A 16 pJ/bit 0.1-15Mbps Compressive Sensing IC with on-chip DWT Sparsifier for Audio Signals

Gaurav Kumar K, Baibhab Chatterjee and Shreyas Sen

Purdue University, West Lafayette, USA

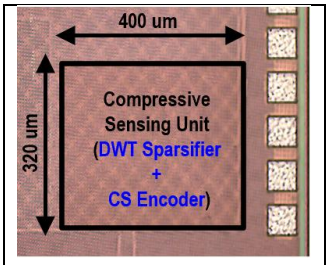
The emergence of Audio-based Augmented Reality has been calling for increasing data-rates for audio signals, with significant reduction in power to enable extremely energy-constrained sensor nodes. Typically, the communication power dominates sensing and computing power in a node [1]. For highly energy constrained scenarios, compressive sensing (CS) have been demonstrated (Fig. 1), where samples are first compressed at the sensor to contain the same information in a smaller number of samples, before transmitting to a receiver, where the signal is reconstructed. Previous CS works [2-5] have focused entirely on "sparse" physiological signals, operating in low speed regime. This work illustrates the first CS design, enabled with a discrete wavelet transform (DWT) sparsifier for catering to non-sparse signals such as high definition audio. Audio recording and playback are quite sensitive to quality, thereby requiring audio codecs, such as .aac, for efficient compression and decompression of audio streams, which usually consume power in the order of mW [6]. Audio inferencing operated in intelligent assistants are more tolerant to input quality, functioning effectively when the *Perceptual Evaluation of Audio Quality Mean Opinion Score* (PAEQ MOS) [7], an ITU-R standard objective metric for characterizing perceived audio quality, exceeds 1.5. CS presents an opportunity to achieve >10X reduction in transmitted audio data with orders of magnitude lower power, as compared to codecs. The design is implemented in 65 nm CMOS and consumes 238 uW power at 0.65 V and 15 Mbps.

Fig. 2 shows the block-level overview of the proposed CS architecture. The CS implementation takes in N input samples and compresses them into M linear combinations, using on-chip DWT sparsifier and the sensing matrix Φ, realizing a matrix-multiplication, i.e. CS Encoding. Thus, instead of N input samples, M compressed samples are transmitted, thereby reducing the communication power of the sensor by approximately N/M, which consequently lowers the overall sensor power. The design works with audio in both raw format (such as .wav – in Mode 1), and compressed format (such as .aac – in Mode 2). Due to additional codec-based compression in Mode 2, it further reduces the communication power with an increase in computation power, whereas Mode 1 provides a new modality with low-power audio compression. Finally, to ensure operation at high speeds, the entire design is realized as a pipelined architecture.

The CS IC comprises of pre-processing unit, the CS encoder and a serializer, followed by a transmit buffer. The signal reconstruction and performance evaluation are performed off-chip in MATLAB. For reconstruction, L1-norm minimization is done by Basis Pursuit algorithm, using the received compressed bits, sensing matrix and signal basis. The performance of compression is evaluated using 1) Signal-to-Noise-Distortion-Ratio (SNDR) [2] post reconstruction, which measures time-domain signal correlation and 2) PAEQ MOS.

The input signal is digitized to 8 bits and is fed serially to the CS IC. The pre-processing units, consisting of a deserializer and sparsifier, deserializes and then sparsifies the input samples, before feeding to the CS encoder for compression. A 2-stage Discrete Wavelet transform (DWT) is chosen for sparsifying the audio over the Discrete Cosine Transform (DCT). Although DCT offers better sparsity than DWT for an audio signal, yet DCT, being a lossy compression, fails to maintain the signal integrity, thereby exhibiting poor signal reconstruction. Fig. 3 presents the circuit implementation of DWT sparsifier, along with the quantitative analysis comparing the performance of DCT and DWT based sparsification. The DWT Sparsifier can be bypassed using an external scan bit.

The output of the sparsifier, supplied to the CS Encoder, includes majority of elements with zero value. Thus, in order to reduce the power consumption, the CS encoder, including sensing matrix generator and matrix multiplier, is implemented in a Wake-up mode, such that the block is power gated when the input is zero, preserving the previously accumulated values in the registers. For every non-zero input sample, the Matrix multiplier accumulates the product of input sample and the row elements of sensing matrix Φ in M 16-bit accumulators. As the value of M directly impacts the power and area overheads, it is chosen to be 30, considering the trade-off between signal reconstruction time and output audio quality as shown in Fig. 3. The degree of compression, defined as compression factor (CF) = N/M, is determined by number of samples to be accumulated (N), which can be set using 3-bit external scan, and can vary in the range 3X-33.3X for both the modes. To prevent data overflow, the accumulator size is chosen to be 16. After every N samples, the seeds of the PRBS generators in the sensing matrix are re-seeded and the serializer serializes the accumulated values, before transmitting through the driver. The accumulators are initialized to zero for next N samples. Fig. 4 shows the implementation of the proposed Wake-up CS encoder.



