SCALING ISSUES AND SOLUTIONS IN ULTRA SCALED MOSFETS USING PREDICTIVE MODELING

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Mehdi Salmani-Jelodar

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

Dec 2014

Purdue University

West Lafayette, Indiana

To my parents and my sisters.

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Professor Gerhard Klimeck, for the valuable insights he provided for this research and his continous support of my PhD. I would like to thank him for providing all the necessary computing resources. He introduced me to the world of high performance computing and its applications, the problem solving in Physics and Engineering, which will be an invaluable experience for my future career pursuits. I would also like to thank Dr. K. Ng, Professor M. Lundstrom and Professor P. Ye for serving on my advisory committee. I am grateful to Hesam Ilati, Matthias Tan, Tarek Amin, Saumitra Mehrotra and Dr. Sebastian Steiger as great friends and colleagues for their supports and help in life, education and research. Also, I am thankful to Dr. SungGeun Kim, Prof. Mathieu Luisier, Yaohua Tan, Dr. Abhijeet Paul, Daniel Mejia, Dr. Ganesh Hegde, Dr. Seung Hyun Park, Dr. Parijat Sengupta, Dr. Sunhee Lee, Zhengping Jiang, Daniel Valencia for stimulating discussions. I would also like to thank our current NEMO5 team Dr. Jim Fonseca, Dr. Bozidar Novakovic, Prof. Michael Povolotsky, Prof. Tillman Kubis who have been of great support. I would like to thank all my colleagues from the NCN and the Klimeck group for providing a stimulating and fun environment in which to learn and grow. I also would like to special thanks to NCN staff members Megan Rush, Vicki Johnson, Cheryl Haines, Amanda Buckles and Leslie Schumacher for their help over the last six years. Outside of my office, I always had very strong support from my wonderful friends, including Behtash Behin-aein, Nichole Ramirez, Ali Jalali, Trinity Overmyer, Ali Jahanian, Sofia Singlton, Alicia English, Behtrooz Jalalahmad, Alicia Philips, Teimor Maleki, Lars Suderland, Golsa Moaieri, Ali Ahmadian, Ali Razavieh and Sebastian Feste.

Lastly, and most importantly, I wish to thank my parents and family for their unconditional love and support.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABSTRACT

Salmani-Jelodar, Mehdi Ph.D., Purdue University, Dec 2014. Scaling Issues and Solutions in Ultra Scaled MOSFETs using Predictive Modeling . Major Professor: Gerhard Klimeck.

Channel length of metal oxide semiconductor field effect transistors (MOSFETs) are scaling below 20 nm. At this scale, quantum mechanical effects, including source to drain tunneling and quantum confinement play an increasingly important role in predicting device performance. Accurate projections of device characteristics are of high interest in the semiconductor industry. This work presents a semi-empirical model based quantum transport tool, which is used for accurately predicting the performance of double gate MOSFETs over the next 15 years as part of the International Technology Roadmap for Semiconductors (ITRS). The results show ON-current and performance degradation as a result of source to drain (SD) tunneling, and band structure alteration and supply voltage reduction due to scaling. Furthermore, the impacts of SD tunneling in ultra-scaled devices are investigated. In particular, heavy mass materials and the lightly doped drain are proposed as solutions for SD tunneling. Thick gate stacks can degrade electrostatics in ultra-scaled MOSFETs. Here, we present an approach to find optimum oxide thicknesses in order to prevent gate leakage and optimize device electrostatics. Lastly, we analyze the confined SiGe based P-type MOSFET as a promising candidate for the next technology node in the semiconductor industry. Reducing the thickness of the $< 110 >$ SiGe channel improves ballistic ON-current and increases the band gap, which is favored for reducing leakage. These changes are quantified and shown to become more pronounced for higher germanium percentages. Moreover, while strain plays an important role in improving the ballistic ON current, this effect diminishes for very thin channel thicknesses.

1. INTRODUCTION

With the continual scaling of CMOS transistors, the semiconductor industry is facing new challenges. For the last four decades, Silicon, as a channel material has been the main driver behind the growth in semiconductor based technologies. For channel lengths under 100 nm, conventional Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) scaling with Si has faced different challenges. These challenges include: meeting the criteria for the ON-current, short channel effects (SCEs) and OFF-current requirements. Different engineering innovations have helped improve the device performance, while pushing MOSFETs scaling far beyond expected, as reflected in the CMOS scaling roadmap of Intel depicted in Fig. 1.1. One of these innovations is the application of uniaxial strain for higher ON-current [1]. Later, High-k/ metal gate technology to alleviate gate leakage, that allows further scaling [2]. Also, introduction of new device architectures including Silicon-on-insulator (SOI) and FinFET to improve the electrostatics and reduce the leakage current [3].

As depicted in Figs. 1.1 and 1.2, the number of critical atoms and electrons under the gate will be countable with the scaling of the minimum feature size. This brings different quantum mechanical phenomena to the picture. For example, electrical properties of materials depend on geometry parameters such as confinement, orientation and strain. As depicted in Fig. 1.3, Si nanowire band structure can change drastically with orientation and strain. Additionally, gate and source-to-drain (S/D) tunneling occurs with the miniaturization of the gate oxide and channel length [4–6]. All of the aforementioned phenomena raise the complexity of projecting the next technology nodes in the semiconductor industry [4]. At such small scales, predictive models that provide insight and assist device engineers towards designing and innovating ultra-scaled MOSFETs are crucial. These models should be able to capture the essential physics for ultra-scaled devices [7].

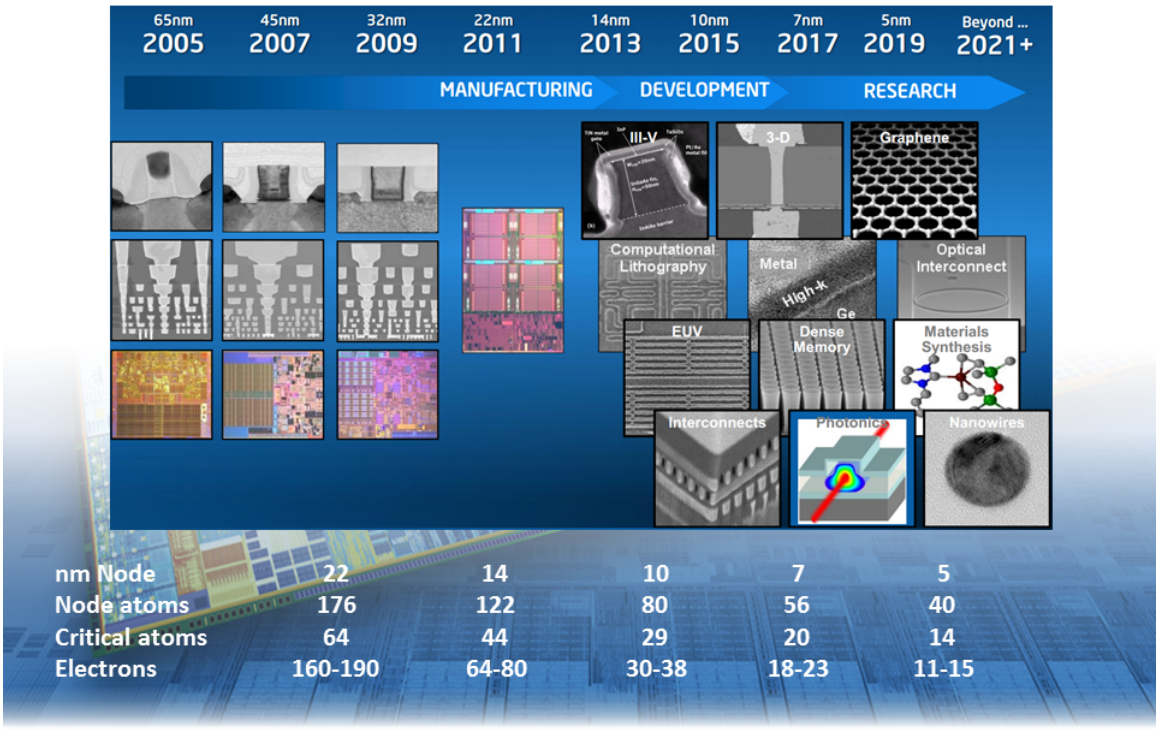

Fig. 1.1. The Intel MOSFET scaling trend from 2005. Multiple innovations have pushed CMOS scaling this size and more innovations are needed to continue the scaling to 5 nm and lower. Number of atoms along the gate length (node atoms) and the gate width in FinFETs (critical atoms) and number of electrons in the channel are shown. These numbers are scaled down to a few atoms and electrons, which are not possible to accurately model with conventional drift diffusion or compact models. The image is adapted from [8].

For atomic scale MOSFETs, non-equilibrium green's function (NEGF) is a computationally feasible method to project transistor's performance. NEGF, quantum transmission boundary method (QTBM) and similar atomistic approaches work, based on semi-empirical models. Using these atomistic models we projected transistor scaling and explored possible issues. By scaling down the gate length, source to drain tunneling deteriorates the device performance. We explored the impact of S/D tunneling and provided some solutions. By replacing thin SiO_2 with thick high-k gate stacks, the gate control over the channel weakens. We explored the impact of using

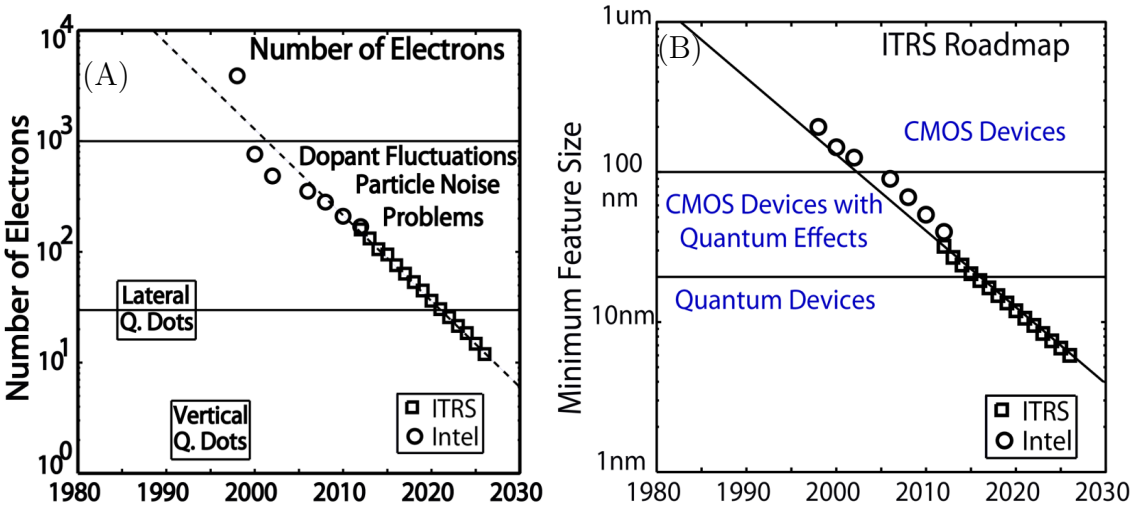

Fig. 1.2. Scaling effects on number of atoms and electrons. (A) Shows the number of electrons under the gate based on ITRS [9] and Intel's published data as $NumberOfElectrons = (V_{DD} - V_T) \times \frac{C_G}{q}$, where V_{DD}, V_T, C_G and q are the supply voltage, the threshold voltage, the gate capacitance and electron charge, respectively. (B) Shows scaling of minimum feature size in CMOS technology based on Intel and ITRS roadmap data.

high-k gate stacks on device performance and provided guidelines for using different high-k oxide material.

Silicon has been the dominant channel material in the semiconductor industry. Recently, Silicon Germanium (SiGe) has received a lot of attention for the next technology nodes. Ge is integrated into the technology process in contacts to apply strain to the channel, which makes the integration process of the SiGe to CMOS a more feasible choice compared to III-V materials or other exotic semiconductors such as MoS2. In the last part of this thesis, using confined SiGe as a channel material is analyzed. Next, sub-sections briefly explain the entire work.

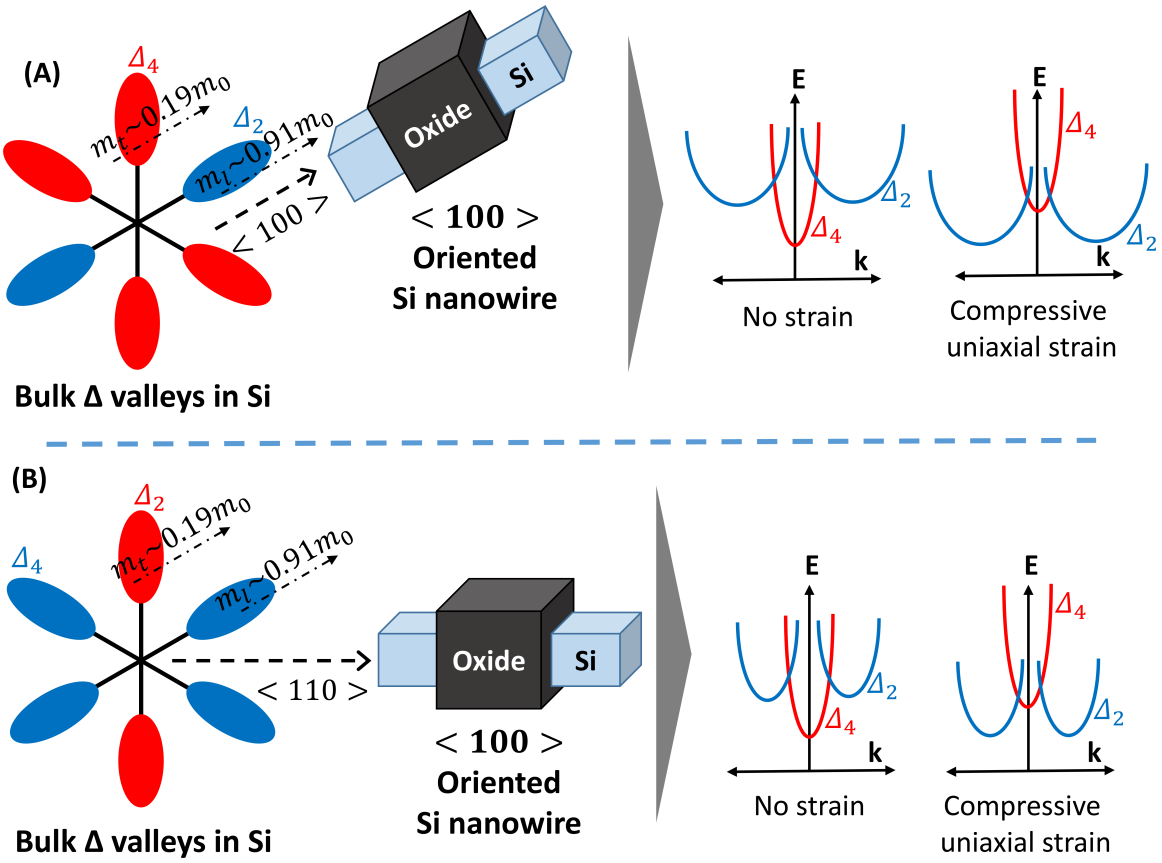

Fig. 1.3. Schematic description of the variation of the electronic band structure (a) < 100 > and (b) < 110 > oriented Si nanowire under no-strain and compressive-strain conditions. Δ_2 and Δ_4 has 0.19 m_0 and 0.91 m_0 effective masses, respectively. Both orientation and strain can vary band structure, and consequently, the material behavior in MOSFET channel.

