of Global Navigation Satellite Systems," Journal of Geodesy, vol. 83, pp.191–198, 2009.

[3] Y. Mireault, P. Tétreault, F. Lahaye, P. Héroux, J. Kouba, "Online Precise Point Positioning: A New, Timely Service from Natural Resources Canada.," GPS World, vol. 19, no. 7, September 2008.

[4] K. Senior, P. Koppang, J. Ray, "Developing an IGS Time Scale," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 50, no. 6, pp. 585-593, 2003.

[5] J. Boehm, R. Heinkelmann, H. Schuh, "Short Note: A global model of pressure and temperature for geodetic applications," Journal of Geodesy, vol. 81, no. 10, pp. 679–683, 2007.

[6] J. Kouba, "A simplified yaw-attitude model for eclipsing GPS satellites," GPS Solutions, vol. 13, no. 1, pp. 1-12, 2009.

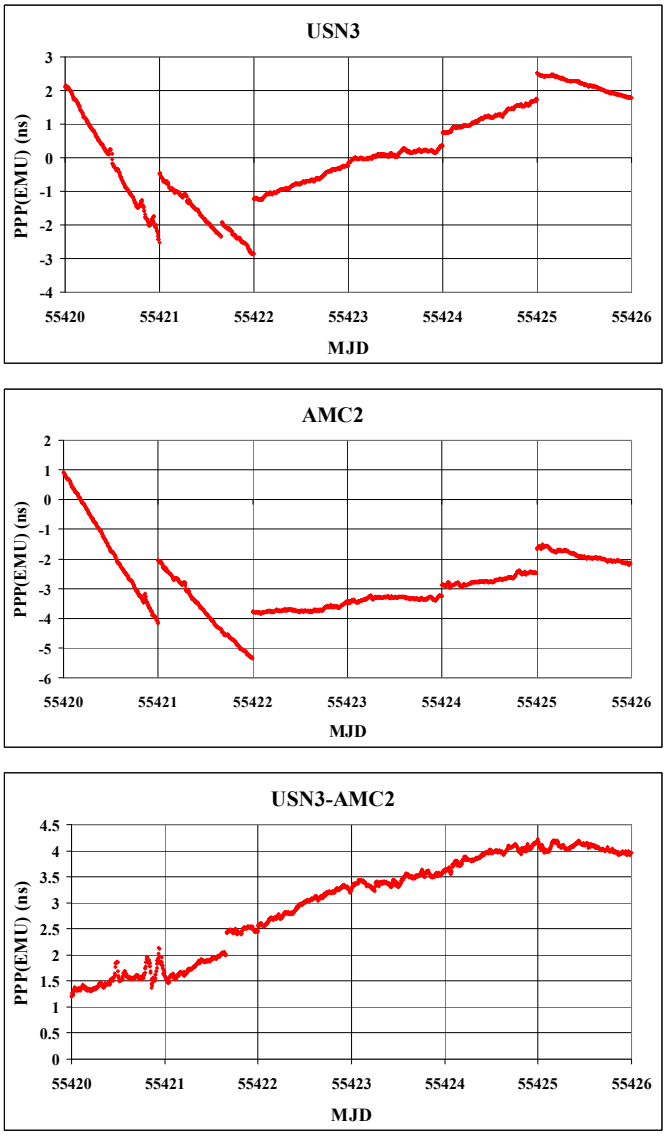


Figure 2. Single station PPP(EMU) solutions for USN3 (top) and AMC2 (middle) and between station differences (bottom).

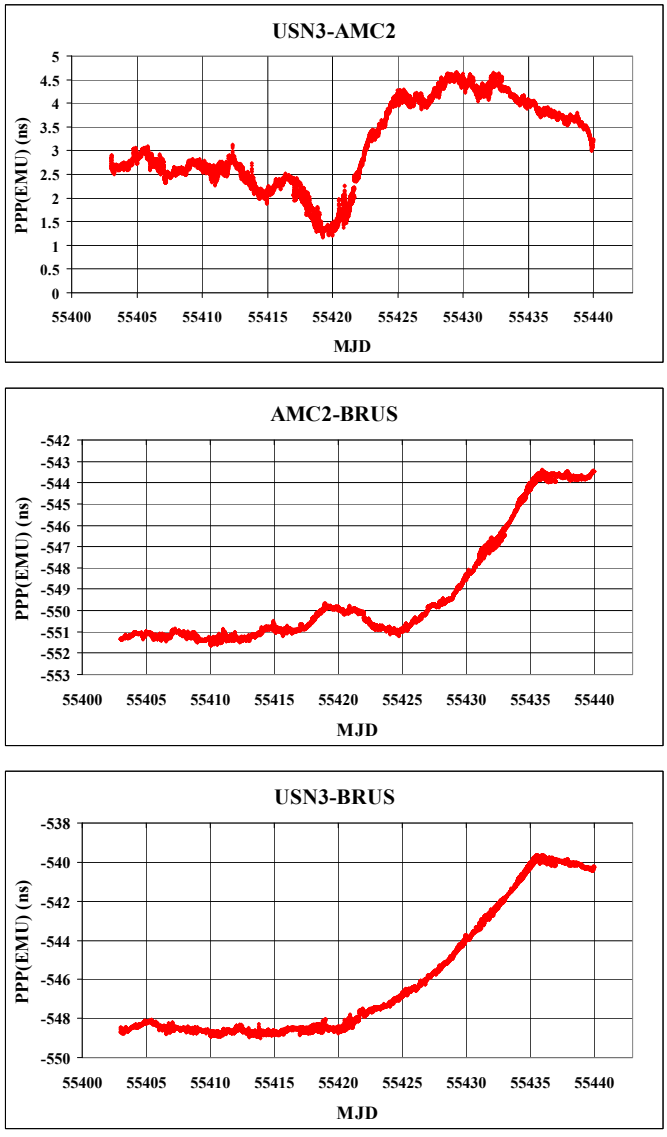


Figure 3. Between station differences results for three stations PPP(EMU) solutions.

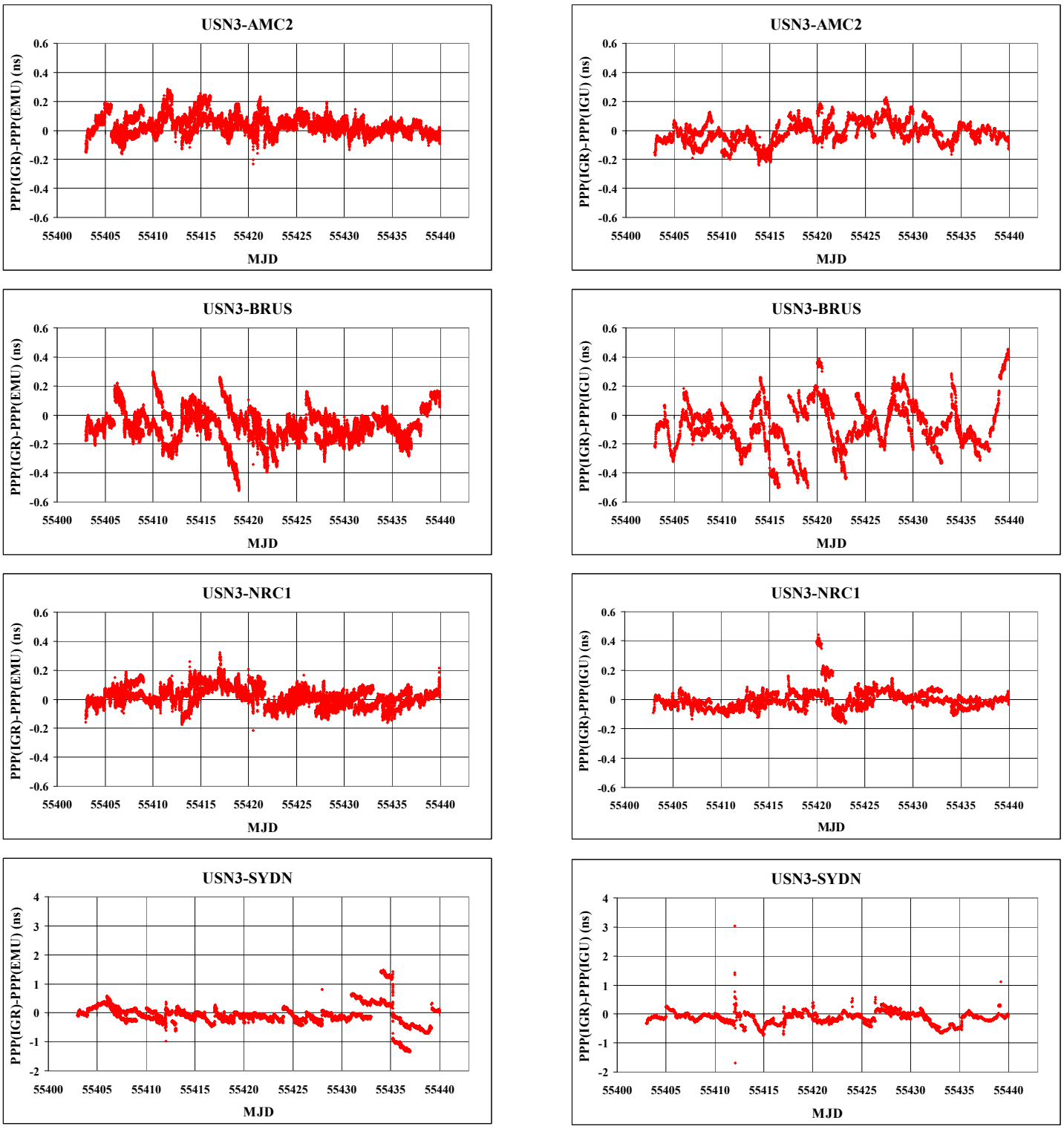


Figure 4. Double difference clock phase results of PPP(EMU) vs PPP(IGR) (left) and PPP(IGU) vs PPP(IGR) (right) for selected baselines with USN3.

Absolute calibration of GNSS time transfer systems : NRL and CNES techniques comparison

Amandine Proia [1,2], Gilles Cibiel[1]
[1]Centre National d'Etudes Spatiales
Toulouse/France
[2]Bureau International des Poids et Mesure
Sèvres/France
E-mail :amandine.proia@cnes.fr

J. White, D. Wilson, K. Senior
Naval Research Laboratory
Washington/United States

Abstract — **The GNSS reception chains (timing receiver + antenna + antenna cable) are used to perform time comparisons of remote atomic clocks. This time transfer system represents the basis of the time laboratories contributions for the realization of the Temps Atomique International (TAI). The calibration of these chains are necessary to ensure accuracy and long term stability of time links used in TAI but also in precise time station (PTS) dedicated to GALILEO.**

Currently, the most widely used approach to determine the electrical-delay of these devices is the differential method developed by the BIPM (Bureau des Poids et Mesures). Another solution is the absolute calibration of each reception chain element in using artificial signals. This method was first defined and put into operation by NRL (Naval Research Laboratory) and the University of Colorado. In 2005, CNES (Centre National d'Etudes Spatiales) decided to develop this technique with a similar approach.

This paper will compare the results of NRL and CNES absolute calibrations performed on two different time transfer systems. First of all, it will describe the calibration method used by each laboratory and the differences which characterize them. An uncertainty budget will be presented for the results of both calibration techniques. Then, both time transfer systems, composed of different receivers and antennas models, will be described briefly. Finally, the absolute calibrations results performed using the NRL and the CNES methods will be compared..

I. INTRODUCTION

The construction and the diffusion of the time scales rely on the utilization of navigation and telecommunication satellites system. Thanks to this ease of access and use, the GNSS (Global Navigation Satellites System) time and frequency transfer is the most used tool to compare remote clocks. The present GNSS reception chains (multi-channel, multi-frequency receiver + cable + antenna) and the evolution of data treatment techniques (P3 or PPP methods) allowed remote comparison of clocks with a statistical uncertainty of a few nanoseconds [1, 2]. Currently, the global uncertainty of these measurements is dominated by the systematic uncertainty of the reception chain calibration. This measurement is ensured by the BIPM (Bureau International des Poids et Mesures) in using the differential calibration [3]. One GNSS system is designated as the reference and is in constant circulation among time laboratories. A relative calibration is performed between this reception chain and the laboratory GNSS equipment where both are referenced to the same clock. With this method, the electrical delay of the GNSS implement is determined with an uncertainty of about 2.5 ns for P1 or P2 (k=1).

Another method, the absolute calibration, using artificial signals showed it is possible to calibrate the reception chain with an uncertainty budget inferior to 1 ns (k=1) [4,5].