Die micrograph.

For reliable reconstruction at the receiver, the rows of sensing matrix are required to be "uncorrelated" with each other. A Dual PRBS-based sensing matrix [2], consisting of two independent PRBS generators, produces the required random Bernoulli matrix by XOR-ing the output of one PRBS generator with every state of other PRBS at lower area and power costs. A Seed update signal is sent to seed generator to update the seed of the sensing matrix for every incoming input sample. Fig. 4 also shows the circuit implementation of the sensing matrix generator.

The entire CS block occupies 400 µm X 320 µm area in 65nm CMOS technology, consuming 2.4-238 µW for data rates varying from 0.1-15 Mbps. The output waveforms, observed at the oscilloscope, are captured, and processed in MATLAB. Fig. 5 shows the timing diagram of the input audio signal and output reconstructed signal for CF of 3. The test-chip power and specifications are also shown in Fig. 5. The consumed power scales with the input data rate, while staying almost independent of CF for a particular data rate. The leakage power dominates the low speed regime and increases with supply voltage. The shmoo plot shows the functionality of the chip at different supply voltages. Fig. 6 provides the optimal supply voltage needed for operating at different data rates. It is found that for a 15 Mbps input signal, the minimum supply voltage required is 0.65 V. Fig. 6 also shows that CF below 15X provides acceptable signal reconstruction. These points correspond to PAEQ MOS >1.5 and SNDR > 0.2 dB. Note that inherently sparse signals can be operated at higher CF because the signal information present in a compression window is much less and thus can be represented by a lower value of M.

The pipelined CS implementation with Wake-up Encoder and on-chip DWT sparsifier is the first high-speed CS targeted for audio signals in energy constrained sensor nodes. It achieves the highest input data rate (up to 15 Mbps) in literature, while the Wake-Up operation facilitates low power operation (2.4-238 uW). Fig. 6 compares the present work with other related works in the literature. The implemented architecture is generic enough to be applied to both sparse and non-sparse signals (all four quadrants in Fig. 1).

References:

[1] N. Cao *et al.*, "A 65nm Image Processing SoC Supporting Multiple DNN Models and Real-Time Computation-Communication Trade-Off Via Actor-Critical Neuro-Controller," *VLSIC*, June 2020.

[2] F. Chen *et al.*, "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *JSSC*, March 2012

[3] X. Liu *et al.*, "A Fully Integrated Wireless Compressed Sensing Neural Signal Acquisition System for Chronic Recording and Brain Machine Interface," *TBioCAS*, Aug. 2016.

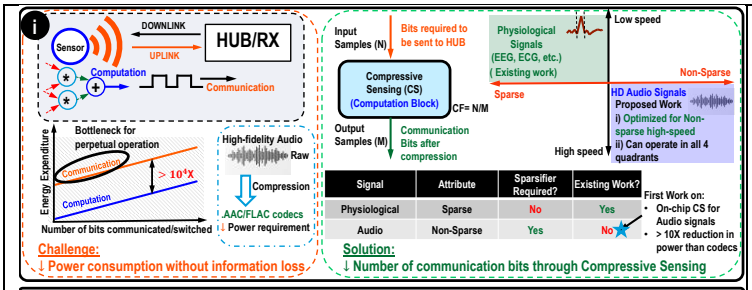
[4] V. Behravan *et al.*, "A compressed-sensing sensor-on-chip incorporating statistics collection to improve reconstruction performance," *CICC*, 2015

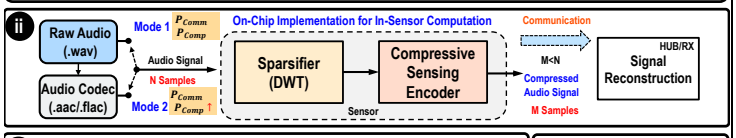
[5] D. Gangopadhyay, *et al.*, "Compressed Sensing Analog Front-End for Bio-Sensor Applications," *JSSC*, Feb. 2014

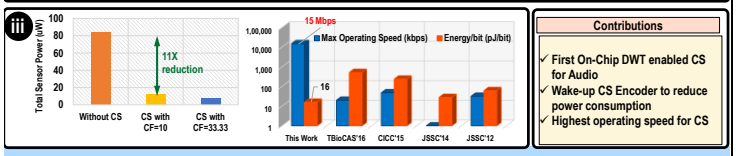
[6] X. Jiang *et al.*, "A Low-Power, High-Fidelity Stereo Audio Codec in 0.13 um CMOS," *JSSC*, May 2012