1.1 Semi-empirical Model Development

To continue MOSFET scaling and designing, nanoscale device predictive modeling is crucial [7]. As the size of devices are scaling below 30 nanometers, atoms are countable and different critical physics come into the picture. This creates the need for atomistic models that can accurately capture quantum mechanical effects. There are several competing methods, each with its own advantages and disadvantages.

The introductory nature of this text prevents us from going into details of these approaches. One finds excellent descriptions in different books devoted to electronic structure calculations [10]. The choice of a particular atomistic calculation generally comes down to an optimization of the following factors:

a) Speed of operation: Modeling realistic devices with non-periodic structures in a reasonable time **over feasible computational resources** is usually the prime consideration.

b) Precision: While time is critical, obtaining physically feasible results that match relevant experiments is another deciding factor.

The most well developed and accurate calculations found in commercial software are a category of methods called the *ab-initio* which try to adopt a "first-principles" approach, i.e. no pre-assumptions. The accuracy of these methods is offset, however, by their disadvantages. Commercial codes can typically handle bulk crystalline materials very well. However, they become prohibitively time consuming and computationally expensive in the case of realistic device sizes having millions of atoms and defects [11]. Other methods, such as semi-empirical methods, are usually preferred for applications, such as tight-binding and enhanced valance force field models. These semi-empirical methods have an acceptable trade-off between the above deciding factors.

Semi-empirical models with physics-based symmetries and properties are usually developed by fitting a few to hundreds of parameters to reproduce the experimental or first-principle calculated data. Enhanced Valence Force Field (EVFF) method for lattice properties is one of the models that we developed by fitting a few system parameters using an evolutionary computing method. An accurate understanding of lattice properties provides a stepping stone for the investigation of thermal phenomena and has a large impact in thermoelectricity and nano-scale electronic device design. The VFF method allows for the calculation of static properties such as elastic constants, as well as dynamic properties including the sound velocity and phonon dispersion. In this thesis, a parallel genetic algorithm (PGA) is employed to develop

the optimal VFF model parameters for Gallium-Arsenide (GaAs) and InAs. This methodology is also used for fitting parameters for strained Silicon, and can be used for other semi-empirical models. The achieved results agree qualitatively and quantitatively with the experimental or *ab-initio* data. Chapter 2 explains semi-empirical model development and an application of it in detail based on published research in [7, 12–16].

1.2 CMOS Transistor Roadmap Projection

Due to the significant resources and investments required to develop the next generation of complementary-MOSFET (CMOS) technologies, it has been necessary to identify clear goals and put collective efforts towards developing new equipment and technologies. The semiconductor roadmap represents a consensus among industry leaders and projects needs based on past trends. The International Technology Roadmap for Semiconductors (ITRS) [4,9] is the standard accepted roadmap to identify future requirements of CMOS technology.

The target performance metric values (i.e. on current and DIBL) for different MOSFETs in ITRS device tables previously has been extracted with compact models, which do not capture the required physics for predictions in ultra-scaled devices. As device sizes are approaching sub-25 nm and their body thicknesses are going down to a few nanometers, different quantum mechanical effects come into the picture and simple compact models or standard Drift Diffusion method cannot capture these effects. In ultra-scaled MOSFETs some major electrical properties of materials are geometry [**?**]. Models that capture essential physics are required [7]. We used a quantum transport tool to help the semiconductor industry to identify the issues for scaling. It is found, that in addition to the short channel effects strengthening, S/D tunneling will be a major issue in scaling MOSFETs. S/D tunneling increases the subthreshold swing (SS), consequently degrading the device performance. S/D tunneling effects are analyzed; Two solutions are proposed based on doping profile and band

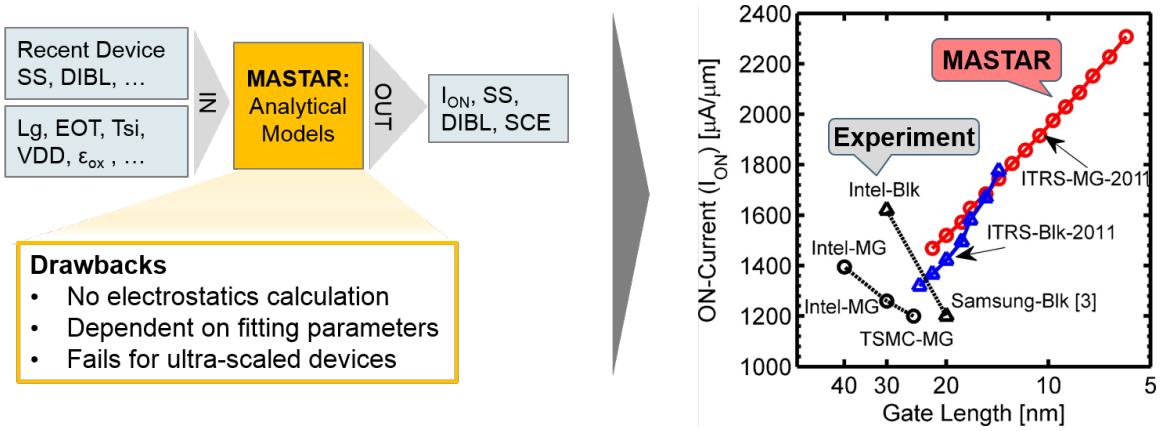

Fig. 1.4. On the left, the flowchart shows MASTAR's inputs and outputs. MASTAR gets SS and DIBL form recent devices [17]. Therefore, the calculations are very dependent on the recently fabricated MOSFETs. Specifically, for future nodes, analytical approaches cannot capture critical physics, such as quantum confinement and source to drain tunneling. Therefore, the projection results can be far from practical MOSFETs performance. Current MASTAR model lacks critical physics, which leads to inaccurate projections for even near future ITRS nodes, which are shown in the right side.

structure engineering to reduce the negative effects of S/D tunneling [19]. Quantum confinement reduces the transport mass and increases the injection velocity. However, in ultra scaled MOSFETs, it raises the S/D tunneling ratio, which leads to lower V_{th}, and consequently reduces the ON-current. Chapter 3 explains CMOS transistor scaling projection for the next 15 years in detail, based on published research in [4, 20–22].

1.3 Design Guidelines for ultra-scaled MOSFETs (L_{ch} < 12 nm)

Traditional thinking assumes that using a light effective mass (m^*) high mobility material will result in higher performance transistor. However, sub-12 nm metal-oxide-semiconductor field effect transistors (MOSFETs) with light m^* may underperform compared to standard Si, as a result of source to drain (S/D) tunneling. An

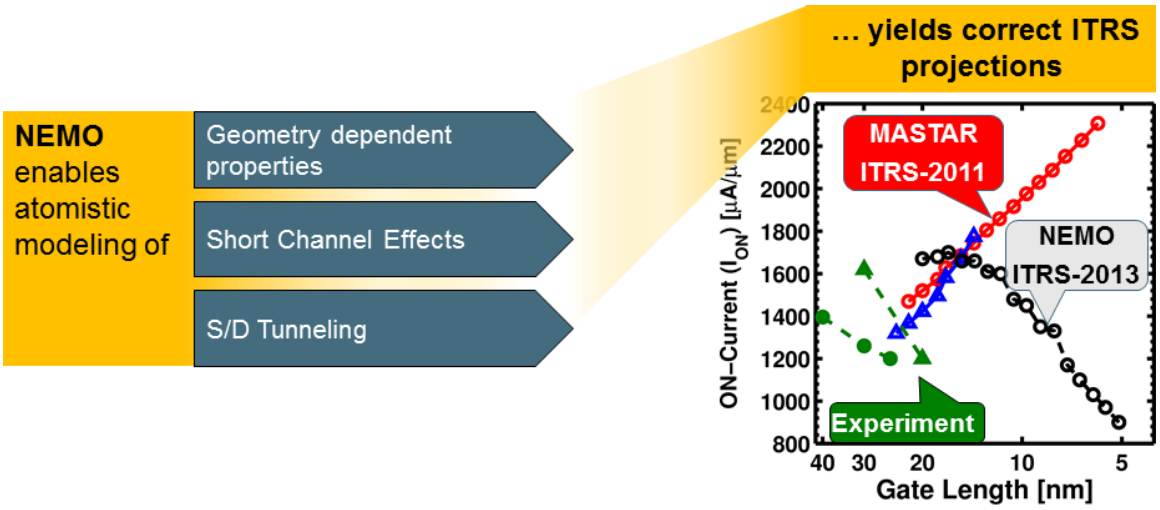

Fig. 1.5. The calculated ON-current using MASTAR [17] for DG and bulk is compared with the recently reported best experimental data. The results for DG MOSFETs from NEMO tool suite [18] is shown and these results are aligned with the trend of experimental reports [4].

optimum heavier mass can decrease tunneling leakage current, and at the same time, improve gate to channel capacitance as a result of an increased quantum capacitance (C_q). A single band effective mass model has been used to provide the performance trends independent of material, orientation and strain. This paper provides guidelines for achieving optimum m^* for sub-12 nm nanowire down to channel length of 3 nm. Optimum m^* are found to range between 0.2 - 1.0 m_0 and more interestingly, these masses can be engineered within Si for both p-type and n-type MOSFETs. m^* is no longer a material constant, rather a geometry and strain dependent property of the channel material.

Currently, using light effective mass materials as channel has been appealing due to their high injection velocities. However, it has been shown at 12 nm channel length InGaAs ($m^* = 0.07m_0$) provides very similar SS, DIBL and ON-current compared to strained-Si for triple-gate MOSFETs [23]. By scaling the channel length further (below 12 nm and based on ITRS device table, this is the device gate length after

2019), the effect of S/D tunneling is playing a more important role in determining device performance metrics [6, 24]. This effect increases the direct tunneling leakage current, and consequently, the sub-threshold swing (SS) [6, 25, 26]. Achieving low SS is the key factor in scaling devices. To lower SS, FinFET structures are used and NW MOSFETs have emerged as promising candidates for ultra-short devices [27]. Gate-all-around NWs offer the best electrostatic gate control over the channel [6, 27]. Previous theoretical studies have explored the effects of channel materials, orientation and strain on tunneling in NMOS devices [6, 24, 26, 28–30]. These studies showed the benefit of using heavier mass materials over light effective mass materials [6]. However, these studies did not estimate the range of optimum heavy effective mass that might be needed to achieve both low SS and high ON-current for different channel lengths [6,24,26,28–30]. Additionally, these studies did not discuss the effect of heavy effective mass on quantum capacitance and electrostatics. PMOS devices were not discussed in detail in previous studies.

In Chapter 4, using a single band effective mass transport model we attempt to answer the following questions: 1) What is the optimum effective mass range, independent of channel material, orientation, strain and transport type (n-type or p-type MOSFETs), for NW MOSFET with a given channel length? 2) How does a heavy mass channel material impact device electrostatics? 3) Since m^* is a silicon geometry and strain dependent parameter, is it possible to engineer silicon in such a way to achieve optimum effective masses? Chapter 4 is written based on research work published as [4, 20, 31].

1.4 Optimum High-k Oxide for the Best Performance in Ultra-scaled MOSFETs

High-k dielectrics are a widely used technique to mitigate the gate leakage in the ultra-scaled metal oxide semiconductor field effect transistors (MOSFETs). High-k dielectrics provide the same equivalent oxide thickness (EOT) as SiO_2, as well as

thicker physical layers for the same EOT. However, using a thicker physical dielectric for the same EOT has a negative effect on the device performance due to the degradation of 2D electrostatics. In this thesis, the effects of high-k oxides on double-gate (DG) MOSFETs with gate length under 20 nm are studied. We find that there is an optimum physical oxide thickness (T_{OX}) for each gate stack, including SiO_2 interface layer and one high-k material. For the same EOT, Al_2O_3 (k=9) over 3 Å SiO_2 provides the best performance, while for HfO_2 (k=20) and La_2O_3 (k=30) SiO_2 thicknesses should be 5 Å and 7 Å, respectively. The effects of using high-k oxides and gate stacks on the performance of ultra-scales MOSFETs are analyzed. While thin oxide thickness increases the gate leakage, the thick oxide layer reduces the gate control on the channel. Therefore, the physical thicknesses of gate stack should be optimized to achieve the best performance.

In Chapter 5, we attempt to answer these questions: 1) What is the impact of using different high-k materials with the same EOT on an ultra-scaled DG MOSFET? 2) What is the optimum thickness of high-k gate stack for a fixed EOT? 3) How do we analytically estimate the gate leakage in the off state for a specific gate stack? Chapter 4 is written based on research work at [5, 32].

1.5 Quantum Confined SiGe MOSFETs

High mobility channel materials have been widely pursued for the continued miniaturization of transistors. Silicon Germanium is a leading candidate as a replacement channel material for PMOS. High mobility materials including SiGe and III-V have also been incorporated in modified FinFET geometries to lower the leakage resulting from short channel electrostatics. For example, SiGe cladding FinFET employs a thin SiGe cladding layer as the channel around a non-conducting Si fin to provide stronger gate control. In this work, atomistic simulations are used to obtain band structure and transport characteristics in the ballistic regime for SiGe channels between 2 to 8 nm in the presence of strain for various Ge mole fractions between zero to one.