This paper proposes a comparison of absolute calibrations performed at NRL (Naval Research Laboratory) and at CNES (France space agency) with two different GNSS reception chains. The first part is dedicated to the receiver calibration method when the second part describes the antenna calibration method. In both cases, the general principle of methods and the differences characterizing the approaches used by CNES and NRL will be presented. Finally, a calibrations comparison of the different reception chains performed at CNES and NRL will be discussed in the last part.

II. RECEIVER CALIBRATION

The absolute calibration method was first defined and put into operations by NRL [6] and the University of Colorado [4]. Since 2005, CNES is developing this method with a similar approach [5,7].

A. *Principle*

The approach of the receivers calibration consists in an artificial reception free of delays, effects and noises upstream at the output of antenna: atmospheric delays (troposphere and ionosphere), multipath effects or antenna delay. These conditions can be only conducted with a GNSS Signal Simulator (GSS) (Figure 1). The GSS allows to simulate the signals emitted by the GNSS satellites in real conditions. It is able to produce the GNSS signal but also all the noises resulting of natural reception (signal propagation or clock error,…).

The absolute calibration needs a emitted signal without noises. The simulator generates only the navigation message

and two carrier frequencies: L1 modulated by the C/A and P codes and L2 modulated by P code.

The power of the radiofrequency (RF) signal is -70 dBm and must be adjusted to correspond to the reception level acceptable by the receiver (-110 dBm generally). It provides also a 1PPS (1 Pulse Per Second) signal which lets the receivers to be synchronized with the simulator.

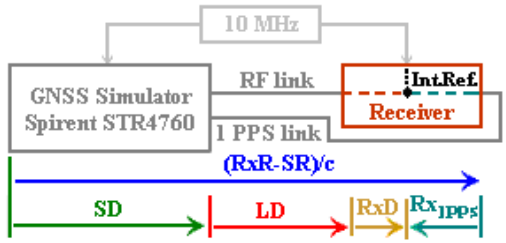


Figure 1. Schematic of receiver absolute calibration

The internal electrical time delay of the receiver is calculated thanks to (1):

$$RxD = \frac{RxR - SR}{c} - LD - SD + Rx_{1PPS} \qquad (1)$$

where:
- *RxD:* Receiver delay [ns]
- *RxR-SR:* Difference between the receiver and simulator pseudo-ranges [m]
- *c:* Light celerity (299 792 458 $m.s^{-1}$)
- *LD:* 1PPS and RF links delays difference (LD_{RF}-LD_{1PPS})
- *SD:* Simulator delay [ns]
- Rx_{1PPS}: Time delay between the receiver internal reference and the external 1 PPS [ns]

During the calibration measurement, due to their temperature sensitivity, the receivers and the simulator are located in a room regulated in temperature.

B. Main differences between both techniques

There are several hardware differences between the CNES and the NRL calibration benches. The NRL receiver calibration bench is equipped with a GPS signal simulator Spirent GSS8000. The reference signals are generated by a Symmetricom 5071A clock. At CNES, the simulator used is a Spirent STR4760 and a MASER EFOS supplies the reference signals. The oscilloscope used by CNES and NRL is the same model: a fast digital Tektronix DPO7000.

At NRL, the calibrations are performed in room where the temperature is equal to 297 ± 3 K (24°C). The temperature regulation of the CNES laboratory is of 293 ± 1 K (20°C).

Three points differentiate the CNES and NRL methods used to calibrate the time receivers:

- *The simulator calibration method*

One of principal differences between the NRL and CNES absolute methods is the way to calibrate the simulator. The internal delay of the simulator corresponds to the delay between the beginning of the PRN C/A code and the internal 1 PPS synchronized with the GPS time. In amplifying a single channel C/A code, the time difference between the rising 1PPS and the code transition can be measured with an oscilloscope.

The NRL method consists in determining the time difference by direct reading (Tick-to-Code) with an oscilloscope. With this approach, the measurement uncertainty is about 0.2 ns [4]. The CNES method is based on the same measurement but where a post-treatment is added. The technique used is the correlation which looks like a standard GPS signal acquisition. It consists of determining the code offset versus the PPS generated by the simulator and the residual carrier phase offset. This method improves upon the simulator calibration because the time delay is estimated over more than one code transition [5]. The measurement uncertainty is the integration time of the correlation which is equivalent to the oscilloscope resolution. The oscilloscope used has a resolution of 0.1 ns which will be therefore the uncertainty of the correlation.

The CNES uses a specific function of the GSS software : the "Single Channel" setting. This program generates one signal C/A on one channel without the navigation message. for To calibrate the simulator, the NRL uses a scenario of a geostationary satellite because for the Spirent simulator models 7700 and higher the "single channel utility" gives incorrect results. It works fine on the 2760 and 4760. It was checked that both methods give the same results on a STR4760 simulator.

- *The pseudo-ranges of the simulated constellation*

The NRL constructs the pseudo-ranges of the simulated constellation from the navigation files generated by the receiver. The CNES uses the pseudo-ranges directly supplied by the simulator. The calibrations comparison give the same results in using both treatments with the GSS8000..

- *The utilization of the RF and 1 PPS links*

At CNES, the links used for the simulator and the receiver calibration are not the same. Two different cables are used to transmit the 1 PPS signal between the simulator and the oscilloscope (simulator calibration) and between the simulator and the receiver (receiver calibration). Two RF links are also used. During the receiver calibration (RxD), the RF link is composed of a cable, attenuators and a bloc DC to obtain a acceptable power level for the receivers. For the simulator calibration (SD), the RF link is different, consisting of a cable, an isolator and an amplifier to visualise the signal with an oscilloscope. The RF and the 1 PPS links of the NRL absolute calibration bench are the same for the simulator and the receiver calibration. The RF link is made up of a cable, an amplifier, variable attenuators and a DC block.

C. Uncertainty budget of the receiver calibration

- CNES typical uncertainty budget

The uncertainty of the receiver delay is the quadratic sum of the measurement uncertainty of the simulator calibration, the RF and 1 PPS links, the pseudo-ranges difference and the

parameter Rx_{1PPS}. The uncertainty varies according to the receiver kind [8]. This difference is due to the calculation and the measurement technique of the Rx_{1PPS} parameter which is defined by each manufacturer.

TABLE I. UNCERTAINTY BUDGET OF DIFFERENT RECEIVERS ABSOLUTE CALIBRATION

Uncertainty source	Receivers calibration uncertainty [ps] (k=1)	
	Septentrio	*Dicom*
SD (Simulator STR4760)	351	351
LD_{RF} (RF link)	36	36
LD_{HF} (1 PPS link)	25	25
RxR-SR	135	161
Rx_{1PPS}	184	-
U(receiver)	**421**	**389**
Expanded U (Rx)	**842**	**778**

The simulator uncertainty takes into account the uncertainties of RF and 1 PPS links measurement, the correlation uncertainty (oscilloscope resolution), the pseudo-ranges error and interchannel bias (provided by the constructor). The RF link delay is measured through a VNA (Vector Network Analyzer). The 1 PPS link delay is determined with a Time Counter Interval in employing the double weight method. For the Septentrio receiver, the Rx_{1PPS} is measured with a time interval counter. It is not necessary to determine the Rx_{1PPS} value for the Dicom receiver because the output data are automatically corrected of this bias at each GPS acquisition.

- NRL typical uncertainty budget

The uncertainty of the receiver delay is the quadratic sum of the measurement uncertainty of the Tick-to-Phase, the Tick-to-Code, the delays of RF pads, in the connectors, in the DC block and the simulator accuracy [4].

TABLE II. ERROR ESTIMATE OF THE RECEIVER ABSOLUTE CALIBRATION

Uncertainty source	Error estimate [ps] (k=1)
Tick-to-Code	200
Delay in Rf pads	100
Delay in the connectors	100
Delay in the DC block	100
Simulator accuracy	1000
U(receiver)	**1000**
Expanded U (Rx)	**2000**

III. ANTENNA CALIBRATION

A. *Principle*

The antenna calibration consists of estimating the time-delay due to the signal propagation through the antenna. It will be therefore necessary to determine the input and output ports of the signal in the device. An antenna allows the transition of the electromagnetic wave between a guided environment (cable) and a free environment (air). One of the access ports will be defined by the connector (antenna/cable). The other port is defined by the position of the phase center which is the interface point between the guided environment and the free space.

Antennas used in GNSS reception chain for the time transfer are active (included an internal amplifier) and have a Right-Hand Circular Polarization (RHCP). The amplifier does not allow the reciprocal signal crossing through the antenna. To calibrate this kind of equipment, it is necessary to use the method in transmission presented in Figure 2. This experiment is performed in an anechoic chamber in order to limit the multi-path effect. The measurement of the transmission delay between both ports of a VNA gives access to the value of the antenna group delay. The signal supplied by the VNA is transmitted by a passive antenna to the receiving antenna (element to calibrate). The receiving antenna sends then the signal to the second port of the VNA.

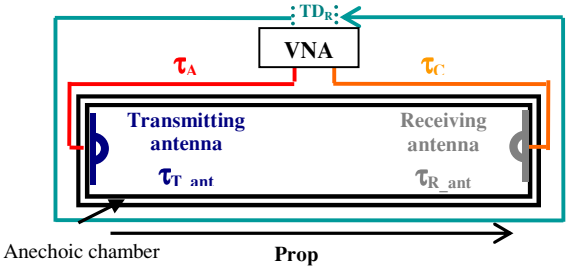


Figure 2. Schematic of a GNSS antenna absolute calibration

In using the method in transmission, the group delay of the receiving antenna is defined by the following equation (2):

$$\tau_{R\_ant} = TD_R - \tau_A - \tau_{T\_ant} - Prop - \tau_C \qquad (2)$$

where:

- $\tau_{R\text{-}ant}$: Electrical delay of the receiving antenna [s]
- TD_R: Total time delay of the transmission measurement [s]
- τ_X: Electrical delay of the cable X [ns]
- $\tau_{T\text{-}ant}$: Electrical delay of the transmitting antenna [s]
- *Prop:* Delay due to the signal propagation in the air [s]

B. *Main differences between both techniques*

There are also hardware differences between the NRL and CNES antenna calibration benches. The VNA used by CNES is a Wiltron 37200A. The transmitting antenna is a GPS antenna 704-X. The anechoic chamber is covered of *Eccosorb CV* absorbents that the rejection rate at 1.5 GHz is of 10 dB.
And another hand, the NRL uses a Broadband Dual Polarized transmitting antenna, a TECOM antenna (1-4 GHz). The NRL VNA is a Agilent 8722ES. The reference of the anechoic chamber absorbents is not known.

The principal difference between the NRL and CNES approaches to determine the internal delay of a GNSS antenna is the calibration of the transmitting antenna.

- *The transmitting antenna calibration*

To perform the antenna calibration with the method in transmission, the delay of the transmitting antenna must be known; thus it must be calibrated. Two methods are available to calibrate a passive antenna: the method in transmission (with two identical antennas or with the three-cornered hat method) or the method in reflection.

CNES uses a GPS antenna with a right circular polarization. To calibrate this antenna, it is necessary to use a measurement in transmission. With both identical antennas, the group delay of the antenna is defined by the following equation (4):

$$\tau_{T\_ant} = (TD_R - \tau_A - Prop - \tau_C)/2 \quad (4)$$

where:
- τ_{T_ant}: Electrical delay of the transmitting antenna [s]
- TD_R: Total time delay of the reflection measurement [s]
- τ_X: Electrical delay of the cable X [s]
- *Prop*: Delay due to the signal crossing in the air [s]

The NRL uses a bipolarized antenna that can be calibrated with the method in reflection. During this measurement, the signal supplied by the VNA and emitted by the antenna is reflected by an element placed downstream to the output port at a known distance.

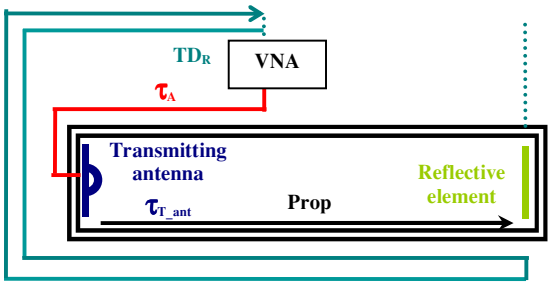


Figure 3. Antenna calibration with the method in refection

In using the method in reflection, the group delay of the transmitting antenna is defined by the following equation (5):

$$\tau_{T\text{-}ant} = (TD_R/2) - \tau_A - Prop \quad (5)$$

where:
- τ_{T_ant}: Electrical delay of the transmitting antenna [s]
- TD_R: Total time delay of the reflection measurement [s]
- τ_A: Electrical delay of the cable A [s]
- *Prop:* Delay due to the signal crossing in the air [s]

C. *Uncertainty budget of the antenna calibration*

- CNES typical uncertainty budget

The CNES antenna calibration uncertainty in using the method in transmission is composed of the measurement uncertainty of the total delay, the RF links, the distance between both phase centers and the uncertainty calibration of the transmitting antenna..