[7] T. Thiede *et al.*, "PEAQ---The ITU Standard for Objective Measurement of Perceived Audio Quality", *Journal AES*, Dec 1999

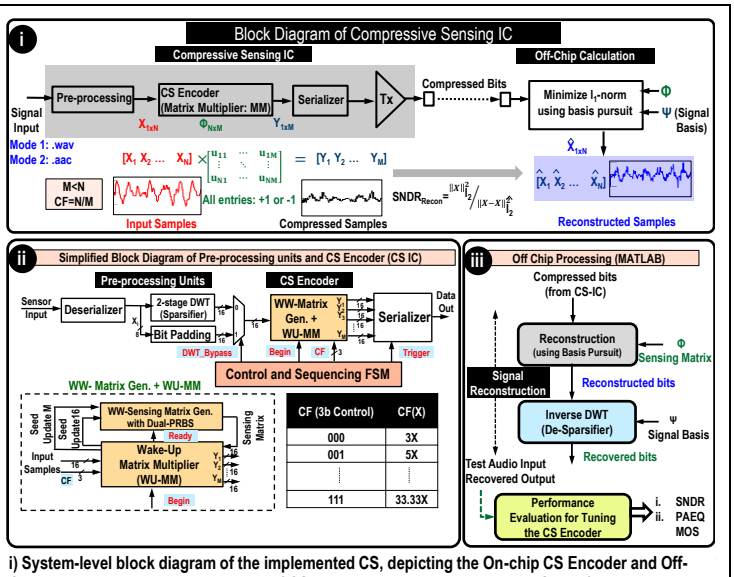
2021 IEEE Custom Integrated Circuits Conference (CICC) | 978-1-7281-7581-2/20/$31.00 ©2021 IEEE | DOI: 10.1109/CICC51472.2021.9431569



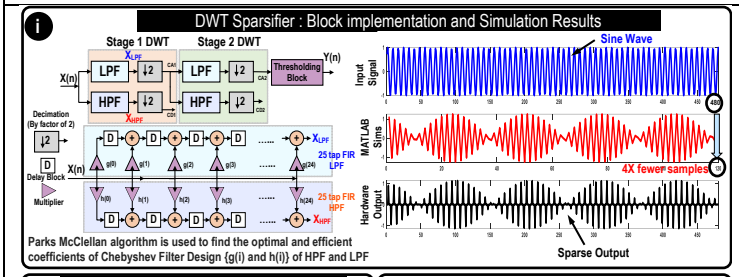




comparison with related works and Contributions of the proposed work.

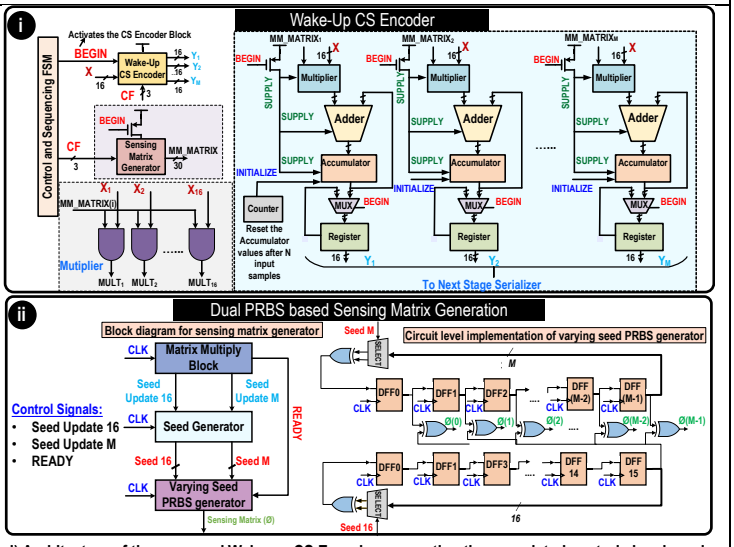


i) System-level block diagram of the implemented CS, depicting the On-chip CS Encoder and Off-Chip Reconstruction; ii) Architecture of CS Block, utilizing the 2-stage DWT Sparsifier and Wake-Up Encoder; iii) Design Flow of Off-Chip processing for signal reconstruction.