We show that reducing thickness of the <110> SiGe channel improves ballistic ON-current and increases the band gap, which is desirable for lowering leakage. These changes are quantified and are more pronounced for a higher germanium percentage. Moreover, while strain has an important role in improving ballistic ON current, its effect diminishes for channel thicknesses below 3 nm. Lastly, we compare a cladding FinFET composed of a 3 nm SiGe channel over a Si fin with a bulk FinFET of the same material composition and footprint. The simulations show that cladding FinFET outperforms regular FinFET; largely as a result of quantum confinement (QC) effects. Chapter 5 is written based on research work, which has been done in collaboration with GLOBALFOUNDRIES and [33].

1.6 Thesis Organization

This thesis is organized as follows: Chapter 2 is a detailed discussion about semi-empirical model development. In Chapter 3, the need for rewriting the ITRS device tables is explained. Additionally, the process for the calculation of the tables are described along with some discussion regarding the results. In Chapter 4, S/D tunneling as an important performance degradation factor in ultra scaled devices is studied in detail. Chapter 5 provides the guideline regarding designing optimum high-k gate stack in ultra-scaled MOSFETs. Utilizing confined SiGe as a near-term solution in CMOS technology nodes is discussed in Chapter 6. In Chapter 7, the thesis is summarized and further development and research efforts for the future are suggested.

The presented work in this thesis is based on papers previously published or submitted to different journals and conferences. Some figures and content in this thesis have been reused from these publications. The permission for the reuse of the content and figures from the publishers has been obtained and appear in the Appendix.

2. SEMI-EMPIRICAL MODEL DEVELOPMENT

2.1 Introduction

Computational nanoscience is attracting aditional attention as the power of computers and clusters is increasing. The ability to accurately model nano-sized structures and their constituting materials is pivotal to understand and improve the state-of-the-art semiconductor materials and devices technology. Many modeling approaches use experimental and mechanical data to adjust required parameters for their models. These methods are termed semi-empirical. The underlying mathematical expressions are based on physical concepts such as specific types of interactions or symmetries. While the fundamental interactions are well understood, they can be extremely hard to quantify from *ab-initio* models. The interaction strengths are therefore treated as fitting parameters. Usually such approaches are based on regulating the model's input parameters to fit the model's output with measured experimental values or higher-order mechanical models. In some cases the number of inputs may be very few (two or three) and scientists find them by trial and error. However, it is more typical that the models are very complex and it is hard if not impossible to solve them analytically. The semi-empirical tight binding method for the material's electronic structure, for example, has between 16 to 200 parameters to fit [13]. In the case of many input parameters,an optimizer is needed to fit the inputs for specific experimental target values. The authors employed evolutionary computing approaches for semi-empirical tight binding successfully in the past [13, 34]. In these models the parameters have a physical meaning and the range of permitted values is usually constrained. The fitting process is also used to tweak and fine-tune the final model itself.The semiconductor device simulation tools which implement the models are monolithic and need several tenths of a second to several hours of computation

time. Developing accurate models for lattice properties is essential for the design of nanoscale electronic and optoelectronic devices [35]. The development of a model based on an atomistic representation of the crystal is still challenging. One of these approaches is the semi-empirical Valence Force Field (VFF) model [36, 37]. Based on accuracy and complexity different types of VFF models have been proposed in the literature [37][7]. These models are comparatively simple and have deficiencies in capturing the physics. By adding additional physical interaction terms to the model and using genetic algorithm to find the related parameters, a more precise model can be constructed. Here, an 8-parameter VFF models developed which includes all nearest-neighbor as well as the co-planar second nearest-neighbor interactions. The model is explained in the next section. For fitting the VFF model with experimental data three issues need to be considered: a) objectives must be met to obtain a variety of physical characteristics of the material, b) model inputs have constraints in order to retain their physical meaning, and c)the dependence of the model results on the 8 input parameters is very complex and nonlinear. These issues and the complexity of the fitness landscape enforced us to use an evolutionary computing approach to solve this multi-objective problem [38]. Genetic algorithms(GAs) have been employed in the past to produce reasonable results for such problems.Each chromosome's evaluation takes a few tens of a second. For these reasons a parallel genetic algorithm approach(PGA) is employed over an eight-core cluster to speed up the optimization process. In related work, Kane used a weighted least squares approach to fit six input parameters of his model against a few selected points of the dispersion relation [39]. Martin's model has only two parameters that are fitted against the elastic constants −a single objective [40]. Lazarenkova et al. proposed three additional parameters where one of them was derived analytically and the remaining two were found by trial and error [41]. All these approaches had a single objective and few input parameters. This chapter features a more complex model with 8 input parameters which is fitted against three different objectives. Due to this complexity GA is employed. The model is applied to the semiconductor gallium arsenide (GaAs) [35], but the method-

ology can also be used for other zincblende materials like InAs and InP. The chapter is organized as follows: Section 2.2 provides an overview on the lattice dynamics and the Valence Force Field (VFF) model. Section 2.3 is a brief introduction to standard GA and PGA. In Section 2.4 we explain the fitness function and the used parameters for the PGA and then show implementation results for fitting parameters of the VFF model to multiple objectives. After that we show the model is predictive and its capability to be extended for other materials. In Section 2.5, an application of the model for calculation of thermal properties in nano-structures is described. Section 2.6 concludes with finishing remarks.

2.2 Lattice Dynamics and the Valence Forced Field Method

This section gives a brief introduction to the valence force field model and its connection to the vibrational modes, or phonons, of a semiconductor crystal. Static and dynamic lattice dynamics of semiconductors play a decisive role in electronics. They determine the thermal conductivity, which is a limiting factor in the performance of today's transistors. Scattering between electrons and phonons typically deteriorates the speed of the device and introduces dissipation and heating, the latter being the main limiting factor in the operation frequency of today's transistors.

2.2.1 Valence Force Field Model of the Crystal Energy

The valence force field (VFF) model provides a fundamental and microscopic description of lattice properties [37]. Semiconductor devices are typically made of materials like GaAs where atomic bonds are to a large extent covalent. This makes the interaction short-range, as opposed to ionic crystals like NaCl where the bond is largely based on long-range Coulomb interactions. The VFF method expresses the total crystal energy as a functional of the bond angles and bond lengths, as depicted in Figure 2.1. The functional reads [42]: Here NN denotes nearest neighbors, COP stands for coplanar bonds, i,j,k,l are atom indices, r and θ are bond lengths and

$$U = \frac{3}{2\cdot 8}\sum_{i,j\in NN(i)}\Bigg[\alpha_{ij}(\delta r_{ij})^2 + \sum_{k\in NN(i)}^{k\neq j}\Bigg(\beta_{jik}(\delta\theta_{jik})^2 + \gamma_{jik}\delta r_{ij}\,\delta\theta_{jik} + \delta_{jik}\delta r_{ij}\,\delta r_{ik} + \sum_{l\in NN(k)}^{COP(j-i-k-l)}\nu_{jikl}(\delta\theta_{jik})(\delta\theta_{ikl})\Bigg)\Bigg] + \frac{1}{2\cdot 4\pi\epsilon_0}\sum_{i,j\neq i}\frac{Z_iZ_j}{|\mathbf{r}_i-\mathbf{r}_j|}. \quad (1)$$

$$\delta r_{ij} \equiv \frac{(r_{ij}^2-d_{ij}^2)}{d_{ij}}, \qquad \delta\theta_{jik} \equiv \frac{(\mathbf{r}_{ij}\cdot\mathbf{r}_{ik}-\cos\theta_0 d_{ij}d_{ik})}{\sqrt{d_{ij}d_{ik}}}$$

angles, and $\alpha, \beta, \gamma, \delta$ and ν are empirical force constants. d is the equilibrium bond length and Z are fractional point charges. The most commonly used description, known as Keating model [36], neglects all but the first two contributions [37, 43].

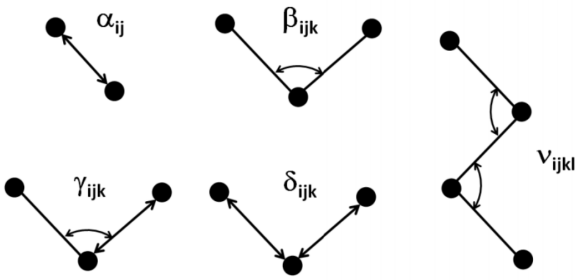

Fig. 2.1. Sketch of the short-range interactions which constitute the valence force field model for the lattice energy.

In mono-atomic crystals like Si the model's inputs are 5 parameters including $\alpha, \beta, \gamma, \delta$ and ν. In diatomic crystals like GaAs two different values of β, γ and δ are possible depending on what atom sits at the apex of the two bonds, bringing the number of parameters up to 8. It is well-known that the bonds in zincblende-type semiconductors like GaAs, which are binary compounds consisting of atoms with different radii, do exhibit a partially ionic character. To account for this, the Coulomb interaction between point charges fixed at the mean atomic positions was added (rigid ion model), which is represented by the last term in Eqn. (1). In 3D-periodic lattices this long-range interaction can be evaluated using Ewald summation [44].

2.2.2 Elastic Constants

Nanostructures, i.e. structures at the size of few to hundreds of nanometers which are composed of different materials, typically comprise some lattice mismatch because the natural bond lengths of crystals differ. Structures consequently exhibit strain where the atoms are distorted from the natural positions they would take if only a single material was present. The microscopic description of the VFF model allows finding these atomic positions by writing down the energy functional for the structure and minimizing with respect to the atomic coordinates, a process known as strain relaxation. Calculations of strain in macroscopic materials, such as buildings or vehicles, are commonly performed using continuum elasticity theory. The model parameters for this theory are the elastic constants, well-established material parameters that are easily accessible by experiment. Due to the symmetry of zincblende-type lattices, only three constants C11, C12 and C44 suffice to describe zincblende nanostructures. A connection to the microscopic VFF model is established by performing a Taylor expansion of the energy in (1) for a bulk crystal around the equilibrium bond length. This expansion can be rewritten in terms of the strain tensor. Coefficient comparison yields analytical relations between the microscopic VFF parameters and the macroscopic elastic constants [43].

2.2.3 Lattice Dynamics and Phonons

The elastic constants are at the center of the description of static lattice properties of crystals. Dynamic properties, i.e. vibrations, are typically classified by harmonic oscillations of the lattice called phonons. Knowledge of the so-called dispersion relation $\hbar\omega(q)$, where $\hbar\omega$ is the phonon energy at the phonon wave-vector q, permits the calculation of manifold quantities ranging from thermal conductivity to phonon-limited electron mobilities. The vibrational frequencies ω of the phonons are the eigenmodes of the dynamical matrix, which is closely related to the crystal energy: It is the Hessian (second derivative) of Eqn. (1) with respect to all atomic positions,

augmented with a wavevector-dependent phase factor that reflects the periodicity of the structure. The details to solve such a dynamical matrix are outlined in Ref. [37]. Zincblende lattices are periodic continuations of two-atomic unit cells.Since every atom has three spatial degrees of freedom, the size of the dynamical matrix for any wave-vector is consequently 6. The resulting six eigen-modes are classified as follows:

- Three modes are low-energy oscillations that have vanishing energy as $q \to 0$. The slope of $\hbar\omega(q\ 0)$ determines the speed of sound in the material. These acoustic modes dominate properties such as thermal conductivity.

- The three high-energy modes are called optical modes. They are characterized by out-of-phase oscillations of the two atomic Ga and As sublattices. In zincblende crystals these modes dominate the scattering between electrons and phonons.

For each type two out of three modes vibrate in directions perpendicular to the wave-vector whereas the vibration of the third mode is in the same direction. This classifies into transversal (T) and longitudinal (L) branches. The crystal hence exhibits TA(2), TO, LA (2) and LO phonons, where O stands for optical and A stands for Acoustic. This classification is illustrated in Figure 2.2.

2.3 Genetic Algorithm

This section is a brief introduction to genetic algorithms and parallel genetic algorithms. Readers who are familiar to these algorithms may skip this part.

2.3.1 Standard Genetic Algorithm

Genetic algorithms are adaptive heuristic search techniques which are inspired by the principles of evolution and natural selection. Due to this, they represent an intelligent exploitation of a random search within a defined search space to find solutions for a given problem. First pioneered by John Holland, they are widely used and experimented in the science and engineering area [45]. In the standard GA, candidate

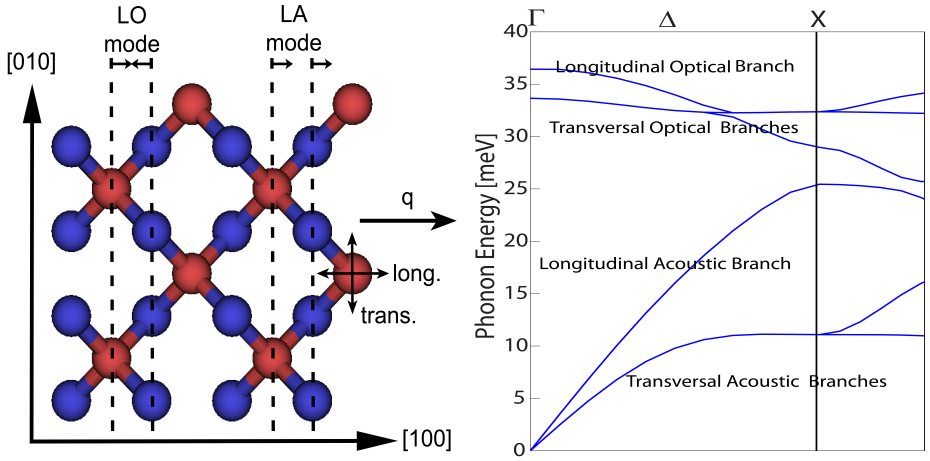

Fig. 2.2. Phonon modes in zincblende lattices. A) Sketch of the movement of atoms in the example of a phonon wavevector parallel to the [100] crystal direction. B) An example dispersion relation with labeling of the optical and acoustic branches. The horizontal axis denotes the wavevector and the vertical axis the phonon energy. The Greek labels (Γ, Δ and X) are symmetry points of the crystal.

solutions are encoded as fixed-length binary strings (or vectors), where the bits of each string are considered to be the genes of an individual chromosome and where the tuple of these individuals called a population. The initial solution population is usually chosen randomly. These chromosomes, which are candidate solutions, are allowed to evolve over a number of generations. In each generation, the fitness of each chromosome is evaluated. This is a measure of how well the chromosome optimizes the objective function. Subsequent generations are created through a process of selection, recombination, and mutation. A chromosome fitness measure is used to probabilistically select which individuals will recombine. Recombination (crossover) operators merge the information contained within pairs of selected "parents" by placing random subsets of the information from both parents into their respective positions in a member of the subsequent generation, or a child. Due to random factors involved in producing "children" chromosomes, the children may or may not have higher fitness values than their parents. Nevertheless, because of the selective pressure applied

through a number of generations, the overall trend is towards a generation of higher fitness chromosomes. Selection is a costly process which is usually based on the chromosomes' fitness. Mutations are used to help preserving diversity in the population. Mutations introduce random changes into the chromosomes [46]. The main objective of the mutation is exploring the solution space versus the main objective of the crossover which is exploiting. The pseudo-code of the canonical GA is depicted in Figure 2.3.

```
Begin
   INITIALIZE population with random candidate solutions;
   EVALUATE each candidate;
   REPEAT UNTIL (TERMINATION CANDITION is satisfied ) DO
      1 SELECT parents;
      2 RECOMBINE pairs of parents;
      3 MUTATE the resulting offspring;
      4 EVALUATE new candidate;
      5 SELECT individuals for the next generation;
   DO
END
```

Fig. 2.3. Pseudo-code of the canonical genetic algorithm.