TABLE III. TYPICAL UNCERTAINTY OF THE ANTENNAS ABSOLUTE CALIBRATION

Source uncertainty	**Antenna calibration uncertainty [ps] (k=1)**		
	Novatel 702-GG	***Ashtech Choke-ring***	***Leica Choke ring 3D***
TD_T	301	308	309
τ_A/τ_C	36	36	36
Prop	210	210	210
$\tau_{T\text{-ant}}$	267	267	267
U(Antenna)	**456**	**461**	**461**
Expanded U (Ant)	**912**	**922**	**922**

The total delay uncertainty corresponds to the instrument measurement noise, the VNA. The propagation error takes into account the variations of the phase centers and the measurement error of the distance between both antennas. The time delay of the transmitting antenna was determined with the method in transmission with two identical antennas.

- NRL typical uncertainty budget

The NRL uncertainty estimation of the antenna calibration takes into account the errors due to the cable A, the anechoic chamber, the cable C and the total delay measurement. The total uncertainty of the antenna calibration is the quadratic sum of these error sources. [4].

TABLE IV. ERROR ESTIMATE OF ANTENNAS ABSOLUTE CALIBRATION

Source uncertainty	**Error estimate [ps] (k=1)**
CableA /Chamber	100
Cable C	100
Tota delay	100
U(Antenna)	**200**
Expanded U (Ant)	**400**

IV. RESULTS COMPARISON

The Figure 4 presents the results of the receivers absolute calibration in using the NRL and the CNES methods. This comparison was performed for a Dicom GTR50 receiver and a Septentrio PolaRx receiver.

The NRL measurement will be taken as the comparison reference for all the graphs presented in this section.

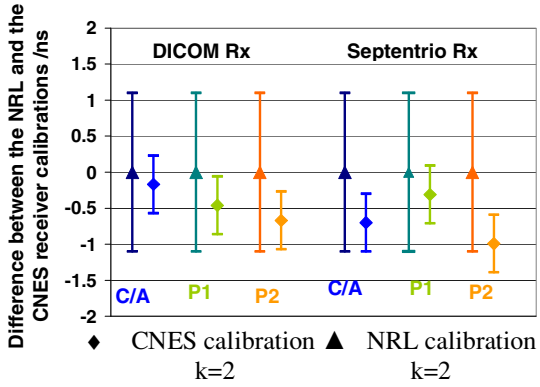


Figure 4. Comparaison between CNES and NRL receiver calibration

The results of the CNES and NRL receiver calibrations are of the same order of magnitude although all the CNES measurements are inferior to the NRL measurements. This difference could arise from the simulator thermal sensitivity. The CNES measurement points and their uncertainties are enclosed in the NRL error bars except for the Septentrio receiver P2 point. Only one measurement of the Septentrio receiver calibration was performed at NRL. It is therefore impossible to confirm the truthfulness of this value.

The figure 5 presents the results of the antennas absolute calibration with the NRL and the CNES method. These calibrations are performed with two types of antennas:

- A Novatel 702-GG antennas
- An Ashtech Choke-ring

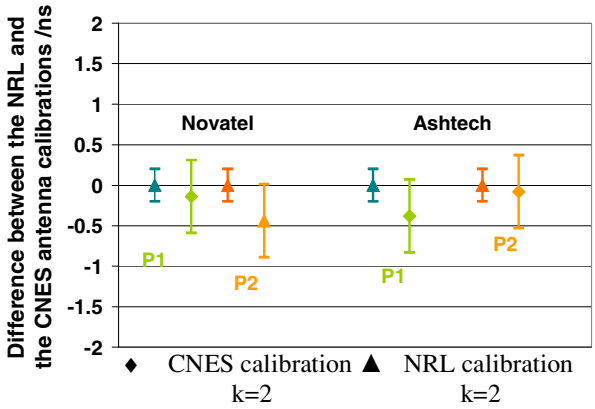


Figure 5. Comparaison between CNES and NRL antenna calibration

The antenna calibration results are in the same order of magnitude. There is an overlap of the CNES and NRL uncertainty bars but not a complete lapping for all measurements. The reproducibility is not also excellent than for the receiver calibration. The explanation can be arisen from the NRL and CNES measurement environments. At NRL the anechoic chamber is aging and the absorbents performance deteriorate. Moreover, the absorbents rejection rate of the CNES anechoic chamber is not very high. The replacement of the absorbents at CNES and at NRL could perhaps allow to obtain a better compatibility between the measurements.

V. CALIBRATION OF THE COMPLETE RECEPTION CHAIN

Thanks to the calibration of the antennas and the receivers, it is possible to determine the electrical delay and the uncertainty measurement of the entire reception chains.

Two reception chains will be composed with the calibrated elements :

Chain 1 : Dicom GTR50 receiver + Novatel 702-GG antenna

Chain 2 : Septentrio PolaRx2 receiver + Ashtech Z12-T antenna

The figure 6 presents the delay difference of these reception chains for the L1 frequency code C/A and code P and for the L2 frequency.

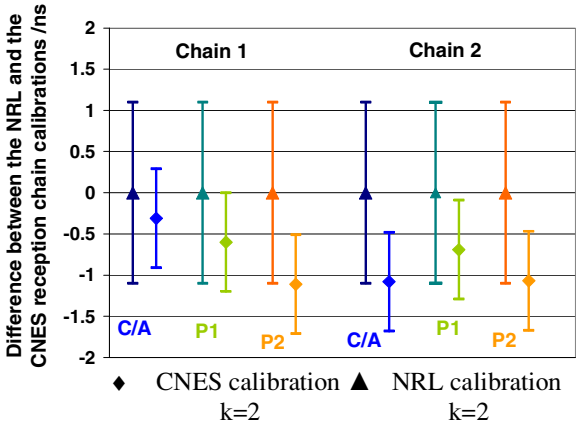


Figure 6. Comparaison between CNES and NRL reception chain calibration

The results of the reception chain calibrations performed at CNES and at NRL are of the same order of magnitude even if the error bars do not completely overlap. The difference between the NRL and CNES reception chain calibrations is inferior to 1 ns even if the measurement environments are not identical. These close results allow to validate the absolute methods put in operations at NRL and CNES and to prove the reproducibility of this technique.

The table IV presents the uncertainty of the reception chains for the L1 frequency code C/A and code P and for the L2. The value for a iono-free utilization is also presented.

Because the ionosphere delay is frequency dependent, one can compute P3, the ionosphere-free linear combination of P1 and P2. Knowing the absolute delay for the P1 and P2 pseudorange codes, the ionosphere-free delay is calculated as follows :

$$RxD_{(P3)} = \frac{f_1^2}{f_1^2 \quad f_2 2^2} RxD_{(P1)} \quad \frac{f_2^2}{f_1^2 \quad f_2^2} RxD_{(P2)} \qquad (6)$$

As P1 = 1.575420 GHz et P2 = 1.22760 GHz, the following approximate equation is used :

$$RxD_{(P3)} = 2.5457 * RxD_{(P1)} - 1.5457 * RxD_{(P2)} \quad (7)$$

The P3 uncertainty is calculated in the same manner :

$$u(P3) = \sqrt{(2.5457 * u(P1))^2 + (1.5457 * u(P2))^2} \quad (8)$$

This formula gives a result three times superior to the P1 or P2 uncertainty but does not take into account any correlations between the L1 and L2 uncertainties.

The global uncertainty presented in the table V is the quadratic sum of the uncertainty of the receiver and the antenna.

TABLE V. UNCERTAINTY COMPARISON OF THE NRL AND CNES ABSOLUTE CALIBRATION OF GNSS RECEPTION CHAINS

Chains	**Calibration method**	**Uncertainty Calibration [ns] (k=1)**	
		C/A/P1/P2	***P3***
Chain 1	NRL	1.03	3.07
	CNES	0.60	1.78
Chain 2	NRL	1.03	3.07
	CNES	0.65	1.93

The uncertainty for the reception chains is about 0.6 ns with the CNES method and 1 ns for the NRL method for the C/A, the P1 or the P2 code. For the iono-free combination, the CNES uncertainty is slightly inferior to 2 ns when it is estimated to 3 ns for the NRL method.

VI. CONCLUSION

It's the first time that a comparison of the absolute calibration results is performed. Two reception chains are used for this comparison : The first is composed of a Dicom GTR50 receiver and a 702-GG Novatel antenna and the second is constituted a Septentrio PolaRx2 receiver and a Ashtech Choke-ring antenna.

The CNES uncertainty of the receiver and antenna calibrations is the same order of magnitude included between 0.39 ns to 0.46 ns. The NRL uncertainty is dominated by the receiver calibration uncertainty estimated to 1 ns when the antenna uncertainty estimation is of 0.2 ns. The uncertainty for the complete reception chain is of 0.6 ns with the CNES method and 1 ns for the NRL method for the P1 or P2 code. For the iono-free combination, the CNES uncertainty is slightly inferior to 2 ns when it is of 3 ns at NRL.

The comparison between the CNES and NRL results shows a good compatibility for the receiver and the antenna calibrations. The results of the NRL and CNES reception chain calibration are close with a maximum difference of 1 ns even if the measurement environment are not similar. This comparison allows therefore to validate the absolute methods put in operation at NRL and at CNES in proving the reproducibility of this technique.

REFERENCES

[1] JP. Defraigne, G. Petit, " Time transfer to TAI using geodetic receivers" *Metrologia, Vol.40, pp 184-188*, 2003

[2] J. Kouba, P. Heroux, "Precise Point Positioning using IGS Orbit and Clock products", *IEEE UFFC, vol. 46 No. 4*, 2002

[3] Petit G., Jiang Z., Uhlrich P., and Taris F., "Differential calibration of Ashtech Z-12T receiver for accurate time comparisons*", Proceedings of 14th EFTF, pp.40-44*, 2000

[4] Plumb J, Larson K., White J. and Powers E. , "Absolute Calibration of a Geodetic Time Transfer System", IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol. 52, no. 11, November 2005

[5] A.Proia, G.Cibiel, "Progress report of CNES activities regarding the absolute calibration method", Proceedings of the 42 th PTTI, 2010

[6] J. White, R. Beard, G. Landis, G. Petit, and E. Powers, "Dual frequency absolute calibration of a geodetic GPS receiver for time transfer", *Proceedings of the 15th EFTF, pp. 167–172*, Mar. 2001

[7] A. de Latour., G. Cibiel, J. Dantepal, J-F. Dutrey, M. Brunet, L. Ries, J-L. Issler, "Dual-frequency absolute calibration of GPS receiver for time transfer", in Proceedings of 19th EFTF, pp.360-365, 2005

[8] A. Proia, G. Cibiel L. Yaigre, "Time stability and electrical delay comparison of dual frequency GPS receiver", Proceedings of 41th PTTI, Santa Ana, 2009

BIPM Calibration Scheme for *UTC* Time Links

- BIPM pilot expermiment to strenghen Asia-Europe very long baselines

Jiang Z., Petit G., Arias F., Lewandowski W. and Tisserand L.

Time Department
Bureau International des Poids et Measures
Pavillon de Breteuil, F-92312, SEVRES CEDEX France
zjiang@bipm.org

***Abstract*— The calibration uncertainty (u_B) is the dominant part in the total uncertainty budget of the *UTC* generation [1,3]. There are two calibration strategies: (a) Receiver calibration and (b) Time link calibration. The state-of-the-art [1,4,5] of the u_B is of order 5 ns for GNSS time transfer based on the strategy (a) and is 1 ns for TWSTFT based on the strategy (b) in which the absolute offset is canceled. In the *BIPM Circular T*, the uncertainty in *UTC-UTC*(k) is obtained from the uncertainties in the links [3]. To reduce the uncertainty in *UTC-UTC(k)*, it is optimal to use the calibration strategy (b). Earlier investigations gave theoretical and practical proofs [6-12]. Based on these studies, we propose the BIPM calibration scheme using a calibration system composed of three GNSS geodesic receivers: one is stationary at BIPM and the other two travel between the laboratories. We describe the set-up of the system, the schedule, the uncertainty u_B and aging of the u_B. The u_B is expected to be better than 2 ns for even very long inter-continental *UTC* baselines.**

I. INTRODUCTION

The recently used PPP technique greatly improves the precise time transfer and is comparable to the impact of the introduction of TW links 10 years ago. Both have produced significant progress and trigged the studies in new methods for the accurate *UTC* time links. One of the important issues is the calibration. There are two basic concepts:

a) *Receiver*-calibration
b) *Link*-calibration

As discussed in [3], time link calibration uncertainty u_B is the dominant part in the uncertainty of *UTC-UTC*(*k*) published in Section 1 of *BIPM Circular T* [1]. The so far TW calibration is based on the (b) concept, i.e., the time link calibration, and that is why the u_B of TW link reaches 1 ns [4]. While the GNSS receiver calibration [5] using the concept (a) still remains at 5 ns. However for the computation of *UTC-UTC*(k), the calibration concept (b) is required. It is potentially easy to reduce the u_B.