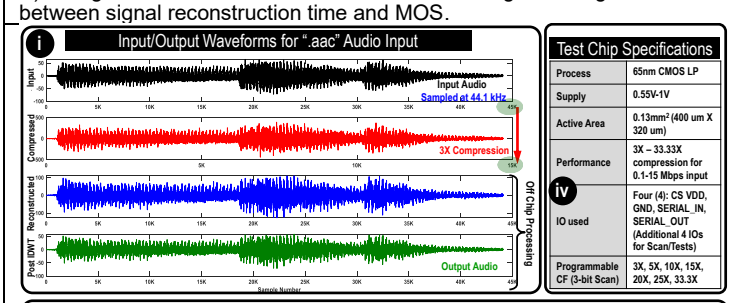
i) Circuit implementation of DWT Sparsifier, along with functional validation with sinusoid input; ii) Benefits of DWT over DCT in terms of MOS and SNDR; iii) Design choice of value of M in the CS Encoding, showing the trade-off between signal reconstruction time and MOS

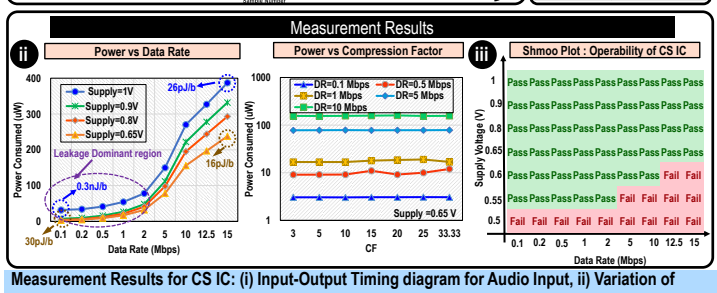
between signal reconstruction time and MOS.



i) Architecture of the proposed Wake-up CS Encoder, presenting the associated control signals and the power-gating to reduce the overall power consumption; ii) Control flow and Circuit Diagram of the Dual PRBS sensing matrix

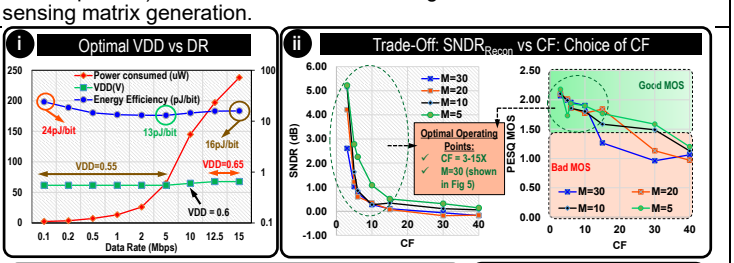
sensing matrix generation.

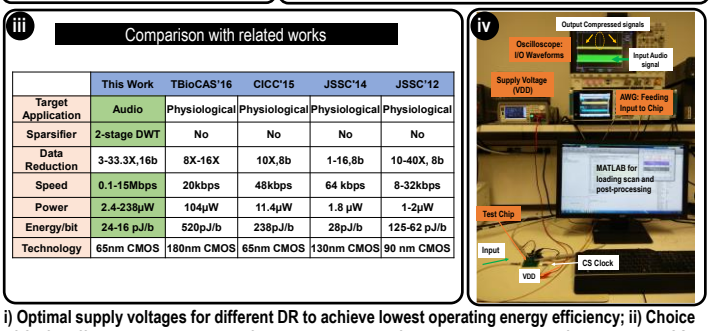




Measurement Results for CS IC: (i) Input-Output Timing diagram for Audio Input, ii) Variation of power with input Data Rate and Compression factor, iii) Shmoo plot, depicting the operability of the IC at different DR and Supply voltages, iv) Chip Specifications

Supply voltages, iv) Chip Specifications.





	This Work	TBioCAS'16	CICC'15	JSSC'14	JSSC'12
Target Application	Audio	Physiological	Physiological	Physiological	Physiological
Sparsifier	2-stage DWT	No	No	No	No
Data Reduction	3-33.3X,16b	8X-16X	10X,8b	1-16,8b	10-40X, 8b
Speed	0.1-15Mbps	20kbps	48kbps	64 kbps	8-32kbps
Power	2.4-238μW	104μW	11.4μW	1.8 μW	1-2μW
Energy/bit	24-16 pJ/b	520pJ/b	238pJ/b	28pJ/b	125-62 pJ/b
Technology	65nm CMOS	180nm CMOS	65nm CMOS	130nm CMOS	90 nm CMOS

i) Optimal supply voltages for different DR to achieve lowest operating energy efficiency; ii) Choice of CF for effective reconstruction of audio signal; iii) Performance comparison of implemented CS IC with state-of-the-art CS Encoders; iv) Measurement Setup

Encoders; iv) Measurement Setup.