2.3.2 Parallel Genetic Algorithm

The parallel GA (PGA) is an algorithm used to accelerate computation using parallel and distributed computing. The PGA has the potential capability to solve problems faster than simple GA. However, it was mainly used by the software. Due to the nature of the problem and available cluster the parallel GA that is used in this project is coarse-grain (or island) [47].

The island or coarse-grained model divides the population in multiple sub-populations. The sub-populations evolve independently from each other for a certain number of generations (isolation time). After the isolation time, a number of individuals are distributed among sub-populations by the migration operator. This model also is more capable of finding global optimum chromosome in complex fitness landscapes. Petty and Leuzestudieda coarse-grain PGA in Ref. [48]. In this research we employed a coarse-grain PGA. For more detail about other PGA methods please read [7]

2.4 Phonon Dispersion for GaAs

To fit the model with experimental data using a PGA, a sophisticated multi-objective fitness function is defined to help the optimization method to find the optimum solution with minimum fitness evaluations. This section first describes how the fitness function is developed and what the used parameters were for PGA. Then the three-step parameter fitting for the objectives is explained: acoustic branches, whole dispersion relation, and at the end whole dispersion and elastic constants. The extracted parameters are displayed in [7,14]. This problem has an 8-dimensional vector of real numbers as a gene, where each of the genes is related to the model's inputs. The order of genes in the chromosomes is as follows:

α	β_{Ga}	β_{As}	γ_{Ga}	γ_{As}	δ_{Ga}	δ_{As}	v(Nu)

2.4.1 Fitness Function and PGA Parameters

We minimize the mean-squared-error between the model's simulation outputs and experimental data. Three objectives needed to be fit. The first objective (e.g. F1) was the dispersion relation. By this objective an attempt was made to match the acoustic (and optical) branches with experimental data points . This is shown in the following equation:

$F_1 = \Sigma_{allkpoints} range(VFF_{output} - Exp_{value})^2$

At high symmetry points (Γ,X and L in Figs. 2.6) weight was added to push the evolution process to converge to a better match for these points, as they are particularly important, thus counted as a second objective (e.g. F2). The third objective (e.g. F3), which was added to the third part of the fitting process, was the mean-square-error between calculated and target elastic constants . The total cost value is then composed as follows:

$CostValue = W_1F_1 + W_2F_2 + W_3F_3$,

where W_i is the weight of every distance function. These weights were assigned manually based on tolerable deviations from target values. To handle the constraints over the range of input parameters, a very large cost value (10^6) was applied when the inputs were out of the acceptable range. The acceptable ranges were $0 < \alpha, \nu < 100$ and -100 to +100 for all other parameters.

In the semi-empirical methods, experimental data is typically used as target values. In this work, we used experimental data from [49]. For some experimental points like Γ, this data was extended with generally accepted values (e.g. zero). Experimental data has some error which depends on the precision of the measurement devices. To include this issue in the minimizing process a "range" fitness function was used: if the model's outputs were within a certain range of the target values they would get the same high positive weight in fitness value. If the model's outputs are out of the range then the square of the distance to the target range enters the fitness value (see Fig. 2.4.1). Other parameters for the island PGA are a population size of 2000, 100 for number of generations, crossover rate 0.7, crossover type was one-point, mutation rate 0.05, migration rate 1 per each island. The PGA ran on an 8 core cluster. Due to the random nature of evolutionary algorithms, we ran each optimization process about 10 times and reported the best achieved results.The model was developed in NEMO5 using C++ [18]. Each optimization process takes about 10 hours to be done on an eight 2.5 GHz cores shared memory cluster with 32 GB memory and 10 GibE (Gigabit Ethernet) connection.

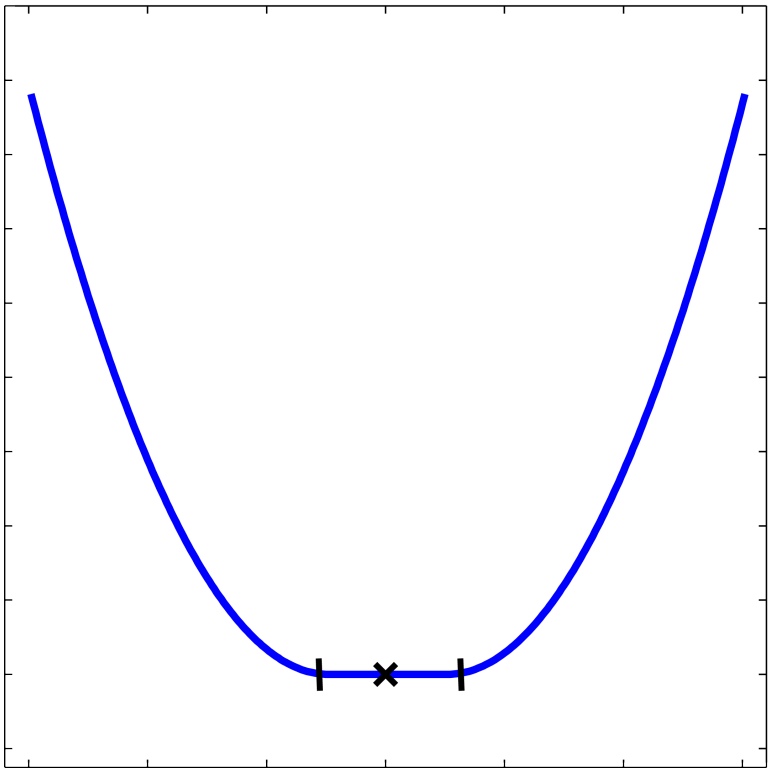

Fig. 2.4. Fitness evaluation with variation around the target value.In this minimization problem the fitness value increases with the distance between a target range and the model's output.

2.4.2 Predictiveness of the model

The strength of any model is given by its ability to predict experimental data not included in the fitting process. A first indication of the model's predictive abilities is given by phonon frequencies, which were not included in the fitting process. Figure 2.4.2 compares the model versus experiment for the phonon dispersion path $L - X - W - L$. Good agreement is observed, suggesting a solid physical foundation of the chosen approach.

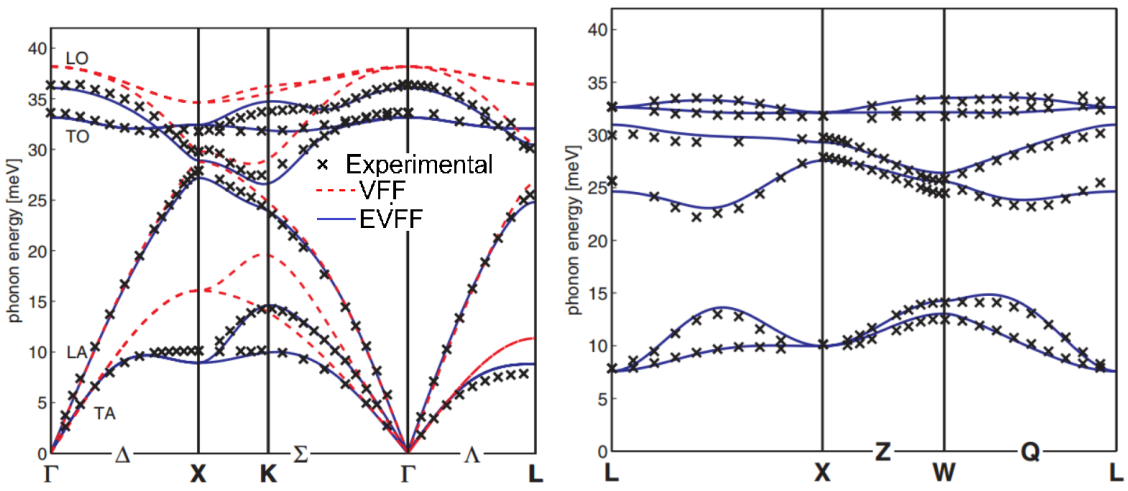

Fig. 2.5. (Left) Phonon dispersion of gallium arsenide. Solid line: 8-parameter set (see Ref. [14]) fitted against the dispersion, the sound velocities, and the elastic constants. Dashed line: Keating model with $\alpha = 41.19N/m, \beta = 8.94N/m$. Crosses are experimental data from [49]. (Right)Phonon dispersion of GaAs along the path LX W L. Solid line: 8-parameter set (see Ref. [14]). Crosses are experimental data from [49].

2.4.3 Extending the model for other materials: InAs

The same approach as explained for GaAs, is used for a EVFF model for InAs. The lack of experimental data [50] made the process less accurate and as it is depicted in the figure there is no experimental/*ab-initio* data point at L symmetry point. However, the overall phonon dispersion looks to be fitted well (Figure 2.4.3). Further information is provided in Ref [12].

2.5 An application of the EVFF model in thermoelectricity

Using the EVFF model for GaAs and InAs, the phonon spectra in zincblende InAs, GaAs and their ternary alloy nanowires (NWs) are computed using an enhanced valence force field (EVFF) model (Figure 2.5) . The physical and thermal properties of these nanowires such as sound velocity (Figure 2.5-a), elastic constant, specific heat

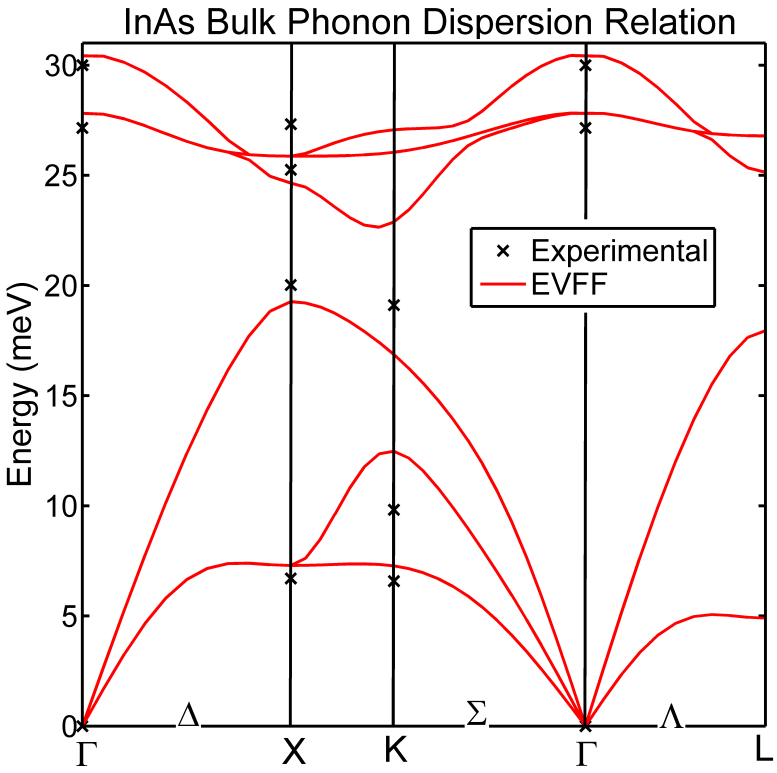

Fig. 2.6. Phonon spectrum of InAs in bulk are depicted. The cross points are experimental values from [50] and the solid lines are EVFFs output with listed parameters in [12].

(C_v)(Figure 2.5-b), phonon density of states, phonon modes, and the ballistic thermal conductance are explored. The calculated transverse and longitudinal sound velocities along $< 100 >$ direction in these NWs are 25% and 20% smaller compared to the bulk velocities, respectively. These velocities along $< 111 >$ direction are about two times smaller than bulk values. The C_v for NWs are about twice as large as the bulk values due to higher surface to volume ratio (SVR) and strong phonon confinement in the nanostructures. The temperature dependent C_v for InAs and GaAs nanowires show a cross-over at 180K and 155K along $< 100 >$ and $< 111 >$ directions respectively. It happens due to higher phonon density in InAs nanowires at lower temperatures. With the phonon spectra and Landauer's model the ballistic thermal conductance is

reported for these III-V NWs. The results in this work could help to engineer the thermal behavior of III-V NWs. Further detail is provided in Ref. [12].