Since 2007, studies of the GNSS link calibration have been conducted based on the concept (b). TW links have already been used to calibrate GPS PPP [6] and MC links. The GPS MC link BEV-PTB [7] is the first calibration using the concept (b); at the moment of the calibration (Sep. 2008), its u_B was 3 ns. It was and is still the lowest value reached for the GNSS calibration in the *BIPM Circular T* [1]. A study [8], see also section IV, proved that we can convert the link calibration to the absolute receiver calibration if at least one of the receivers in the network is absolutely calibrated. [8, 9, 12] generalized the link calibration to all the *UTC* time link calibrations and estimated the reachable u_B. [10, 11] proposed a link calibration scheme using a traveling GPS receiver and realized the link calibrations on the baselines ROA-PTB and CH-PTB. The u_B declared is less than 2 ns.

All these studies and exercises gave encouraging results. Based on these studies and towards a standardization of the *UTC* time transfer calibrations, this paper proposes a new BIPM calibration scheme that will be tested in a pilot experiment to strength the Asia-Europe time links. The goal of the pilot experiment is to reach the expected uncertainties in the *UTC* time links:

$u_A \leq 0.3$ ns
$u_B \leq 2$ ns

If this pilot experiment is successful, we will consider using the BIPM scheme to calibrate more *UTC* time links, and in particular, the inter-continental TW and GNSS baselines.

In the following sections, we first define the 'total delay', then the set-up of the 'total delay calibration', the schedule of the calibration tour, named **BIPM scheme,** and finally the uncertainty estimation and conclusion.

II. THE BIPM CALIBRATIN SCHEME

The typical values of time transfer uncertainty are u_A(TW)=0.5 ns, u_A(PPP)=0.3 ns and u_B(TW)=1.0 when using the TW portable station. However the u_B(TW)=1.0 ns has never been realized between Asia-Europe. The present value is 5 ns. The ratio between u_A and u_B is 1:10. Clearly, the major difficulty to improve the quality of Asia-Europe time transfer is the calibration.

When a real TW link calibration is not available, we apply in the *UTC* computation a *link* calibration correction by an alignment operation, e.g., we align a TW link to the corresponding GPS link. This calibration correction is the "total delay of the link" which is the key concept in the BIPM scheme. Therefore the so called new scheme, in term of the

link calibration, has always being practiced in *UTC* computation. The so far BIPM GNSS links are *receiver* calibrated in general while in the new scheme the GNSS links are *link* calibrated.

The following notations are used in the paper:

- **$Dly_{R/i}$**: the total delay of a *receiver* system at Lab(*i*), *i* = BIPM, PTB, Lab(*k*), …
- **$Dly_{L/ij}$**: the total delay of a time *link* system which is the difference of Dly_{Ri} and Dly_{Rj}
- **Ref_i**: the reference receiver at Lab(*i*), *i* = BIPM, PTB, Lab(*k*), …
- **Std_B**: the BIPM traveling standard equipment consisting of N ≥2 GNSS receivers
- **$UTCP_i$**: the *UTC* point at Lab(*i*), *i* = BIPM, PTB, Lab(*k*), … cf. Figures 1 and 4
- **$CLBP_i$**: the calibration point at Lab(*i*). *i* = BIPM, PTB, Lab(*k*), … The $Dly_{R/i}$ is the total delay between this point and the phase center of the antenna at laboratory *i*, cf. Figures 1 and 4
- **Δ**=$CLBP_i$–$UTCP_i$ is the cable delay between the *UTC* point and the calibration point. This is the only measure of the delay that should be made and supplied to BIPM by the Lab(*i*). Δ is considered as known in the following discussion.

III. THE DEFINITION OF THE TOTAL DELAY

The total delay of a *link* is the difference of the total delays of two GNSS receivers at the ends of a baseline. It is defined with the help of the total delay of the *receivers*.

A. The total delay of a GNSS receiver

The total delay of a GNSS receiver (Dly_R) at Lab(*k*) is the total electronic delay between the phase center of the antenna and the $UTCP_k$. It is the sum of all the delays of the cables plus the pieces of equipment, which include the receiver, the splitter(s), the frequency distributor(s), the amplifier(s), the phase micro stepper(s) … and the cables between the antenna and the receiver as well as the cables between all the laboratory equipments, including the **Δ**, on the path of the satellite signal from the antenna to the $UTCP_k$ (Figures 1, 4). If the $CLBP_i$ and $UTCP_i$ are not the same, the delay between them is the Δ which should be measured by the Lab(*k*).

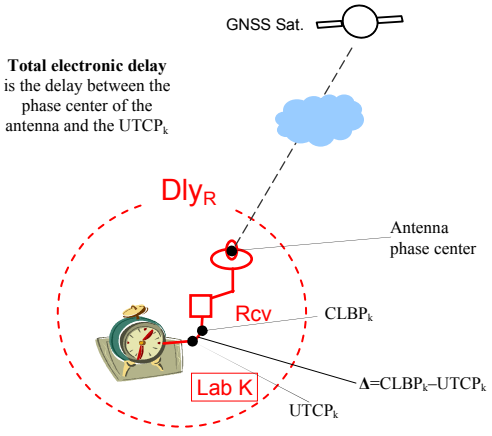


Figure 1. The total delay of a GNSS *receiver* (Dly_R) at Lab(*k*)

In the present BIPM calibrations of C/A receivers, some components, e.g. BIPM antenna cable etc., are measured separately by BIPM and the Lab(*k*). But this practice is not necessary and may introduce errors. Table I shows the discrepancies of the BIPM antenna cable delay and its uncertainties (SD) measured by 18 laboratories. If we take into account only the results accompanied with the uncertainties, the minimum delay is 183.0 ns ±0.5 ns and the maximum is 186.2 ns ±0.01 ns. The difference is 3.2 ns, largely over the given uncertainties. If we also consider those results that are not accompanied by the uncertainty, the difference reaches 6.6 ns. The standard deviation is 1.6 ns. This is only a simple cable, so that it is clear that discrepancies in the delays for other complex parts of the equipment could be worse. All these errors will go to the final error budget of the calibration. It is wiser to avoid all these non necessary measurements by defining and using the *total delay*. However, the laboratories could optionally measure all the components of the all delays. This information would be helpful in case of doubt.

TABLE I. THE DELAYS OF THE BIPM CABLE C101 MEASURED BY 18 DIFFERENT *UTC* LABORATORIES

Cable C101:

Laboratory	Delay /ns	SD /ns	Measurement method	Date (YYMM)
BIPM	184.34	0.4	Double Weight Pulse method	0105
OP	184.6	0.3	Dual weighing method	0112
IEN	185.6	0.1	Pulse method	0202
ROA	186.21	0.01	Sartre Modem	0203
PTB	182.8	0.5	Mitrex Modem	0203
PTB	183.8	0.5	Pulse method	0203
USNO	184.06	0.25	-	0204
NIST	183.0	0.5	Network analyzer	0204
VSL	183.1	-	Mitrex Modem	0206
VSL	183.2	-	1GHz sine + oscilloscope	0206
NTSC	183.75	0.01	TIC method	0208
CRL	184.4	-	TIC method	0208
CRL	182.69	-	Network analyser at GPS freq. L1	0208
CRL	182.71	-	Network analyser at GPS freq. L2	0208
NMIJ	189.6	-	TIC method	0209
TL	183.3	0.5	Network analyzer	0209
NML	183.94	1.0	TIC method	0210
OP	184.3	0.3	Double Weight Pulse method	0211

B. The total delay of a time link

The total delay of a receiver (Dly_R) as defined above is not measurable with only the BIPM scheme. The Dly_R serves in this section only to define the Dly_L, the total delay of a time *link* between Lab(*i*) and Lab(*j*) in the BIPM scheme:

$$Dly_L(i\text{-}j)=Dly_R(i)-Dly_R(j)$$

for a *UTC* link with *j*=PTB, it becomes:

$$Dly_L(k\text{-PTB})=Dly_R(k)-Dly_R(\text{PTB})$$

The total delay of a link Dly_L is the CALR in the case of a TW link. However, the methods to obtain them are quite different. Figure 2 illustrates the configuration of the geometry of the Dly_L which is directly measurable in the BIPM scheme. By default, the total delay in the following sections is always that of the time *link* Dly_L.

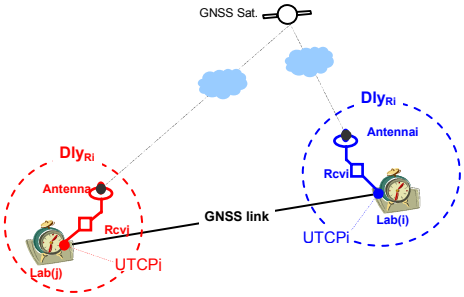


Figure 2. The total delay of the *link* Lab(*i*)-Lab(*j*): Dly_L (i-j)=Dly_R(i)-Dly_R(j)

C. *The BIPM scheme and the computation of a Dly_L*

For performing a *link* calibration with the BIPM scheme, the travelling receiver Std_B does not need to be calibrated and, furthermore, long-term stability of Std_B is not a critical issue. The only requisite is that the Std_B must be stable during the calibration tour. This is in fact a usual requisite for any calibration.

Let us first state that measurement noise (assumed white noise) is not a limiting factor in our calibration scheme, as in any calibration scheme: because hundreds of measurements are taken each day, the measurement noise averages out below 0.1 ns and therefore will not be further considered. On the other hand, the total uncertainty of the calibration depends highly on the instability of the travelling Std_B over the duration of the calibration trip. In order to limit the disturbance of this instability, a multi-receiver system and a rigorous schedule are used in the BIPM scheme. These measures can greatly reduce the impact of jumps in the Std_B reference during the calibration tour because we can know exactly, by the comparison between the traveling and the stationary receivers, the closure measurement at BIPM and the side by side set-up at Lab(k), where andwhen a jump occurred, as well as its size.

Compared with the present BIPM calibration, the new scheme is characterized by:

- the key point is the stability of the BIPM standard which is composed of N+1 GNSS receivers, one of them staying at BIPM (in fact, there are more than one GNSS receivers operational at BIPM and all of them can back up the fixed reference) and a traveling standard composed of N other receivers (N≥2). For the pilot experiment, the BIPM standard consists of three GPS receivers: a reference receiver Ref_B stays at BIPM plus a traveling standard with two GPS receivers: Std_{B1} and Std_{B2} (Figure 5);
- the clock result of a GPS PPP solution is associated with the phase center variation of the antenna which will be determined with a cm level of accuracy by PPP. The laboratories can put the antennas of the Std_B anywhere suitablebut trying to avoid the multi-paths' effect. Therefore, the Std_B ensemble must be of geodetic receivers;
- the common point of the two traveling receivers of the Std_B is a splitter sharing a cable leading to the calibration point (CLBP), cf. Fig. 5. All the calibrations are made with the CLBP as reference;
- the $CLBP_k$ of Lab(*k*) is referred to $UTCP_k$. The Lab(*k*) needs only to measure the delay $\Delta=CLBP_i-UTCP_i$;

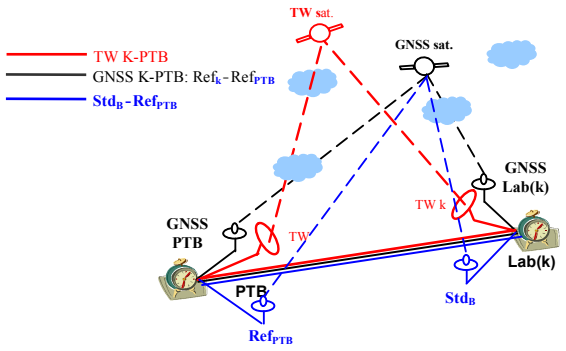


Figure 3. The side-by-side common-clock short-baseline set-up and the total delay Std_B-Ref_k: $Dly_L(Std_B\text{-}Ref_k)=Dly_R(Std_B)\text{-}Dly_R(Ref_k)$

- the calibration result is the total delay of a link Dly_L. We calibrate the total delay of all the time links on the route of the traveling Std_B between PTB and Lab(*k*). The link(s) to calibrate can be of any types: GPS, GLN and TW (Figure 3);
- since the goal is to calibrate links for *UTC*, PTB should always be included. The scheme may be simplified without PTB only if experiences prove the long-term stability of the Ref_{PTB} and the Std_B. But this needs time.