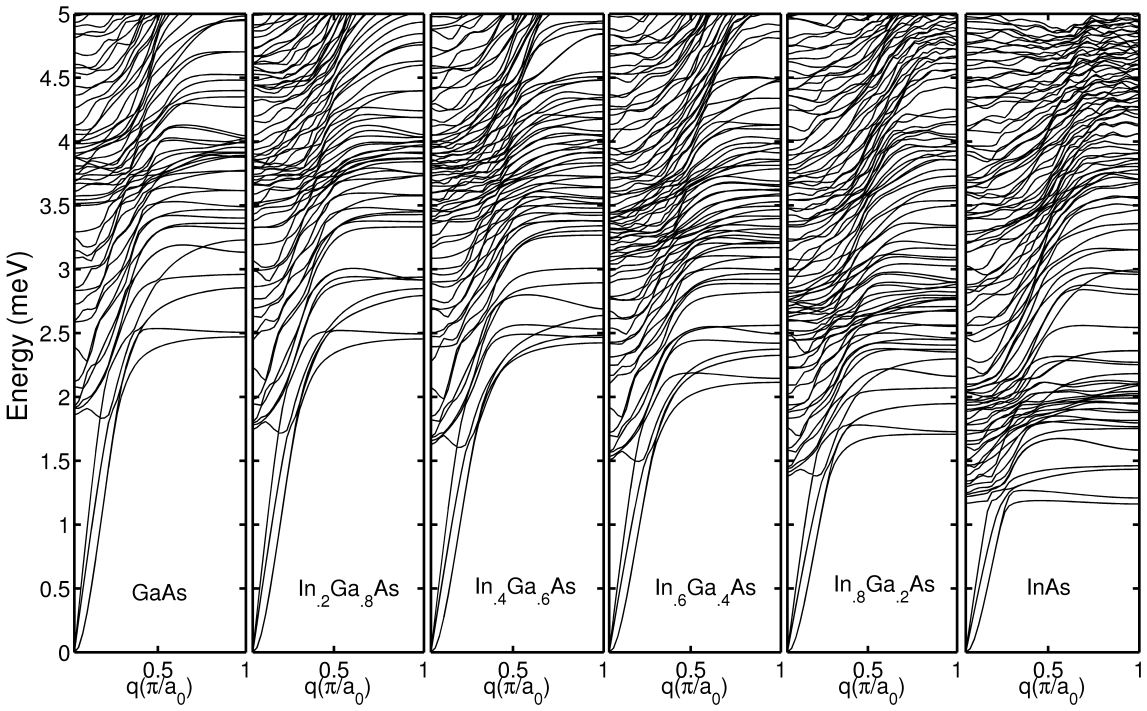

Fig. 2.7. Low energy branches of phonon band structure for different alloys of InGaAs. The phonon wave vectors, q, are all in the nanowire periodic direction which is <100>.

2.6 Conclusion

In this chapter the process for a semi-empirecal model development using parallel genetic algorithm was shown. The semi-empirical model for semiconductor lattice properties was achieved by fitting its constant parameters to experiment. The model was verified to be predictive for other frequencies and we showed it is extendable for other materials. As an application of the model we used it for predicting theroelectric properties of a square cross section nanowire. The process for developing complex

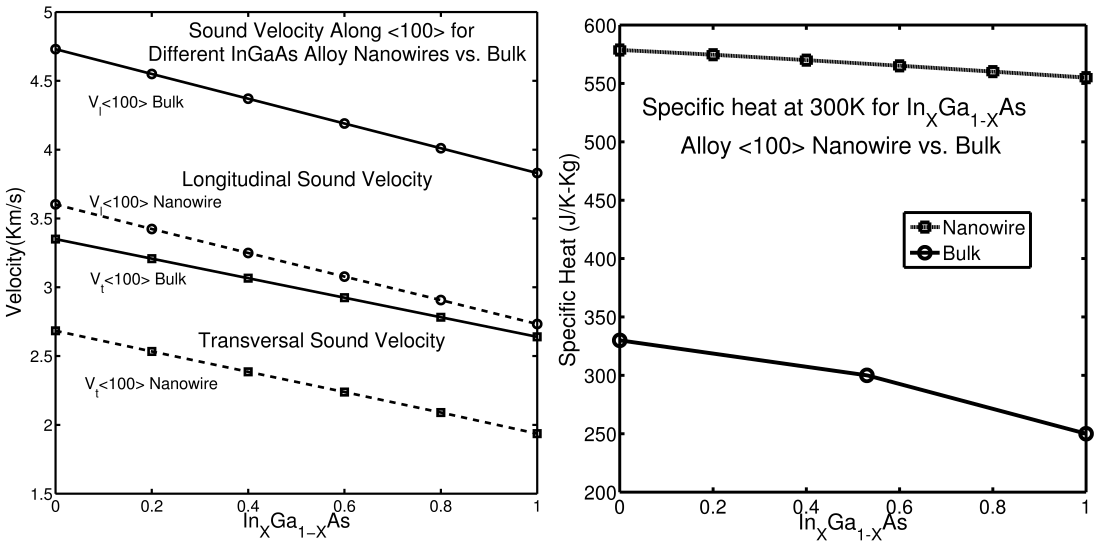

Fig. 2.8. (left)Longitudinal (V_l) and transversal (V_t) Sound Velocity in <100> InGaAs alloy NWs with freely vibrating transverse boundaries. As a reference the bulk longitudinal and transversal sound velocities are shown along the <100> direction [51]. (right) Variation of the specific heat (C_v) at 300°K in InGaAs alloy nanowires with fraction of In and Ga. As a reference the specific heat for bulk InGaAs [28] is shown by a black line with holes. NWs have larger C_v than bulk due to larger surface to volume ratio.

semi-empirecal models are the same. We used the same method to fit parameters for straines Si [13] and MoS2, GaSb and InSb (which we did not publish).

3. CMOS TRANSISTOR SCALING ROADMAP

3.1 Introduction

Due to the significant resources and investments required to develop the next generation of complementary-MOSFET (CMOS) technology nodes, it is necessary to identify clear goals and put collective efforts towards developing new equipment and technologies. The semiconductor roadmap represents a consensus among industry leaders and gives projected needs based on past trends. The International Technology Roadmap for Semiconductors (ITRS) [9] is the standard accepted roadmap with clear targets to identify future requirements of CMOS technology. The target values for different MOSFETs in ITRS device tables used to be extracted with compact models using MASTAR CAD tool (Fig. 3.1). These compact models get SS and DIBL form recent devices [17] as inputs which can cause a huge discrepancy in setting targets for future nodes as it is depicted in Fig 3.1-b. As device structures are getting extremely small (channel length below 20 nm and body thicknesses less than 8 nm) quantum mechanical effects become an issue to the picture. Additionally, the material transport/confinement orientation can change the electronic properties of the channel material. This means that the electrical properties of a material are not inherited from the bulk material properties, but are also geometry dependent. In Figs. 3.1-a and 3.1-b, the effects of confinement are depicted, such that when the body thickness of Si goes from 5 nm to 2 nm, band structure and density of state change. These changes will affect device performance. In addition, in Figs. 3.1-c,-d and -e, the transport orientation and confinement direction can drastically vary the device characteristics. As it is shown in Table I, channel length will go below 10 nm after 2019. This will lead to high source-to-drain tunneling current in OFF-state. To capture tunneling effects accurately, a sophisticated CAD tool will be required. Due to aforementioned reasons,

devices characteristics for ITRS-PIDS tables are calculated with a quantum transport tool which captures all of the essential physics accurately. The scaling trend and ION for double gate devices are depicted in the Fig. 3.1. The calculated ON-current trend with quantum transport tool and compact model based calculations show a drastic difference. This trend is happening due to three major effects: i)Short channel effects (SCEs), ii) S/D tunneling and iii) Quantum confinement. All of these effects will be discussed in the following sections. In section 3.2 we explain the method to calculate the devices characteristics. In section 3.3 the results are provided. In Section 3.4 we will discuss different phenomena in these devices. Finally, Section 3.5 will conclude the work.

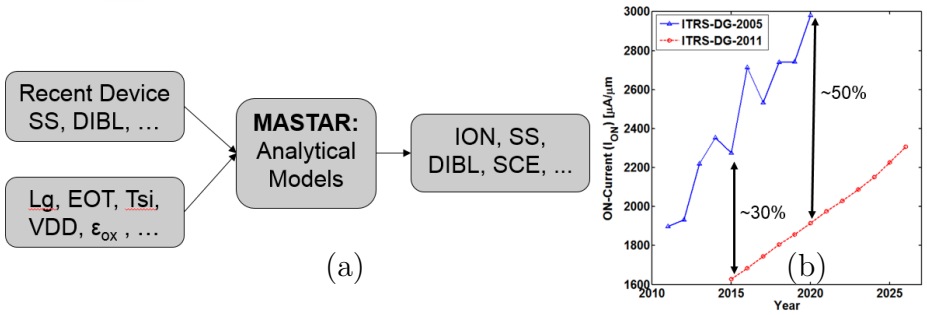

Fig. 3.1. a) A flow chart which shows MASTAR's inputs and outputs. MASTAR gets SS and DIBL form recent devices [17] due to that the calculations are very dependent to the recent fabricated MOSFETs. It can have huge discrepancy over years. Specially, for far future it cannot capture the critical physics like quantum confinement and source to drain tunneling.

3.2 Methodology

An atomistic quantum transport simulator based on Quantum Transmitting Boundary Method (QTBM) method with the nearest-neighbor $sp^3d^5s^*$ tight-binding (TB) [52] is used to calculate intrinsic device characteristics in the ballistic regime. The

Table 3.1.
ITRS-PIDS table for High-Performance(HP) Devices. Will be available online at ITRS website by Jan 2013.

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
L_G	20	18	16.7	15.2	13.9	12.7	11.6	10.6	9.7	8.8	8.0	7.3	6.7	6.1	5.6	5.1
VDD	0.86	0.85	0.83	0.81	0.80	0.78	0.77	0.75	0.74	0.72	0.71	0.69	0.68	0.66	0.65	0.64
EOT	0.80	0.77	0.73	0.70	0.67	0.64	0.61	0.59	0.56	0.54	0.51	0.49	0.47	0.45	0.43	0.41
K	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0	16.5	17.0	17.5	18.0	18.5	19.0	19.5	20.0
T_{OX}	2.58	2.56	2.54	2.52	2.50	2.47	2.44	2.41	2.37	2.34	2.30	2.26	2.22	2.18	2.14	2.10

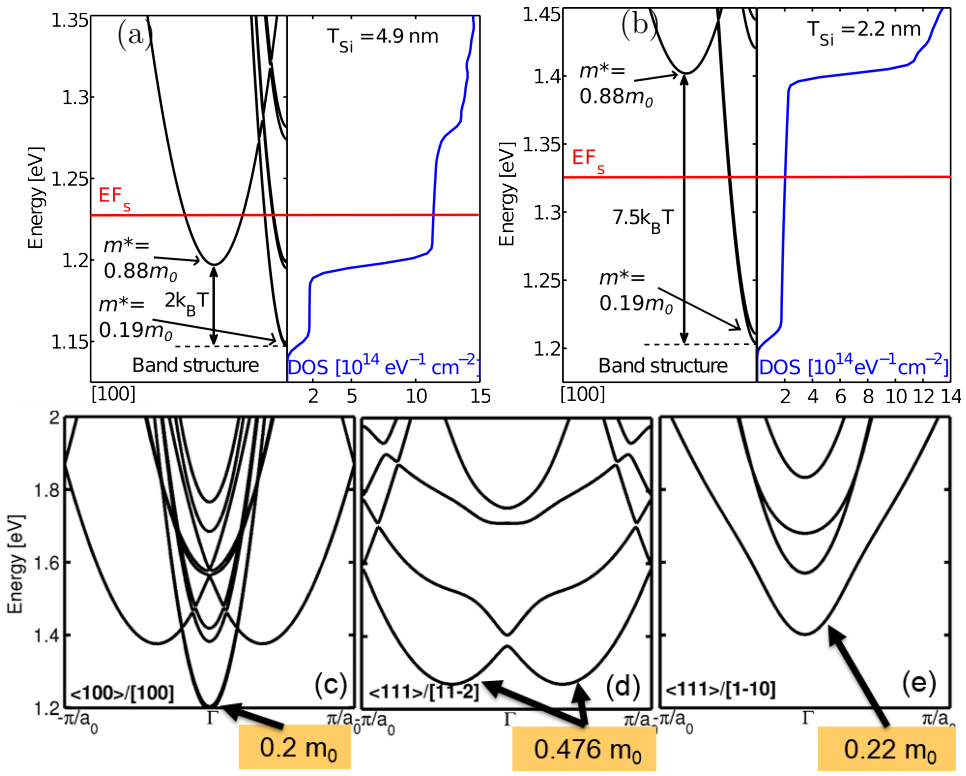

Fig. 3.2. a) Band structure and DOS for T_{si}=4.9 nm. b)Band structure and DOS for T_{si}=2.2 nm. c,d,e) Self-consistant band structure calculated for UTB with 2.2 nm thickness with different transport and confinement orientations. Orientation can change band structure and respectively carrier mass in a device.

validity of the TB band structure is confirmed via comparison against first principle electronic structure calculations in ultra-thin body (UTB) silicon for different UTB body thickness and the result for t_{si}=2.2 nm is depicted in Fig 3.2. These results show agreement between first-principle calculations and tight-binding's output. To capture the scattering effect backscattering model, the Lundstrom model, is used. After calculation of the device ballistic characteristics using backscattering model [54–57] is applied using the following equations [54]:

$$T = \frac{\lambda_0}{(\lambda_0 + 2\ell_{kT})} \tag{3.1}$$

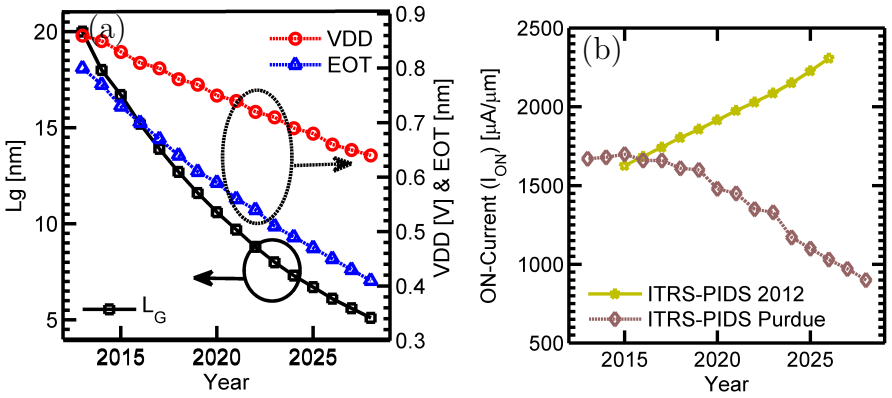

Fig. 3.3. a) Projecting MOSFET scaling geometry such as L_{Eff}, VDD and EOT for next 15 years. b) Analytical model based (MASTAR) calculation shows constant performance improvement with Lg shrinking, however, atomistic modeling shows exactly opposite trend.