The BIPM scheme has four steps. Fig. 3 is the set-up of step 3.

1) start the measurements at the BIPM;
2) visit the PTB;
3) visit successively Lab(*k*), k=1, 2, 3, …;
4) close the measurements at the BIPM.

IV. Discussion

The basic idea of this section was proposed for the first time in [8]. We generalize the principle to adapt it to the new BIPM scheme. We first prove that the BIPM *link* calibration can be converted to the GNSS *relative receiver* calibration and then that this *relative* calibration can be transferred into the *absolute* calibration. For a time link network, if one of the receivers is *absolutely* calibrated, this *absolute* scale can be transferred to all the other receivers.

We then discuss some characters of the BIPM scheme.

A. *Convert a link calibration to a GNSS* receiver *calibration*

We recall first that the BIPM scheme is not an *absolute receiver* calibration. We are interested in the *link* calibration that affects the *UTC*-*UTC*(*k*) and also in an arbitrary time *link* UTC(*k*)-UTC(*j*). In the following discussion, the "*receiver calibration correction*" stands for the delay correction to be added to the RefGps value (the raw receiver recordings in the CGGTTS data files).

We take an arbitrary *UTC* baseline Lab_k-PTB. Suppose the *receiver* delay correction of the GNSS *receivers* is X_p for PTB and X_k for Lab_k; $Link_k$ is the non calibrated GNSS *link* between Lab_k and PTB; $Link_K$ is the *receiver* calibrated *link*; $Link_B$ is the calibrated *link* by carrying out the BIPM scheme. We can compute the calibration value of the GNSS receiver at Lab_k by two steps: first the *link* correction between Lab_k and PTB and then the *receiver* correction at Lab_k, as follows:

$$Link_K = [RefGps(Lab_k)+X_k] - [RefGps(PTB)+X_p] \quad (4.1)$$

By the definition of the total delay of the link, we have $C_k=X_k-X_p$. Introducing it in equation (4.1):

$$\begin{aligned} Link_K &= [RefGps(Lab_k) + C_k + X_p] - [RefGps(PTB) + X_p] \\ &= RefGps(Lab_k) - RefGps(PTB) + C_k = Link_k + C_k = Link_B \end{aligned} \quad (4.2)$$

Above the C_k can be obtained through the BIPM scheme and X_p can be given through earlier calibration (although we do not need it here). The *receiver* calibration correction, i.e. the total delay of the GNSS receiver at Lab_k can be computed with the simple relation: $X_k = C_k+X_p$. The fact that the *receiver* delay at PTB X_p is cancelled in equation (4.2)

confirms that we do not need it to calibrate any receivers and C_k can be considered as the receiver calibration correction because X_p can be set to any value e.g. X_p=0.

Now we should prove that such obtained *receiver* calibration correction is correct not only for the *UTC* link between Lab_k and PTB but also correct for any link between a pair of *UTC* labs. To be general, we take arbitrarily two labs: Lab_k and Lab_j. From the definition of $Link_k$ and the equation (4.2), we have:

$$C_k = X_k - X_p = Link_B - Link_k \quad (4.3)$$

$$Link_K = Link_k + C_k = [RefGps(Lab_k) + C_k] - RefGps(PTB) \quad (4.4)$$

where C_k is the calibration delay correction applied to the Lab_k *receiver*. As for the *receiver* of PTB, calibrated or not is not important to obtain the calibrated $link_K$ (X_p is canceled). Similar as equation (4.3), the *receiver* delay corrections:

$$C_k = X_k - X_p = Link_B(Lab_k - PTB) - Link_k(Lab_k - PTB) \quad (4.5)$$

$$C_j = X_j - X_p = Link_B(Lab_j - PTB) - Link_k(Lab_j - PTB) \quad (4.6)$$

By subtracting equation (4.6) away fromequation (4.5):

$$C_k - C_j = X_k - X_j = Link_B(Lab_k - Lab_j) - Link_k(Lab_k - Lab_j) \quad (4.7)$$

$$Link_B(Lab_k - Lab_j) = Link_k(Lab_k - Lab_j) + C_k - C_j \quad (4.8)$$

Similar as the equation (4.1), the calibrated GNSS link equals:

$$Link_K(Lab_k - Lab_j) = [RefGps(Lab_k) + C_k] - [RefGps(Lab_j) + C_j] \quad (4.9)$$

In the equation (4.9), C_k and C_j are the GNSS *receiver calibration corrections*. Introducing equations (4.6) and (4.7) into (4.9), we have the GNSS link with the traditional *receiver* calibrations X_k and X_j:

$$Link_K(Lab_k - Lab_j) = [RefGps(Lab_k) + X_k - X_p] - [RefGps(Lab_j) + X_j - X_p]$$
$$= [RefGps(Lab_k) + X_k] - [RefGps(Lab_j) + X_j] \quad (4.10)$$

Again, X_p the *receiver* calibration correction for PTB GNSS receiver is cancelled in the arbitrary link. Equations (4.8) and (4.9 or 4.10) give the same result (within their uncertainties) but with completely different calibration set-ups.

From the equations (4.8) to (4.9) and then (4.10), we transfer the *link* calibration obtained through $Link_B$ (the new BIPM scheme) to the *receiver* calibration (the traditional calibration scheme). The *link* calibration corrections differ from the *receiver* delay corrections by a common constant: X_p. However, it is important to understand that:

1. it is not necessary to know the value of X_p in the BIPM calibration scheme for the *UTC* time transfer;
2. if we know one of the X_k or X_j, we can compute the X_p. It is easy to prove this by exchanging the positions of Lab_k and PTB in equation (4.1), i.e. the X_p of the PTB can be determined at *any* Lab_k

Equation (4.10) is that of the traditional BIPM calibration. As mentioned in the introduction, one of its disadvantages is that the uncertainty is poor: u_B = 5 ns. Equations (4.2) and (4.8) use the new BIPM scheme calibration and its u_B is well reduced, potentially less than 2 ns [10,11]. Furthermore the BIPM scheme allows *simultaneous* multi GNSS receiver calibration between the *UTC* laboratories that operate GPS, GLN and TW facilities. *Simultaneous* measurement is impossible for the traditional calibration (eq. 4.10). Possible biases due to, for example, the long-term variations of the internal reference of Std_B receiver and the geodesic hypothesis etc., are not avoidable.

B. Convert a relative *receiver calibration to* absolute *one*

To identify the difference between the BIPM scheme and the absolute calibration, we name the BIPM-scheme obtained receiver calibration correction the "*relative* receiver calibration". An *absolute* calibration is to determine the sum of the absolute delays of the antenna, the receiver, all the equipment and the cables on the way of the GNSS signal passing through, which is *absolutely* measurable. The direct output of BIPM scheme is defined with respect to the standard receiver Std_B. In the relation (4.2), we assume that in $link_K$ the total delay of the GNSS receiver at PTB is *absolutely* calibrated. The absolute delay is $X_{p/A}$. Let us set, $X_{p/A} = X_p + X_o$, X_o here is the difference between the *receiver* calibration and the *absolute* calibration. Therefore, we have at the Lab_k

$$X_k = C_k + X_p = C_k + X_{p/A} - X_o, \quad (4.11)$$

$$X_{k/A} = X_k + X_o = C_k + X_{p/A} \quad (4.12)$$

Therefore, to convert the BIPM scheme total delay C_k to the *total absolute delay* $X_{k/A}$, it is as easy as to add a constant: $X_{p/A}$. Again for the *UTC* generation, we compute the *UTC*-*UTC*(k) and do not need the *absolute* delay. However, the *total absolute delay* is indispensable for the Section 5 where the differences of the *UTC* and GNSS system reference time are given.

It is interesting to point out that:

1) in the *UTC* network, if any of the receivers is *absolutely* calibrated (not necessarily that of PTB), its *absolute* scale can be transferred to all the other receivers;

2) if there are *more than one* receiver *absolutely* calibrated in the BIPM scheme, the redundant calibrations can improve the calibration by a least square adjustment, cf. [17] for details.

C. The BIPM scheme is a unified one-for-all calibration

By Figure 3, the set-up of the BIPM scheme at Lab(*k*) serves to calibrate all kinds of *UTC* time *links* between the Lab(*k*) and PTB: TW and all type of GNSS (C/A, P3 or PPP). This is the so-called unification of all the *UTC* time links by the unique BIPM standard [12].

Advantages of the unified approach are:

1. the calibration results will be affected only by the short-term instability of the Std_B;
2. calibration of the TW links, in particular of the very long inter-continental TW links, is possible;

3. improvement of the calibration uncertainty u_B for most *UTC* time links, and consequently reduction of the uncertainty of [*UTC* – *UTC*(*k*)] in the *Circular T* from 5 to 2 ns, see below V;
4. simplified organization of *UTC* calibration campaigns;
5. simplified monitoring of these calibrations.

D. The BIPM scheme is more accurate and less manpower, time and money consuming

In the traditional BIPM differential calibrations of GNSS equipment, different BIPM standard receivers travel for different GNSS receivers calibrations: GPS C/A code receivers, GPS P3 code receivers and GLONASS receivers. TW calibration is separately. Many of them have never been achieved. During the calibrations, the visited laboratories have to determine the position of the BIPM antennas and to measure all the delays of cables and equipment on the path of the satellite signal. These increase the total uncertainty and the labors.

The new BIPM scheme allows the calibration of all types of links on a baseline; the laboratories measure and supply only the delay Δ between UTCP(*k*) and CLBP(*k*). This saves time, manpower and reduces the measurement uncertainties.

Generally speaking, if the uncertainty of a *receiver* calibration is *u*, the total uncertainty in *UTC*-*UTC*(*k*) is $u_B \approx \sqrt{2} \times u$. However, for a link calibration, we have $u_B \leq u$.

E. Comparison between the new and the old BIPM schemes

We summarize the features of the new BIPM scheme vs. the traditional ones.

- The new scheme calibrates links while current schemes calibrate receiving equipment. Through link calibration, all existing techniques between a pair of laboratories can be calibrated at the same time.
- Two receiving systems travel in the new scheme, providing redundancy and cross-checking. As the equipment is organized in a unique container and as the visited lab should only provide a frequency and a 1PPS and only measures the delay of the 1PPS, the operations are simpler and the risk of mistake is minimized.
- The receivers provide phase and code measurements allowing the precise positioning of antennas with PPP. This is not the case in the present C/A calibrations.
- Finally the link calibration uncertainty u_B is expected to be 2 ns or below (see next section) while, in the present scheme, the combination of the uncertainties of the two receiver calibrations provides a link uncertainty u_B = 5 ns.

V. THE UNCERTAINTY

The proposal in this section was originally raised in [12]. We generalize it to the new BIPM calibration scheme.

*A. The u_B of a UTC link Lab(*k*)-PTB*

The calibration uncertainty for the link Lab(*k*)-PTB can be estimated by:

$$u_B^2(k\text{-PTB}) = [u_A^2(k\text{-PTB}) + u_A^2(\text{Std}_B) + \boldsymbol{c}^2(\text{Std}_B)]/N \qquad (5.1)$$

N is the number of BIPM traveling receivers; $\boldsymbol{c}$ is the conventional short-term-link-instability of the Std_B for a typical calibration tour that consists of 4 occupations of 4 steps: BIPM➔PTB➔Lab(*k*)➔BIPM. If we set $\boldsymbol{c}$ = 2 ns (as is shown to be achievable by Esteban et al [10]) and N=1, the u_B of the calibrations are: u_B(GNSS P) ≈ 2.23 ns (for u_A = 0.7); u_B(GNSS C/A) ≈ 2.60 ns (for u_A = 2.0) and u_B(TW) ≈ 2.17 ns (for u_A = 0.5). The above estimations are conservative as in fact N=2 so that the value of u_B obtained by equation (5.1) may be reduced, e.g. to about 1.6 ns for the GNSS P3 code calibration [17].

B. Aging of the u_B

The above estimation is valid only at the moment of the calibration without taking into account of the evolution of u_B with time. It is logic that without re-calibration or an independent check, the maximum life of a "*free running*" u_B is limited, saying about 10 years, e.g., for each of the techniques u_B(>10 year) = 10 ns. A "free running" u_B increases with time. For the purpose of automatic computation, it is useful to design an analytical model for u_B evolution. For example, a model based on a random walk phase noise or on a flicker phase noise (or on a combination of both) could be used. In the following, examples are shown for a random walk model, although the resulting values are probably overly pessimistic for the long-term:

$$u_B^2(T - T_0)_{\text{k-PTB}} = u_B^2(T_0)_{\text{k-PTB}} + (T - T_0) \times u_{A^2\text{k-PTB}} \qquad (5.2)$$

Here T-T_0 is the time since the calibration in months and u_A is the random walk part of the 1-month Type A uncertainty of the link taken to be: u_A(GNSS C/A code) ≈ 0.7 ns; u_A(GNSS P codes) ≈ 0.5 ns and u_A(TW) ≈ 0.3 ns. The values of u_B for different calibration ages calculated on the basis of the above hypothesis are presented in Table III. The unit is ns.

TABLE II. EVOLUTION OF U_B INCREASING WITH TIME AS A RANDOM WALK PROCESS

Link	$u_B(T_0)$ /ns	u_A /ns	u_B(1yr) /ns	u_B(3yr) /ns	u_B(5yr) /ns	u_B(10yr) /ns	u_B(>10yr) /ns
TW	1.0	0.3	1.4	2.1	2.5	3.4	10
GNSS P	2.0	0.5	2.6	3.6	4.4	5.8	10
GNSS C/A	3.0	0.7	3.9	5.2	6.2	8.2	10

VI. THE EXPERIMENT SET-UP

Figure 4 illustrates the general set-up up of the BIPM calibration scheme at the Lab(*k*). Here the 'time transfer equipment' represents one or all the time transfer techniques: TW, GLN and GPS C/A, P3 and PPP. The configuration between the pulse distributor driven by the master clock and the receivers of Lab(*k*) as well as the BIPM traveling standard is demonstrated. The BIPM traveling Std_B is installed in the BIPM calibration station (Figure 5), composed of two GPS geodesic receivers driven by the same master clock as the receivers at Lab(*k*). The pulse passes first through the CLBP_k of the distributor then through the BIPM receiver cables goes to the BIPM splitter in the station where it is separated into two signals going to the two BIPM individual standard receivers, Std(1) and Std(2). The delay of the ensemble of the BIPM Std from the antennas to the CLBP_k is same at all the

Labs and will be canceled in the time links. The difference of the links of the Std_B-PTB and that of the Lab(*k*)-PTB is nothing but the *link* calibration correction.

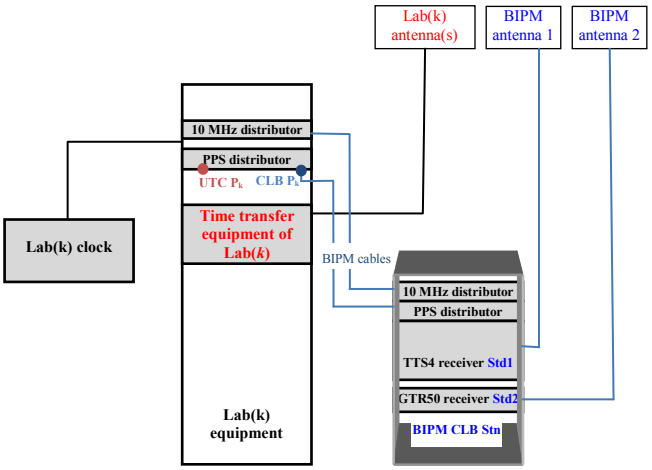


Figure 4. Configuration between the pulse distributor linked with the master clock and the receivers of $Ref_{Lab(k)}$ and Std_B

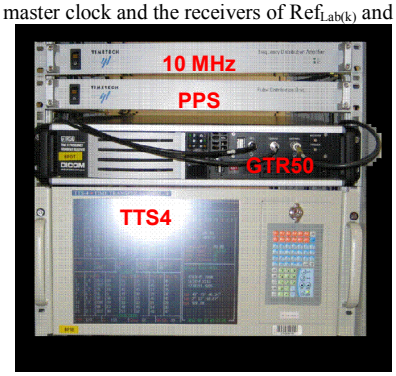


Figure 5. *The BIPM calibration station Std_B*

VII. SUMMARY

We have studied a new calibration approach based on the *link* calibration concept, and named it "the BIPM calibration scheme". We developed a set of calibration equations and computer programs that have been installed in the *UTC/TAI* software package Tsoft. Preparation of the equipment and organization are ongoing. A pilot study has been lunched during the CCTF TW Working Group meeting in Sept. 2010 in Beijing and will plan to perform the first measurements in laboratories in 2011.

TW and GPS PPP are the best performing techniques in *UTC* time transfer. Each has advantages and is indispensable. However, since last year, two phenomena arose: the quality of TW is degrading and that of the PPP has improved. Replacing the degraded TW links by PPP does not seem a good solution; we propose therefore 1) combination of the two techniques to improve the u_A; 2) differential of the two techniques to improve the u_B. In particular, we aim at strengthening the Asia-Europe very long baselines for *UTC* time transfers. This is the goal of the pilot experiment. To do it, we need:

- Combination TW and PPP: TW+PPP
- Application of the TW indirect links composed of two direct ones
- PPP assisted calibration: the BIPM scheme of calibration TW-PPP
- The expected uncertainties: $u_B \leq 2$ ns and $u_A \leq 0.3$ ns

The methods applied in this paper could be extended to the whole *UTC* time transfer worldwide network through coordination with the Regional Metrology Organizations for supporting. A significant improvement of the global uncertainty is expected.

REFERENCES

[1] BIPM Circular T 267, ftp://ftp2.bipm.org/pub/tai/publication/cirt.267

[2] Jiang Z, Petit G, Combination of TWSTFT and GNSS for accurate UTC time transfer, Metrologia 46 (2009) 305-314

[3] W. Lewandowski, Matsakis D, Panfilo G, Travella P, The evaluation of uncertainty in [UTC-UTC(k)], Metrologia 43, pp 278 – 286, 2006

[4] D. Piester, A. Bauch, L. Breakiron, D. Matsakis, B. Blanzano, O. Koudelka, Time transfer with nanosecond accuracy for the realization of International Atomic Time, Metrologia, Vol. 45, No. 2. pp. 185-192, 2008

[5] G. Petit, Z. Jiang, P. Moussay, J. White, E. Powers, G. Dudle, P. Uhrich, "Progresses in the calibration of "geodetic like" GPS receivers for accurate time comparisons," Proc. EFTF 2001

[6] Jiang Z, Transfer the UTC Calibration from TWSTFT to GPS PPP, BIPM TM 151 May 2008

[7] A. Niessner, W. Mache, B. Blanzano, O. Koudelka, J. Becker, D. Piester, Z. Jiang, and F. Arias, Calibration of the BEV GPS receiver using TWSTFT, Proc. PTTI 2008

[8] Z. Jiang et N. A. Niessner, Calibrating GPS with TWSTFT for Accurate Time Transfer, Proc. PTTI 2008

[9] E.F. Arias and Z. Jiang, Considerations on unifying the TWSTFT and GPS Calibration for UTC time transfer, presented at the CCTF WG on TWSTFT Workshop 2008

[10] H. Esteban, J. Palacio, F. J. Galindo, T. Feldmann, A. Bauch, D. Piester, A GPS calibration trip expeerience between ROA and PTB, Proc. EFTF 2009

[11] Feldmann T., Bauch A., Piester D., Stefanov A.; Bernier L.G., Schlunegger C., Liang K., On Improved GPS-Based Calibration of the Time Links between METAS and PTB, Proc. EFTF 2010

[12] Jiang Z, Arias F, Lewandowski W, Petit G, Toward new procedures in TWSTFT and GNSS delay characterization for UTC time transfer, Proc. EFTF2010

[13] Matsakis D. (USNO): personal comunication, email to Jiang Z, 18 Nov. 2010

[14] Tseng W., Y. Lin (TL): personal comunication, emails to Jiang Z, Sept. 2010

[15] Wen-Hung Tseng, Shinn-Yan Lin, and Kai-Ming Feng, "Analysis of the Asia-Pacific TWSTFT network," Proc. of the 2008 IEEE International Frequency Control Symposium (IFCS) and Exposition, Honolulu, HI, USA, pp. 487-492, 2008

[16] Wen-Hung Tseng, Shinn-Yan Lin, Kai-Ming Feng, Miho Fujieda, Hideo Maeno, "Improving TWSTFT Short-Term Stability by Network Time Transfer, "IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, no. 1, pp. 161-167, Jan, 2010

[17] Z. Jiang, G. Petit, F. Arias, W. Lewandowski and L. Tisserand, 2010, Improvement of Very Long UTC Time Links and the BIPM Calibration Scheme, BIPM TM182

Coordinating GPS Calibrations Among NIST, NRL, USNO, PTB, and OP

Marc Weiss, Victor Zhang
Time and Frequency Division, NIST
Boulder, Colorado, USA
mweiss@nist.gov

J. White [2], K. Senior [2], D. Matsakis [3], S. Mitchell [3], P. Uhrich [4], D. Valat [4], W. Lewandowski [5], G. Petit [5], A. Bauch [6], T. Feldman [6], A. Proia [5,7]

[2] U.S. Naval Research Laboratory, Washington D.C., USA
[3] U.S. Naval Observatory, Washington D.C., USA
[4] Observatoire de Paris, Paris, France
[5] Bureau International des Poids et Mesures, Sèvres
[6] Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
[7] Centre National d'Etudes Spatiales, Toulouse, France

***Abstract*—Reviewing calibration results over the history since early 1980's among several labs shows very mixed results. The best stabilities of GPS receivers, as given by calibrations, are of the order of a few nanoseconds or better over a year, though many results are quite a bit worse. Absolute calibrations show similar potential, though there are problems. We conclude that more calibrations and standard methods are needed.**

I. INTRODUCTION

The purpose of this paper is to evaluate the status of GPS receiver delay calibrations, for a subset of receivers that support the generation of international atomic time (TAI). We first discuss historical and current methods for calibrating receivers. Then we look at the history of differential receiver calibrations in order to estimate receiver stabilities for the receivers in use at the time. Our last consideration of differential calibrations is to compute closures of calibration values by use of the most recent available data. These data give a measure of consistency of differential calibrations for use in current receivers. Finally, we discuss the history, use and potential for absolute calibrations.

GPS data are used in various ways for comparing clocks for generating TAI. All of these comparison methods are differential. We use the data to determine the difference between UTC(lab) among the labs that contribute to TAI. The delays through receivers need to be calibrated differentially in order to reference the data to a laboratory's UTC measurement plane. The best way to do this is with the use of a portable GPS receiver as a calibration system. The method is to do a series of common-clock, short-baseline calibrations between a travelling receiver and the primary receivers of different labs.

For example, to calibrate the differential delay between laboratory A (lab A in the following) and laboratory B (lab B), we could start with a portable system at lab A. To achieve a common-clock short-baseline calibration, one connects the travelling receiver to the same clock signals as the primary receiver, measuring the delay in the reference signal from the lab's UTC reference plane. We will discuss the details of a common-clock short-baseline calibration later. In principle, one collects data from the same satellites, measured against the same clock in two receivers with separate antennas, positioned near each other. With proper calibration of reference signals and coordinates, the difference in the measurements should be due only to the difference in the receiver systems, mainly the delays in the antennas, antenna cables and receivers. Evaluating these data allows us to determine the difference in the delays between each pair of systems. The sequence, in practice, would be to first measure the travelling receiver against the primary system at lab A, which we denote Trav-SysA (1). The travelling system is then brought to lab B, where we measure Trav-SysB. We then bring the travelling system back to lab A, where we determine a closure measurement of Trav-SysA (2). This closure gives a measure of any systematic shift in the delay of Trav during the trip. Typically, we then conclude that the difference in primary receivers is:

SysB – SysA= (Trav-SysA(1) + Trav-SysA(2))/2 - (Travs-SysB).

The actual value of SysB-SysA must be accompanied by the uncertainty of this measurement. We will discuss methods used for determining this in later sections.

There are many laboratories that contribute GPS data in support of generating TAI. There are also several different types of receivers, several different formats for the data, and several different methods for processing these data. In all cases, the differential receiver delays are crucial for

determining the values for UTC(lab A) – UTC(lab B). These delay values cannot be determined remotely. Either differential measurements must be made *in situ* against another receiver, or an absolute calibration of the receiver system must be done, which requires interrupting service. Differential calibrations appear to be the most accurate and convenient method of measuring these relative delays, though an absolute calibration is required for the total delay of one receiver system. Campaigns of calibrations have been done historically, and seem to be required at ongoing intervals to maintain accurate time transfer. We discuss in this paper the status of current differential and absolute calibrations, various hardware and software methods for performing calibrations, and possibilities for the way forward to best use GPS receivers and data in support of TAI.