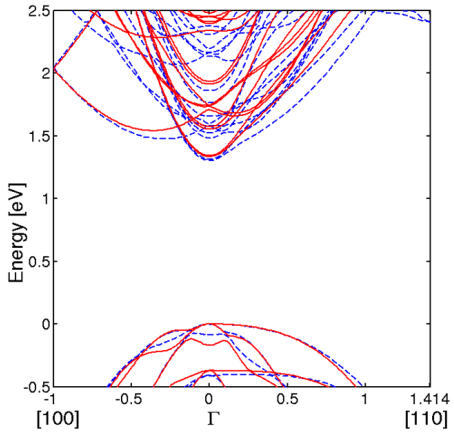

Fig. 3.4. Comparison between DFT/LDA (abinit [53]) band structure (blue dashed lines) and tight binding (red solid lines) for a 2.2 nm UTB. The LDA bandgap has been corrected by scissor operator. Maximum deviation for the lowest conduction and highest valence bands is 251 meV at K point (at the end of [110]) but around the Γ point it matches very well. Similar agreement is obtained for an even smaller body thickness of 1.1 nm.

$$HighVDS : I_{Scattering} = \frac{T}{2-T} I_{Ballistic} \tag{3.2}$$

$$LowVDS : I_{Scattering} = T \times I_{Ballistic} \tag{3.3}$$

where T is the transmission coefficient and ℓ_{kT} is effective ON-current channel length, which is the distance between top of the barrier and one k_BT lower than that [56, 58, 59]. This value is calculated from ballistic potential profile for each bias-point. The mean free path (λ) value is required to include scattering effect by back scattering model. This value is extracted from the experimental reported values [60–63] for different UTB body thickness and charge under the gate. The approach that we used for calculation of mean free path as below [**?**]:

$$\lambda = \frac{VDS}{v_{inj}} \mu \tag{3.4}$$

where VDS is very low, i.e. 5 mV and v_{inj} is related to the equivalent charge under the gate. μ_0 is mobility, which is dependent to the charge under the gate and device body thickness.

3.2.1 Maximum tolerable series resistance calculation for HP devices

At the last post-processing phase, series resistance effect is assumed as maximum 1/3 reduction over ON-current. Based on that assumption the maximum tolerable series resistance is extracted. To calculate the maximum tolerable series resistance, we calculate the second ballistic $I_D - V_G$ with VDS-0.1 V. After calculation of I-V with scattering we calculate the RSD as below:

3.3 Simulations and results

Double-gate device structure is depicted in Fig. 3.3. The source and drain length is fixed between 10-15 nm for all cases. The ballistic characteristics of the channel transport/confinement orientation are set to <100>/[100] and gate length, physical

$$I_{ds} = I_{in} - \Delta V_{gs} \frac{dI_{in}}{dVgs}|_{V_{DD}} - \Delta V_{ds} \frac{dI_{in}}{dVds}|_{V_{DD}}$$

$$\Delta V_{gs} = I_{ds} R_s \; and \; \Delta V_{ds} = I_{ds}(R_s + R_d) \; and \; R_s = R_d$$

$$\rightarrow \Delta V_{ds} = 2\Delta V_{gs}$$

$$I_{ds} = I_{in} - \Delta V_{gs} \left(\frac{dI_{in}}{dVgs}|_{V_{DD}} + 2 \frac{dI_{in}}{dVds}|_{V_{DD}} \right)$$

In HP: $I_{ds} = 0.67 I_{in} \rightarrow 0.67\, I_{in} = I_{in} - 0.67 I_{in} R_s \left(\frac{dI_{in}}{dVgs}|_{V_{DD}} + 2 \frac{dI_{in}}{dVds}|_{V_{DD}} \right)$

$$\rightarrow R_s = \frac{0.5}{\left(\frac{dI_{in}}{dVgs}|_{V_{DD}} + 2 \frac{dI_{in}}{dVds}|_{V_{DD}} \right)}$$

We need to have one other Ids at :

$$Vgs = V_{DD} \; and \; V_{ds} = V_{DD} - \Delta V$$

$$\Delta V = very\ small\ value\ lie\ 0.1V$$

oxide thickness (T_{OX}), oxide's dielectric constant, EOT and VDD are set the same as ITRS Table 3.1) after that the ballistic properties of each device for 2 different VDS, i.e. VDD and VDD-0.1 V are calculated. By post-processing scattering and parasitic resistances effects are applied. The achieved results are shown in Figs. 3.1-b and 3.3. By further geometry scaling the device performance metrics will degrade. The downward trend for ION is primarily due to short channel effects and S/D tunneling. Quantum confinement has different impacts on the device performance. We discuss these effects in the following sub-sections.

3.4 Discussion

3.4.1 Short Channel Effects (SCEs)

In the ideal MOSFET, top-of-the-barrier is controlled only by gate. However, with channel length reduction, drain could affect the barrier too which lead to some negative effects including Vth-rolling. Short channel effects are related to the strength

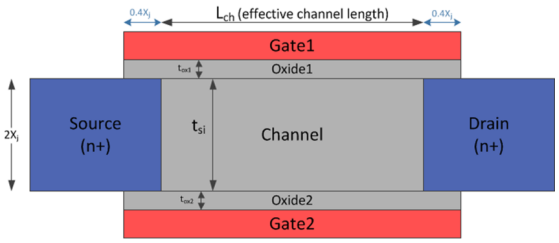

Fig. 3.5. schematic of the DG device structure. From, ITRS table (Table I), the L_{Eff}, t_{si} (which is the equivalent of channel width in tall FinFETs), EOT and VDD varies for each case. The channel material is Si <100>/[100] with no strain. I_{OFF} is set to $100nA/\mu m$ for all high performance devices. L_{Eff} is assumed 80% of L_G and body thickness, t_{si}, is set to 40% of L_{Eff}.

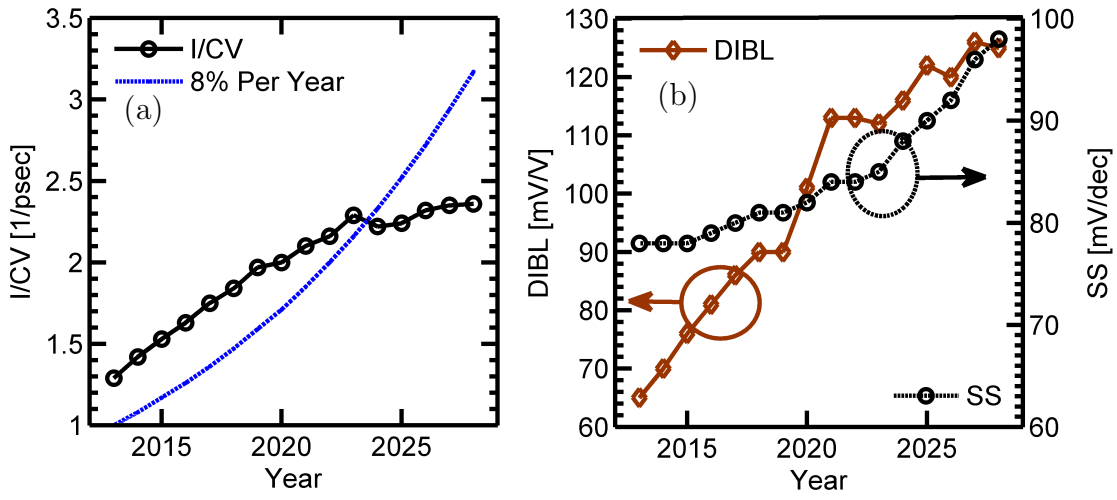

Fig. 3.6. a) Device speeds with calculated current by quantum transport TCAD increase by ideal (8% increase per year) up to 2023. b) Ballistic DIBL and SS for devices with L_{Eff} reduction increase (degrade).

of drain control over the barrier. The shorter the distance between the drain and the top-of-the-barrier the stronger SCEs. It is depicted in a simple cartoon in Fig. 3.4.1. By scaling in device the drain control increases. This side effect can be seen in the DIBL increase(Fig. 3.3).

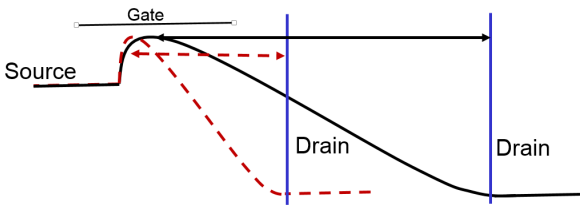

Fig. 3.7. A simple schematic to show the relation between channel length reduction and drain control on the top-of-the-barrier. SCEs increase with drain control on top-of-the-barrier. Shortening channel length, the distance between drain and top of the barrier reduces which causes higher SCEs and drops the device performance.

3.4.2 Source-to-drain Tunneling

By scaling channel length the width of potential barrier reduces which makes it possible for carriers to penetrate through the barrier. This effect in the OFF-state will increase the leakage current and make it harder to turn off the device. As it is depicted in Fig 3.8-a,-b and -c, tunneling current percentage over the whole OFF-current increases with gate scaling. With S/D tunneling, gate voltage needs to be lower to make the barrier taller in order to keep the OFF-current at 100 nA/um, leading to higher SS (Fig. 3.8-d). Higher SS degrades device performance. This issue will be discussed further in the next chapter.

3.4.3 Quantum Confinement

Quantum confinement (QC) varies band structure and density of states (DOS). Thinning Si $<$ 100 $>$ /[100] reduces m^* and DOS. Current can be calculated as $I = Q_{inj}v_{inj}$ where Q_{inj} is charge on top-of-the-barrier and v_{inj} is the average velocity of carriers on top-of-the-barrier ($v_{inj} \propto 1/\sqrt{m^*}$). Quantum confinement can help to reduce the carrier effective mass, but this effective mass reduction results in higher S/D tunneling ratio. As it is depicted in Fig. 3.4.3, with scaling v_{inj} raises but Q_{inj}

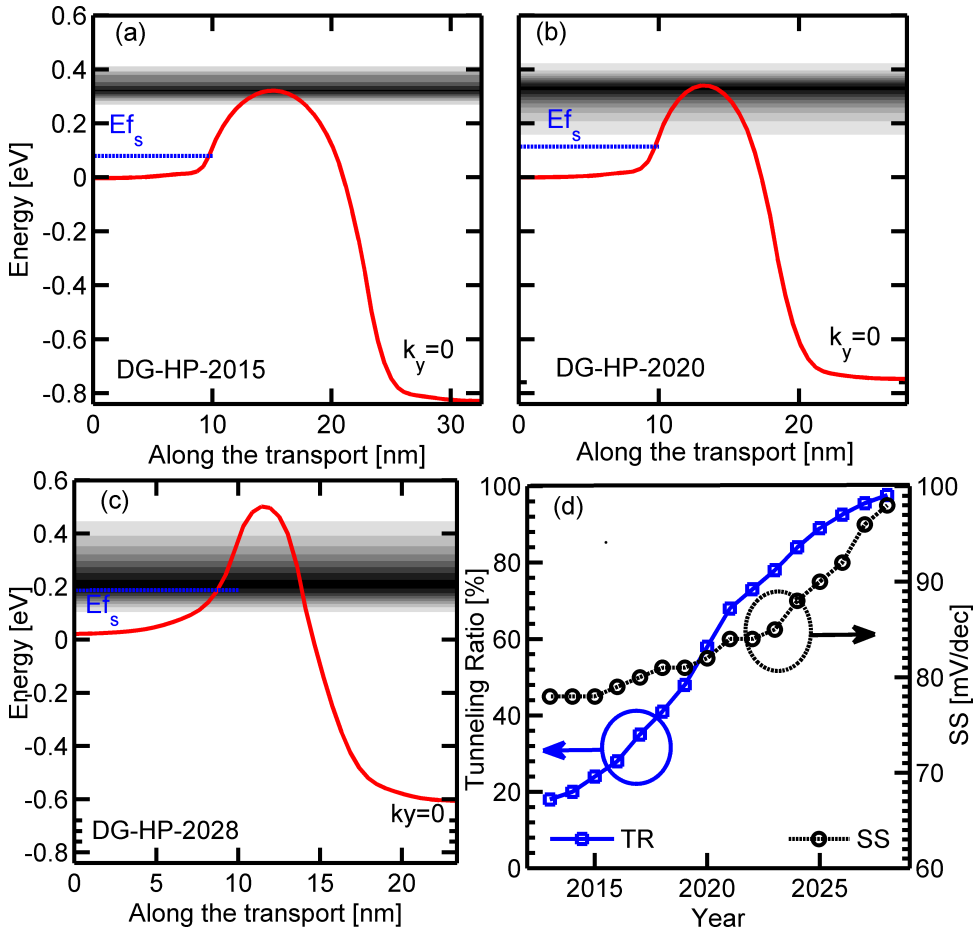

Fig. 3.8. a,b,c) Current density (JE) below top-of-the-barrier (tunneling current) increases due to shortening in the effective channel length, from 24% in 2015 to 58% at 2020 and 98% in 2028. d)S/D tunneling increases SS and degrade device performance with scaling.

drops. As it is shown in Fig. 3.1, thinning in Si body thickness and reduction in the transport effective mass results in v_{inj} increase ($t_{Si} \downarrow \rightarrow m^* \downarrow \rightarrow v_{inj} \uparrow$). With scaling, confinement reduces m^*, however, this raises tunneling and increases SS which leads to larger Vth. In addition to that, by scaling VDD and increasing in Vth, Q_{inj} drops more than the increase in v_{inj} ($Q_{inj} \propto C_G(VGS - Vth)$).

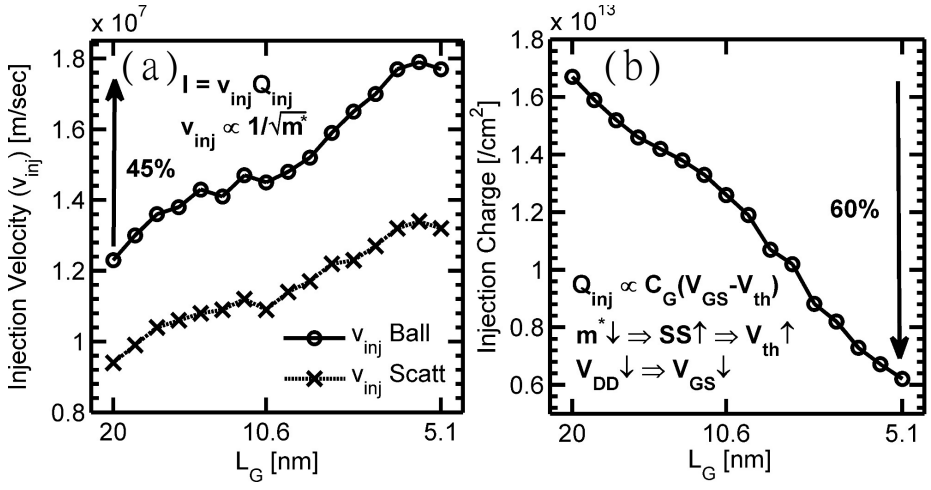

Fig. 3.9. The charge reduction is higher than v_{inj} increase. With confinement along [100] in Si, the transport mass reduces (Fig. 3.3) but tunneling also increases which leads to higher tunneling and SS. Degraded SS increases the Vth and reduces the charge at ON-state (Q_{inj} $C(VGS-Vth)$.