II. Measurement Techniques for Calibration

GPS receivers for timing from 1980 on used a 1 pps signal as a reference. GPS time was measured against this 1 pps usually by use of a time-interval counter. Calibration involved measuring this reference 1 pps against the reference plane for time of the lab [1]. With the advent of receivers that used a 5 MHz or 10 MHz frequency signal as the reference signal, calibration became more complicated. Receivers measure GPS signals against a zero-crossing of the 5/10 MHz. This zero-crossing is determined by a 1 pps signal. So, in addition to knowing the offset of the 1 pps from the reference plane, the so-called tic-to-phase needs to be measured. Complicating this, some receivers lock an internal oscillator to the external frequency, and use that for a reference. Determining the internal reference point may be somewhat ambiguous. The internal reference point may be more than the next internal zero-crossing from the reference 1pps, or indeed may precede it, since the processing is usually digital. In addition, there are tic-to phase zones where the reference point may slip forward or backward a cycle. These forbidden zones may not be near to the value when the measured external tic-to-phase is close to zero. Figure 1 gives a block diagram illustration of the aspects involved with calibrating a frequency-reference receiver.

Some receivers can have a 1 pps out that is derived directly from the internal frequency reference with a fixed delay from the lock point. In that case, it may be possible to use that signal to eliminate any ambiguity for the reference zero-crossing in the receiver.

Thus we see that there are several measurement techniques possible for calibration. In addition, the question of biases in time-interval counters has been considered [2]. One recommendation has been to use a travelling time-interval counter along with the travelling receiver to minimize uncertainties due to counter error.

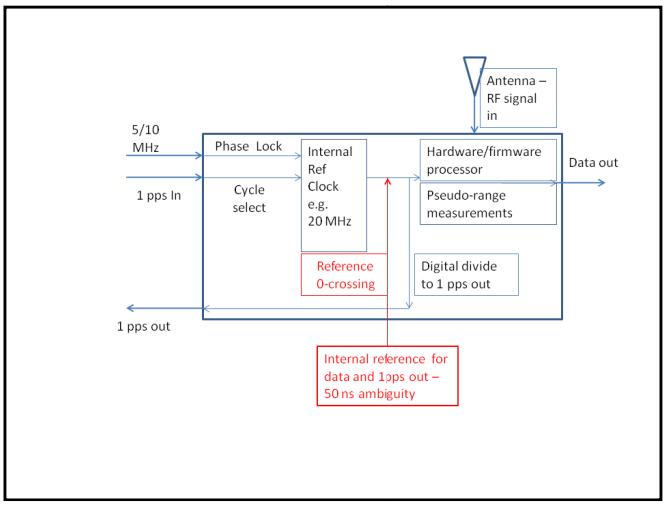


Figure 1. Block diagram of the timing functions in a modern GPS or GNSS receiver that measures the received pseudo-range against an internal oscillator, which in turn is locked to an external frequency reference. This illustrates the difficulty in unambiguously finding the the reference point.

III. Processing Calibration Data

Historically, there have been two different frequencies that have delays through the receiver, L1 and L2, with C/A and P codes modulated on these frequencies, P code not being originally intended for civilian use. While these are the frequencies of most interest to the timing community historically, a new frequency specifically designed for civilian use, L5, is appearing in the new GPS constellation, as well as a version of the C/A code on L2 called L2C. There will also be other constellations of interest. The Russian system, GLONASS, has been available for some time. Though less used for various reasons, GLONASS may become more used for TAI generation as it is modernized. There are also the European system, Galileo, and the Chinese system, Compass or Beidou 2, which are being developed.

There are several formats that data can be taken in and several methods for processing these data. The CGGTTS format [3] provides 13 minute averages every 16 minutes, with all delays from the satellite removed and the times of the on-board satellite clock adjusted for GPS time. For a short-baseline common-clock comparison, as a minimum, pseudo-range data must be adjusted for the range delay, the delay of the signal from the GPS satellites to the receivers. If exact common-view computations are done this is enough. However, the CGGTTS format uses only the C/A code for the main data, the Reference-GPS data. If there are P code on L2 carrier (P2) data, measured ionosphere values are a deterministic combination of both P2 and C/A code data. It is possible to use the combination of Reference-GPS and measured ionosphere data types to compute the differential delays for both the C/A code and the P2 code.

The format called RINEX (for Receiver INdependent EXchange format) [4] has been developed to provide detailed data for flexible processing options. RINEX data are instantaneous pseudo-range data, and are typically recorded every second to every 30 seconds. This format can store data

for all codes from all constellations. It is managed and updated by the International GNSS Service (IGS), and version 3 should provide flexible capability. However, it requires more storage space and processing than the CGGTTS data, which are already pre-processed. Minimal processing for common-clock short-baseline calibrations would be software that corrects only for the range delay from the satellites to the receivers, then differences these data between receivers matching exact common-view data, i.e. exactly matching satellites and transmission times between receiver data. This processing gives the differential calibration for the maximal density of data, with the minimum of extraneous processing.

In the timing community, it is most common to use either of two software systems for processing RINEX data. The P3 software reads RINEX files and outputs a CGGTTS format using either the C/A or P codes on L1 (C1 or P1) and the C/A or P2 codes on L2 (C2 or P2). The P3 data type is the ionosphere-free combination of P1 or C1 and P2 (C2 is available on a very limited number of satellites at this time). This format can, in principal, be used for differencing any of these codes individually, with 13 minute averages every 16 minutes. In practice, the usual output from the software allows for using the P3 combination and the value of the measured ionosphere to derive calibration values. Alternatively, one can use Precise-Point Positioning (PPP) [5,6] software to estimate the value of local clock against IGS system time at both locations for every reference time in the RINEX files. This uses carrier-phase smoothing of the code measurements, fixed local coordinates, and IGS post-processed estimates of satellite orbits to give an optimal estimate of the receiver's reference clock. While the PPP software produces perhaps the most precise estimates of user clock against system time, using it to compute differences for a local calibration introduces terms that do not cancel between receivers, due to the smoothing and averaging techniques.

Thus there are various processing methods for computing calibrations, with trade-offs for convenience, complexity, and accuracy.

A. OP-NIST

Table I and Figure 2 give the values for differential calibrations between OP and NIST since 1983 [1]. The values denoted d are differential time corrections to be added to [UTC(NIST)-UTC(OP)], and u(d) are estimated uncertainties for the periods of comparisons. All calibrations are of NBS10 at NIST compared to NBS51 at OP except for the November 2009 calibration. The NBS10 vs. OP NBS51 calibrations were all C/A code only comparisons. The November 2009 calibration compared the new primary receiver at NIST, NISTn1, with TTR01 and OPMT at OP. NISTn1 is a Novatel OEM4-G2* used at NIST for providing both CGGTTS and RINEX format data to the BIPM and IGS. This provides the backup link to the NIST TWSTFT link. OPMT, an Ashtech Z12-T*, provides the RINEX files to the BIPM that are used for the NRCan PPP computation or for the CGGTTS TAIP3 files. There is no endorsement or critique either implied or intended. This is the source for the backup link to TWSTFT for the link OP-PTB in the TAI network. The NISTn1 vs. OPMT comparison included codes on L1 and L2, yielding a P3 comparison. The best stabilities seem to be on the order of a few ns or better over a year. There are some larger variations that are unexplained, and an apparent walk of almost 10 nanoseconds over 20 years.

* Note that we include product names and model numbers only for reference. No endorsement or critique is implied.

IV. Calibration History

A. OP-PTB

We present first, in Table II and Figure 3, the calibrations from the BIPM website [7] for C/A receiver calibrations between OP and PTB.

TABLE I. Some past calibrations between NIST and OP

OP Receiver	NBS51-NBS10		TTR01-NISTn1		OPMT-NISTn1	
Date	d/ns	u(d)/ns	d/ns	u(d)/ns	d/ns	u(d)/ns
July 1983	0.0	2.0				
September 1986	0.7	2.0				
October 1986	-1.4	2.0				
January 1988	-3.8	3.0				
April 1988	0.6	3.0				
March 1995	-3.7	1.0				
May 1996	-3.0	1.5				
May 2002	-5.0	3.0				
July 2003	-5.6	1.9				
December 2003	-4.6	3.0				
December 2005	-8.7	3.0				
November 2009			-9.4	0.8	-7.3	0.8

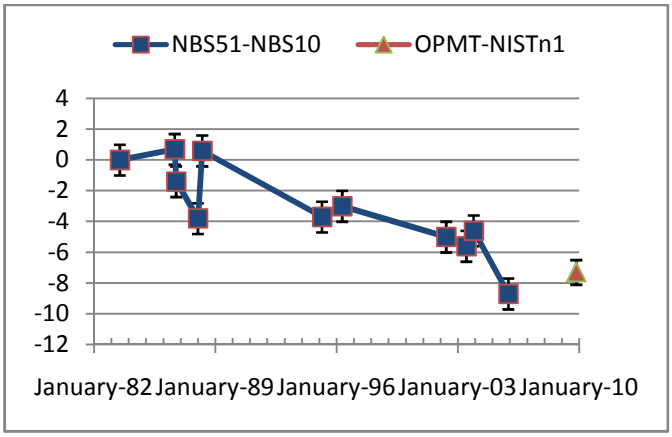


Figure 2. NIST-OP calibration values from Table 1.

TABLE II. DIFFERENTIAL TIME CORRECTIONS D TO BE ADDED TO [UTC(PTB) – UTC(OP)], AND U(D) ARE ESTIMATED UNCERTAINTIES FOR THE PERIODS OF COMPARISONS.

Date	d / ns	u(d) / ns	
October 1986	9.4	2.0	
October 1994	4.0	2.0	
July 1997	2.0	3.0	
November 1997	4.0	2.0	
March 1998	–6.0	2.0	
June 1998	5.0	3.0	
March 2002	–1.0	3.0	
June 2003	–4.7	3.0	
August 2003	0.1	3.0	
June 2004	0.1	3.0	
July 2004	0.6	3.0	TTR5 – TTS2
September 2006	–4.2	3.0	TTR5,TTR6 – TTS3
November 2009	10.3	3.0	
November 2009	4.8	3.0	

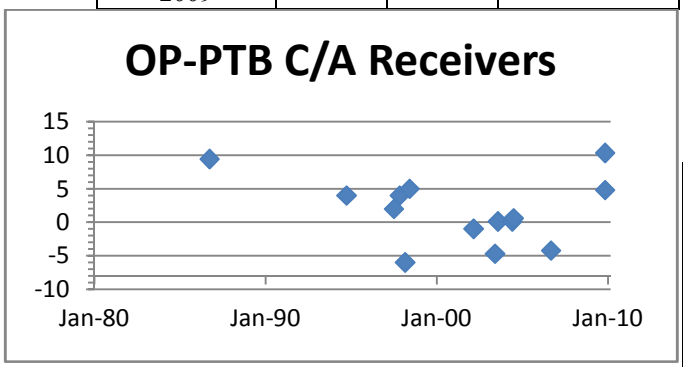


Figure 3. Plot of the data in Table II for the link OP-PTB. The receivers at OP were an NBS TTR5, then an AOA TTR6. At PTB the Code receivers were an NBS TTR5, then an AOS TTS2, followed by an AOS TTS3.

Apparent variations in calibration repeatabilities may be due to using different receivers. The best stabilities seem to be on the order of a few ns or better over a year. Next, we consider the OP-PTB GPS TAIP3 calibration link. Table III, and Figure 4 provide the P3 OP-PTB link calibration results for a direct comparison with other GPS C/A calibration results. All calibration data were computed by the BIPM during calibration campaigns, except the last line, which was released in the frame of the Galileo Fidelity activities from OP and CNES calibration campaigns [8]. Since receiver calibration values are implemented in respective receivers here, contrary to the previous C/A receiver calibrations, it is appropriate to only compare the change in values between these results and other tables. For some periods we see very little change between calibrations. For others there are 5-6 ns over a year or less, though the uncertainties are comparable in magnitude.

Finally in this section, we study the uncertainty budget consistency between GPS and TWSTFT on the OP-PTB link. Here we present an assessment of the uncertainty budget consistency between GPS and TWSTFT on the OP – PTB link. We have built daily averaged differences between UTC(OP) – UTC(PTB) as obtained from TAIP3 GPS CV and by TWSTFT. Figure 5 shows the results over two years, from February 2009 to February 2011, some outliers having been averaged out. The vertical lines materialize some events having potentially affected the measurements, like reference clock changes or a satellite transponder change. The horizontal lines are the plus or minus one sigma limits around 0, obtained from the quadratic sum of the uncertainties claimed for both techniques on the OP-PTB link. Over the whole period, the GPS TAIP3 combined uncertainty was estimated at 3.3 ns (1 σ). The TWSTFT combined uncertainty at 1 σ was estimated to be 1.1 ns until April 2009, 1.2 ns from May to November 2010, and 1.3 ns afterwards. The resulting uncertainty on the difference between both techniques is about 3.5 ns (1 σ). Over the last two years, the results obtained on the OP-PTB link are consistent with the statistics of the respective uncertainty budgets, despite major events like either reference clock changes for both techniques or a satellite change for TWSTFT, which require each time local measurements of new delays. There remains an average bias of about 2.0 ns over the period, which is probably due to some delays not properly taken into account. This calls for a better understanding and new measurements of both equipment chains.

TABLE III. . DIFFERENTIAL TIME CORRECTION D TO BE ADDED TO [UTC(PTB) – UTC(OP)] WHEN MEASURED BY GPS TAIP3 COMMON-VIEWS, U(D) BEING THE ESTIMATED TOAL UNCERTAINTIES

Date	d / ns	u(d) / ns
2002-07	10.8	5.0
2003-05	10.8	5.0
2004-08	15.0	5.0
2006-04	13.3	5.0
2008-04	13.7	5.0
2008-11	7.7	3.2

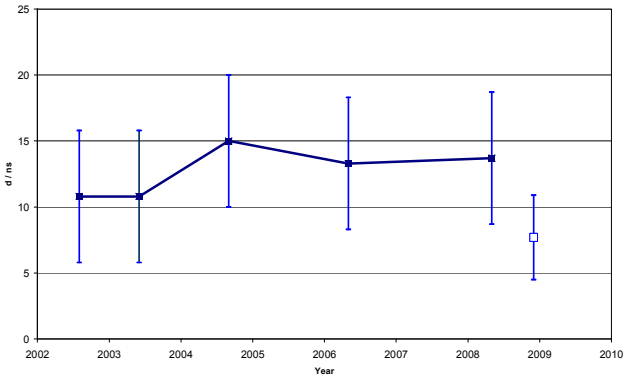


Figure 4. Plot of the data in Table III for the link UTC(PTB) – UTC(OP) by GPS TAIP3. The OP GPS receiver is OPMT. The PTB receiver is PTBB. The last point is kept separated from the others as it is the only one not measured by the BIPM.

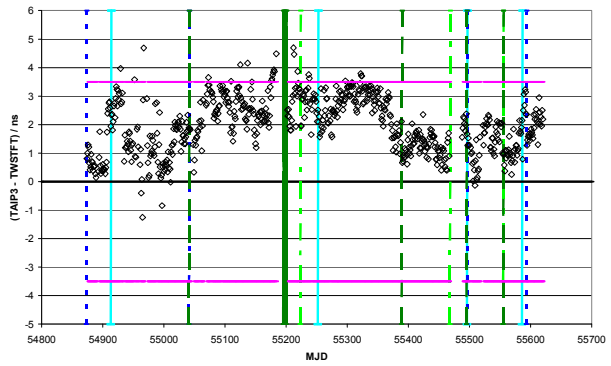


Figure 5. Daily averaged differences between UTC(PTB) - UTC(OP) by GPS TAIP3 CV and UTC(PTB) - UTC(OP) by TWSTFT.

TABLE IV. USNO – NIST CALIBRATIONS

Date	d/ns USNO C/A receivers - NBS10	u(d)/ns	d/ns USNO P3 receiver - NISTn1	u(d)/ns
Sep-86	25.3	2		
Sep-94	-9	3		
Dec-94	-7.6	1		
May-07	-7.5	2		
Aug-10			-6.2	2
Apr-11			-0.5	1

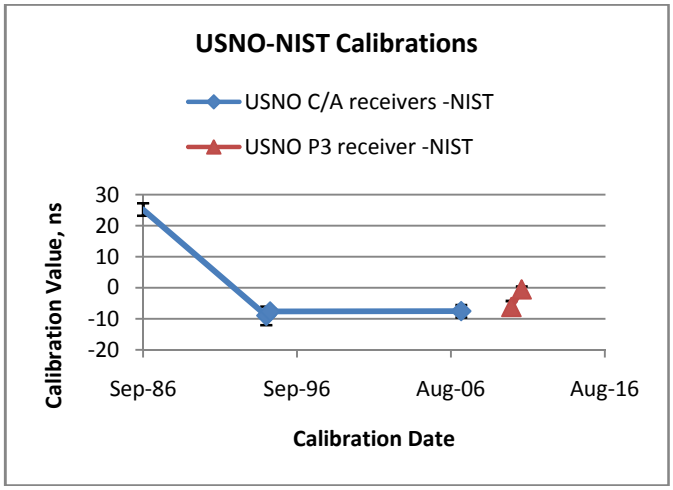


Figure 6. USNO-NIST calibration values from Table III.

B. *USNO-NIST*

Calibrations between USNO and NIST are presented in Table IV and Figure 6. These data come from the BIPM website files [7], except for the most recent calibrations of August 2010 and April 2011. These latter two use the two-frequency receivers at both labs, and are not yet finalized.

C. *USNO-OP*

Data for calibrations between USNO and OP from the BIPM website are presented in Table V and Figure 7. Given that a 14 ns step was introduced in 1997, the data are separated into calibration results before and after this event. The most recent results, from 2003 to 2007, show variations at about 2 ns per year, except for the large variation from 9 ns in 2002 to 2.4 ns in 2003. Given the good values over several years, this suggests the possibility of missing information regarding the large change over the earlier periods.

TABLE V. USNO-OP CALIBRATION VALUES FROM THE BIPM WEBSITE. A 14 NS STEP WAS INTRODUCED IN FEBRUARY OF 1997.

Date	d/ns before +14 ns	u(d)/ns	d/ns after +14 ns	u(d)/ns
Dec-84	32	10		
Oct-86	25.3	2		
Apr-87	15.6	5		
Jun-91	-14			
Jun-94	-13	2		
Sep-94	-9	1		
Dec-94	-7.6	1		
Mar-95	-20	2		
Jan-96	-14	2		
Apr-02			9	3
Dec-03			2.4	3
Jan-06			-4.5	3
Mar-07			-2.3	3

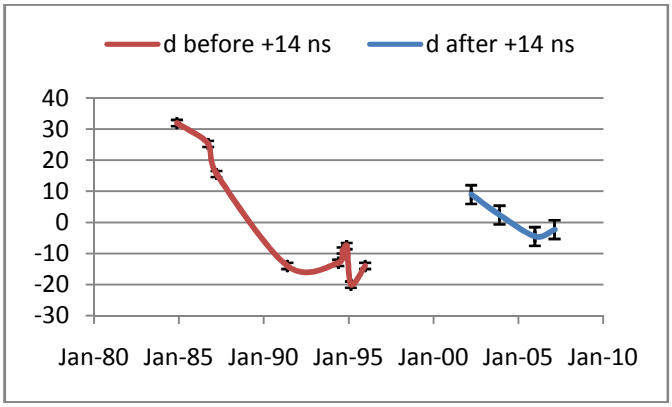


Figure 7. USNO-OP calibration data from Table V.

V. CLOSURE CALCULATIONS

We take some of the most recent calibrations and compute the total around closed loops. Clearly, the results should be 0. Note that the symbol +/- or σ in the Table VI indicates the total uncertainty for the given calibration.

C/A code, L1 frequency receivers:

OP-NIST (TTR6-OEM4)=	9.4 +/- 0.8, 2009/11
PTB-OP (TTR5,6-TTS3)=	-4.8 +/- 3.0, 2009/11
USNO MOT1-PT05=	2.3 +/- 1.2, 2010/06
NIST-USNO MOT3=	-7.5 +/- 2.0, 2007/05
Sum=	-0.6 +/- 3.9 ns

C/A code, L1 frequency receivers:

OP-NIST (TTR6-OEM4)=	9.4 +/- 0.8, 2009/11
USNO-OP=	-2.3 +/- 3.0, 2007/03
NIST-USNO MOT3=	-7.5 +/- 2.0, 2007/05
Sum=	+0.4 +/- 3.7ns

TABLE VI. CLOSURE FOR 2-FREQUENCY RECEIVERS

Labs	C1/P1 mean	C1/P1 σ	P2 mean	P2 σ	P3 mean	P3 σ	Year/ month
OPMT-NIST	-7.3	0.8	-7.2	0.9	-7.5	2.5	2009/11
PTBB-OPMT	3.3	1.0	0.4	1.4	7.7	3.3	2008/11
USN3-PTBB	…	…	…	…	-0.4	0.7	2010/06
NIST-USN3	13.1	2.0	17.6	2.0	6.2	6.0	2010/08
closure					6.0	7.3	

We see that each of the closures agree with a sum of 0 within the uncertainty. However, there is a large variation in the uncertainties for the calibrations.

VI. ABSOLUTE CALIBRATIONS

For supporting TAI, only differential calibrations are required; absolute calibrations are not necessary. Nevertheless, without any absolute calibrations, the entire TAI network could walk off. Also, it is convenient to have absolute calibrations, as this can be accomplished at a single lab without reference to any other [9]. However, absolute calibrations are generally less accurate than differential. We include a history of some absolute calibrations at NIST, as they give a measure of both the accuracy and difficulties of absolute calibration as well as the stability of receiver delays over time.

The first primary GPS receiver at NIST was NBS10.

NBS10

June 1986	53 ns	NRL
April 1987	57 +/−5 ns	
Sep 1998	54.4 +/−2 ns	

A new receiver named NIST has been primary since 2006. The 2006 Differential calibration of L1 against NBS10set the delays as L1, L2=-44.7, -44.7.

2007 NRL calibration receiver+antenna required change (not implemented) of

L1 +9.4 +/−0.4, L2 +2.3 +/−3.3.

2011-April USNO calibration of only the receiver required

L1 -25.2 +/−0.3, L2 -42.9 +/−0.3.

Comparing to the receiver only values from 2007 NRL

L1 -28.5 +/−0.4, L2 -45.4 +/−3.3 ns.

There appears to be a change of +3.3 +/−0.5 ns in the L1 channel, from 2007 to 2011.

VII. CONCLUSIONS

The historical GPS receiver calibration data appear mixed in quality. Some periods indicate potential for good repeatability, perhaps a few ns/year; But other periods show calibrations vary 10's of ns over a year or less. The closures are in line with the combined uncertainties, but these uncertainties can be as large as 7.3 ns for k = 1.

These results show that standard methods for calibrations are needed. Three aspects need standardization: 1) measurement methods, 2) processing methods, and 3) methods for computing uncertainties. This seems mandatory to reach the ns uncertainty level or below.

In any cases, regular calibrations are needed, either relative calibration campaigns or absolute calibration of separate elements of a GPS receiver chain. Coordinating GPS calibrations should be the goal to better support the generation of UTC. Receivers appear capable of supporting 1-2 ns stabilities over a year. Calibration uncertainties may be achievable at the 1 ns level. Hence differential calibrations should be accomplished approximately annually. Achieving this will take collaboration on the part of many laboratories and the BIPM.

VIII. REFERENCES

[1] M.A. Weiss, V. Zhang, W. Lewandowski, P. Uhrich, and D. Valat, "NIST and OP GPS Receiver Calibrations Spanning Twenty Years: 1983-2003," Proc. 2004 EFTF Conf.

[2] T. Feldmann, A. Bauch, D. Piester, M. Rost, E. Goldberg, S. Mitchell, and B. Fonville, "Advance GPS-Based Time Link Calibration with PTB's New GPS Calibration Setup," Proc. 2010 PTTI, pp. 509-526.

[3] D. W. Allan and C. Thomas, "Technical directives for standardization of GPS time receiver software," Metrologia, 31, 69-79, 1994.

[4] RINEX format, http://igscb.jpl.nasa.gov/igscb/data/format/

[5] J. Kouba, A Guide to Using International GNSS Service (IGS) Products, http://igscb.jpl.nasa.gov/components/usage.html, May 2009.

[6] G. Cerretto, P. Tavella, F. Lahaye, "Statistical Constraints on Station Clock Parameters in the NRCAN PPP Estimation Process," Proc. 40th PTTI, pp. 441-458, 2008.

[7] BIPM website, http://www.bipm.org/jsp/en/TimeCalibrations.jsp

[8] P. Uhrich and D. Valat, "GPS receiver relative calibration campaign preparation for Galileo In-Orbit Validation", Proc. of the 24th European Frequency and Time Forum (EFTF), Noordwijk, The Netherlands, April 2010 (CD-Rom).

[9] A.Proia, J. White, D. Wilson, K. Senior, and G. Cibiel, "Absolute Calibration of GNSS Time Transfer Systems: NRL AND CNES Techniques Comparison," to be published in these proceedings.

AUTHOR INDEX