3.5 Conclusion

In this chapter we discussed the need for rewriting ITRS device tables. The methodology to capture essential physics are discussed along with achieved results and discussion about the scaling effects on the device performance metrics. The major challenges for scaling will be short channel effects and S/D tunneling. In the next chapter we analyze S/D tunneling in detail and investigate one possible solution to reduce this effect in ultrascaled MOSFETs.

4. DESIGN GUIDELINES FOR TUNNELING DOMINANT REGIME MOSFETS (LCH<12 NM)

Over the last four decades Si based technology has been continuously scaled down in device dimensions, which are now approaching sub-22 nm channel lengths [4, 9, 64]. Until this point, using light m^* materials as channel has been appealing due to their high injection velocities. However, it has been shown at 12 nm channel length InGaAs ($m^* = 0.07\, m_0$) provides very similar SS, drain induced barrier lowering (DIBL) and ON-current compared to strained-Si for triple-gate MOSFETs [23]. By scaling the channel length further below 12 nm, the effect of S/D tunneling becomes more important [6, 24]. This tunneling increases the leakage current and the sub-threshold swing (SS) [4, 6, 25, 26], counter productive to the goal of achieving low SS as a key factor in scaling. FinFET and nanowire (NW) structures are used to lower SS and NW MOSFETs have emerged as promising candidates for ultra-short devices [27]. Gate-all-around NWs offer the best electrostatic gate control over the channel [6, 27]. Previous theoretical studies have explored the effects of channel materials, orientation and strain on tunneling in NMOS devices [6, 24, 26, 28–30]. These studies showed the benefit of using heavier band minima mass materials over light materials; specifically Si as heavier mass material versus InGaAs as light mass material [6]. However, these studies did not estimate the range of optimum heavy m^* that may be needed to achieve both low SS and high ON-current for different channel lengths. Additionally, these studies did not investigate the effect of heavy m^* on quantum capacitance and electrostatics. Also PMOS devices were not discussed in aforementioned studies.

In this chapter, we attempt to answer the following questions for the first time: 1) What is the optimum effective mass range, independent of channel material, orientation, strain and transport type (n-type or p-type MOSFETs), for NW MOSFET with a given channel length? 2) How does a heavy mass channel material impact device

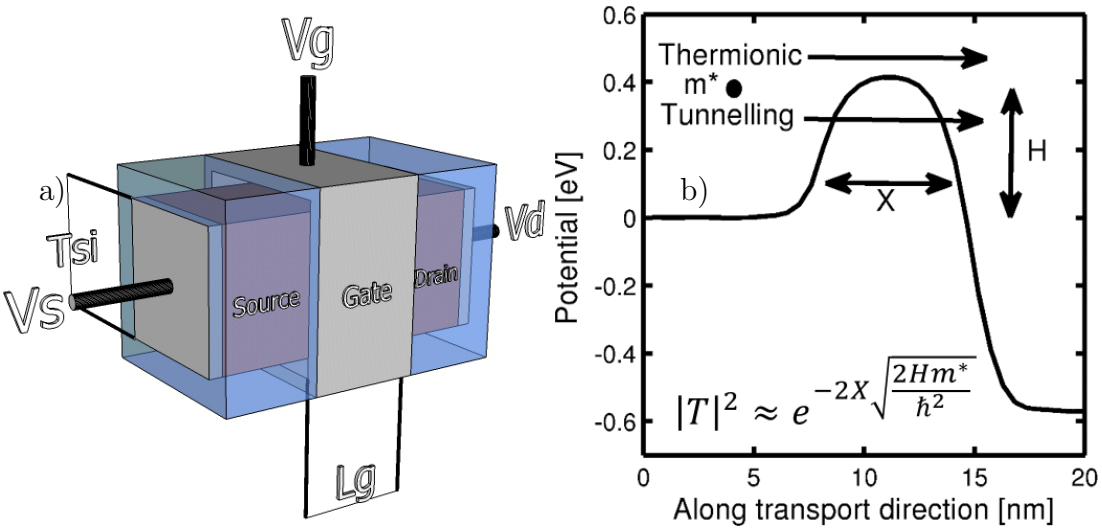

Fig. 4.1. a) A 5x5 nm^2 square cross-section nanowire with different gate length is used as target device. The source and drain length varies with the channel length due to the simulation convergence. Equivalent oxide thickness (EOT) is set to 0.4 nm. The current is normalized to the device width [23]. The OFF-current is set to 100 $nA/\mu m$ for all different channel lengths. b) A schematic to show the potential barrier profile in a short channel device. It explains two different effective factors in tunneling probability 1) Potential barrier profile (in simplified case: height (H) and width (X)) and 2) carriers' effective mass (m^*) [65]. Complex band structure and effective mass of the imaginary band is important to determine the decay rate. Two components of the OFF-current for short channel devices, which include thermionic and tunneling current are shown.

electrostatics? 3) Since m^* is a silicon geometry and strain dependent parameter, is it possible to engineer silicon in such a way to achieve optimum effective masses?

4.1 Methodology

To accurately capture S/D tunneling, an effective mass based real space Non-Equilibrium Green's Function (NEGF) transport simulator is used [52]. To address the impact of m^* we use a single band effective mass model with the m^* ranging from 0.07 m_0 to 2.0 m_0 for a nanowire (Fig. 1) with six different channel lengths

(L_{Eff} = 12, 10, 8, 6, 4 and 3 nm). In tunneling, the imaginary bands within the band gap of channel material are important [29]. The imaginary bands are by default modeled as parabolic with effective mass the same as real conduction band (CB), or valence band (VB) in effective mass model. Quantum confinement in Si increases the band gap (Eg) from its bulk value, i.e. 1.12 eV [23]. At such a large E_g (> 1 eV), it is possible to neglect the effect of VB (CB) in transport calculations for NMOS (PMOS) [24]. The carriers will be tunneling through the barrier within a few hundreds of meV below top-of-the-barrier (Figs. 2 and 3-c). For energies close to the conduction band (shaded area in Fig.2), effective mass approximates the complex band structure accurately when band gap is large enough ($E_g > 1$ eV). This means that for materials like Si or GaAs ($E_g > 1$ eV), effective mass approximation holds well [24]. Furthermore, quantum confinement can cause a significant increase in the band gap for light m^* (low bulk E_g) materials like InAs [66]. Given the ultra-scaled nanowire dimensions, semiconductor materials of research interest like III-Vs and Si are within the purview of this work. To provide the general trends independent of the complexity of band structure for different materials under complex device structures we therefore use a single band effective mass model. More complex methods, such as full-band modeling are required to obtain a quantitatively accurate understanding of specific devices with different materials and geometry specifications [29, 52]. For example, for low band gap materials, tight-binding should be used to capture band to band tunneling accurately [29].

Electrostatics have a direct impact on the potential barrier profile and SS. Stronger electrostatics lead to a wider barrier, which reduces the tunneling ratio and lowers the SS. To achieve the best electrostatics and reduce its effects on SS, a gate-all-around nanowire with 5x5 nm^2 square cross-section with very thin equivalent oxide thickness (EOT = 0.4 nm) is used as the target device (Fig. 1-a) [23]. This device cross-section size is kept the same as ref. [23]. Several approximations are introduced to point to the essential device physics and to avoid secondary effects. Effects due to multiple sub-bands in Si are avoided by setting the transverse masses to 0.1 m_0

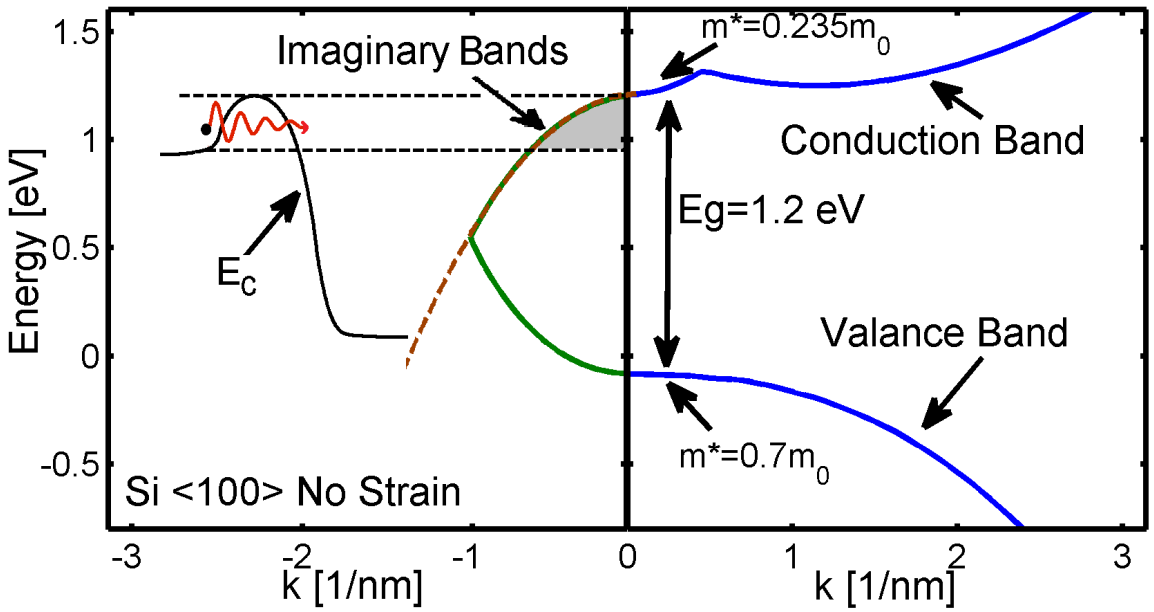

Fig. 4.2. The highest VB and lowest CB for Si <100> NW with no strain with 5x5 nm^2 square cross section is shown. Imaginary band connects the bottom of CB to the top of the VB is depicted in the left side of the image. The complex band structure of Si NW is calculated using $sp^3d^5s^*$ tight-binding (solid lines) [18] and effective mass (dashed lines). In a few hundreds of meV below the bottom of CB, the imaginary band is parabolic similar to the CB. Therefore, there will not be any significant interference from VB imaginary bands in tunneling.

and degeneracy factor (g_v) to 1. Higher g_v and multiple sub-bands can move down Fermi-level which will change the carriers' mass and consequently affect the tunneling ratio. The dielectric constant of bulk Si has been used for the channel.

4.2 Results and Discussion

Fig. 3-a depicts the effects of effective mass and channel length on SS. Heavy m^* have better SS values for any given L_{Eff}. At $L_{Eff} = 12$ nm all m^* have nearly ideal SS. However, as the channel length shrinks, SS degrades. This degradation is more pronounced for lighter m^*. SS is dependent on electrostatic (strong gate control), which is related to oxide thickness and MOSFET design. A good aspect ratio of cross-section size to channel length is required to achieve strong electrostatics in nanowires. Due to the fixed cross-section size in this simulation set, the gate control weakens by reduction of L_{Eff}. EOT reduction is a way to improve the electrostatics. Fig. 3-b examines the effects of m^* and L_{Eff} in the ideal case of EOT = 0 nm. An increase in the effective mass does not improve the device performance above $m^* := 0.7\ m_0$. The tunneling current in the OFF-state increases the SS values. In Figs. 3-c, -d and -e the gate voltage has been adjusted to achieve $I_{OFF} = 100\,nA/\mu m$ for devices with similar geometry but different effective masses. Leakage current in the heavy effective mass material case ($m^* = 1.0\ m_0$) is primarily composed of thermionic current (tunneling ratio = 0.01% at Fig. 3-e). In the case of very light effective mass ($m^* = 0.07\ m_0$), tunneling current is dominant and the potential barrier is higher in order to control the leakage current (Fig. 3-c). It is still possible to achieve $SS \approx 100\,mV/dec$ for 3 nm device with $m^* > 0.7\,m_0$ (Fig. 3-a).

Figs. 4-a and -b depict DIBL versus m^* for L = 3, 4, 6 and 10 nm. DIBL increases for shorter L_{Eff} devices due to short channel effects. Here, the cross-section remains fixed at 5x5 nm^2, and as a result, the ratio of gate length to channel cross-section decreases by shrinking L_{Eff}, which leads to weaker gate control and higher DIBL value. However, for heavier mass for a given L_{Eff}, DIBL is better. This is due to

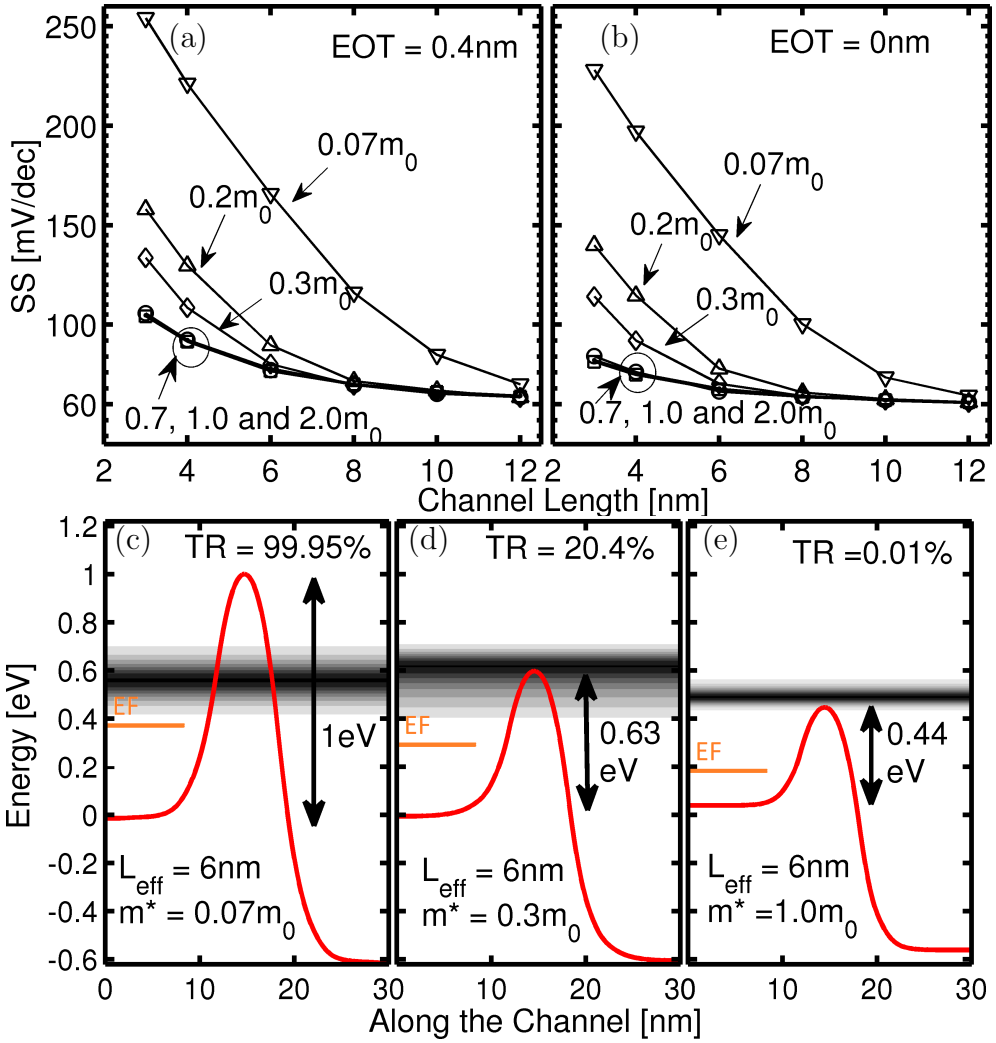

Fig. 4.3. a) SS for devices with EOT = 0.4 nm. b) SS for devices with EOT = 0 nm which leads to near-ideal electrostatics. c, d and e) current density (JEs) vs potential barrier. This shows the current above and below the potential barrier in the cases of $m* = 0.07$, 0.3 and $1.0\, m_0$ for $L_{Eff} = 6\, nm$. Tunneling rate (TR) is the ratio of tunneling current to total current. Heavier m^* materials have larger density of states which makes the distance between contacts Fermi-levels (EF) and bottom of conduction band (i.e. at 0 eV for the source contact) smaller in compare to light m^* materials for the same doping density. For lighter m^* the gate voltage needs to be more negative to produce higher barrier to keep the OFF-current fixed at the 100 $nA/\mu m$.

higher quantum capacitance (C_Q) and respectively higher gate capacitance ($C_G = C_{OX}||C_S$ where C_S is the semiconductor capacitance and $C_S \approx C_Q$). Meanwhile, the drain capacitance (C_D) does not vary much for the given channel length. C_G for 3 different m^* at the same channel length is depicted at Fig. 4-b. Moreover, the low density-of-states (DOS) in the light m^* materials leads to smaller C_G, fewer number of carriers in the channel and lower current for the ultra-scaled MOSFETs [**?**, 67]. This shortcoming of light m^* materials is known as DOS bottleneck. Recent research efforts trying to solve the DOS bottleneck issue still aimed at low transport m^* in ultra-thin-body III-V MOSFETs [66, 68].

C_G is proportional to density of states and higher C_G translates to a better gate control [69]. The calculated C_G value is higher in heavier mass material devices (Fig 4-b).

Figs. 4-c and -d show another important performance metric I_{ON}/I_{OFF} for different devices. In Fig. 4-c, I_{ON}/I_{OFF} ratio for light effective mass ($m^* = 0.07\, m_0$) and $m^* = 0.2\, m_0$ are very close at the length of 12 nm, which is consistent with [23]. Below 10 nm channel length, the device is in **tunneling dominated regime**. For light effective mass (high mobility) materials, high S/D tunneling in the OFF-state degrades the device performance including I_{ON}/I_{OFF} and SS. Fig. 4-d shows the impact of the effective mass for any given L_{Eff} on I_{ON}/I_{OFF} as a function of m^*. There is an optimum m^* for each channel length depicted in the inset of Fig. 4-d.

To achieve the optimal performance for different channel length, it is required to engineer the m^* from $0.2\, m_0$ to $1.0\, m_0$. Table I provides different m^* obtained for Si NW with 5x5 nm^2 square cross-section using the $sp^3d^5s^*$ tight-binding model in NEMO5 [18, 70]. 3 NMOS and 2 PMOS designs are listed. Our models predict the capability of Si to be scaled down to 3 nm.

The achieved optimum mass for each channel length is also valid for hole transport. These m^* are electron effective mass in the case of NMOS and hole effective mass for PMOS. As depicted in Table I, tensile (compressive) strain can degrade device performance for NMOS (PMOS). Using compressive (tensile) strain can help increase

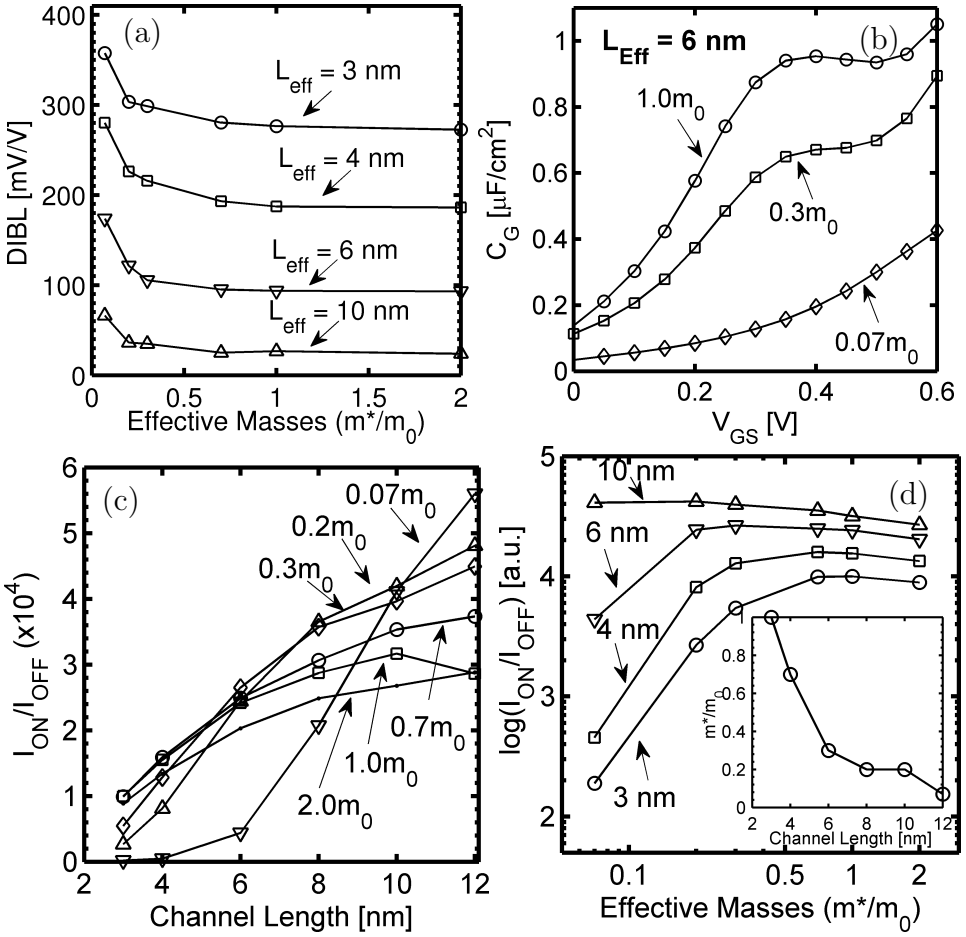

Fig. 4.4. a) DIBL for 4 different gate lengths. Heavier m^* leads to better DIBL due to higher C_G. b) C_G and in the cases of $m* = 0.07$, 0.3 and 1.0 m_0 for $L_{Eff} = 6$ nm, c) I_{ON}/I_{OFF} for different effective masses along different channel length. Effective masses are written as labels. d) I_{ON}/I_{OFF} as a function of effective mass for 4 different gate lengths. The inset shows the optimum m^* for each given L_{Eff}.

m^* for the carriers in NMOS (PMOS) devices. Confinement increases the band gap (Eg), which reduces the band-to-band tunneling ratio. Eg, non-parabolicity, multiple valleys, sub-bands and Fermi-level could change m^*. To get an accurate and quantitative estimation, aforementioned factors should be analyzed separately using a full-band transport tool.

Table 4.1.
Examples to engineer m^* of Si (bottom of CB for NMOS and top of VB for PMOS) for 5x5 nm^2 square cross-section NWs. For NMOS (PMOS), tensile (compressive) strain can reduce the m^*. For example for PMOS with compressive strain in 5th row the m^* is less than the case of without any strain in the 4th row.

Orientation (Si)	Stress [GPa]	Eg [eV]	m^*/m_0
$< 100 > (NMOS)$	0	1.28	0.235
$< 110 > (NMOS)$	-2.5	1.142	0.553
$< 100 > (NMOS)$	-1.5	1.182	0.942
$< 100 > (PMOS)$	0	1.280	0.7
$< 100 > (PMOS)$	-0.5	1.236	0.46

There are a few recent experiments for extremely short channel devices. Migite et. al., in [71, 72] fabricated a 3 nm channel length junction-less FET SOI. They reported the SS and DIBL for PFET (and NFET) as 189 mV/dec (238 mV/dec) and 520 mV/V (960 mV/V). Interestingly, for PFET (with heavier mass) SS and DIBL values are lower than NFET, which is consistent with the message in this paper. These values are higher than the predicted ones in this work, which is a result of weaker electrostatics in 2D (ultra-thin body) devices in comparison to NW [23].

4.3 Conclusion and Future Work

In this chapter, a general understanding of MOSFET performance in the source-to-drain **tunneling dominant regime** (sub-12 nm channel length) is provided. In sub-12 nm, for each L_{Eff} there is an optimum m^*, which provides highest I_{ON}/I_{OFF} and lowest SS. It has been shown that optimum heavy effective mass leads to a better C_q, which means improved DIBL and SS. For heavy effective mass, the ON-current is lower due to lower injection velocity. It is found that for nanowires, the optimum m^* lies between $0.2 - 1.0\, m_0$. Finally, a guideline to achieve the optimum effective mass for Si NMOS and PMOS at a given channel length is provided.

Other properties of Si MOSFETs such as: different transport orientations, stress, cross-section size and shape can be explored to obtain the optimum masses for different channel length. Other factors that affect S/D tunneling, include dielectric constant, high-k oxides, doping profile, band gap, non-parabolicity and scattering, which need further investigation. Also, using heavy mass materials might impact on device power consumption, f_{max} and gate leakage, which should be studied separately.

5. THEORETICAL ANALYSIS FOR OPTIMUM HIGH-K OXIDE FOR THE BEST PERFORMANCE OF ULTRA-SCALED DOUBLE-GATE MOSFETS

A widely used technique to mitigate the gate leakage in the ultra-scaled metal oxide semiconductor field effect transistors (MOSFETs) is the use of high-k dielectrics, which provide the same equivalent oxide thickness (EOT) as SiO_2. High-k materials provide thicker physical layers for the same EOT. However, using a thicker physical dielectric for the same EOT has a negative effect on the device performance due to the degradation of 2D electrostatics. In this letter, the effects of high-k oxides on double-gate (DG) MOSFET with the gate length under 20 nm are studied. We find that there is an optimum physical oxide thickness (T_{OX}) for each gate stack, including SiO_2 interface layer and one high-k material. For the same EOT, Al_2O_3 (k=9) over 3 Å SiO_2 provides the best performance, while for HfO_2 (k=20) and La_2O_3 (k=30) SiO_2 thicknesses should be 5 Å and 7 Å, respectively. The effects of using high-k oxides and gate stacks on the performance of ultra-scales MOSFETs are analyzed. While thin oxide thickness increases the gate leakage, the thick oxide layer reduces the gate control on the channel. Therefore, the physical thicknesses of gate stack should be optimized to achieve the best performance.

The scaling of transistors requires the thinning of SiO_2 gate oxide [9], which can induce significant gate tunneling below 1 nm oxide thicknesses. To mitigate the gate tunneling current in a thin oxide layer of ultra-scaled MOSFETs, high-k dielectrics are used [73, 74]. However, due to the thicker high-k gate oxides, performance drops have been observed in the ultra-scaled MOSFETs with $k > 30$ [75–80]. Thicker T_{OX} and larger k values worsen the short channel effects even if the EOT is kept the same [75]. This happens due to the effect of lateral field in the oxide, which is more pronounced in higher k materials [75, 76], and fringing capacitance due to spread of

potential between the gate and the source and drain [77]. It is known that not only EOT, but also the physical oxide thickness play important roles in SCEs. However, it is not clear what the solution is as the gate lengths of MOSFETs approach below 20 nm [9, 74, 81].

In this chapter, we attempt to answer these questions: 1) What is the impact of using a different high-k materials with the same EOT on an ultra-scaled DG MOSFET? 2) What is the optimum thickness of high-k gate stack for a fixed EOT? 3) How do we analytically estimate the gate leakage in the off state for a specific gate stack? We show that for ultra-scaled DG MOSFETs, there is an optimum oxide thickness that balances gate leakage of a thin oxide layer and SCEs of thick oxides. We provide an analytical model for the effect of high-k on the gate control and gate leakage current.

5.1 Methodology

- A self-consistent Schrödinger-Poisson solver, based on the real-space effective-mass approximation and the wave function formalism, including direct gate and Fowler-Nordheim tunneling currents is used to simulate DG MOSFETs [70, 82–84]. This work only considers electrons for calculations, while holes are not taken into account for the n channel devices. Neglecting holes means that band-to-band tunneling is ignored in these simulations. This effect is negligible in devices, whose source-to-drain voltage (V_{DS}) is smaller than the band gap of the channel material [**?**, 52]. DG device specifications are in correspondence to the ITRS table data for the 2015 node (Fig. 1). The transport direction is aligned with the <110> crystal axis, the confinement direction is (001). The source and drain regions are doped with a donor concentration ND = [10^{20}/cm^3]. For all simulations EOT is fixed at 0.86 nm. The gate dielectric constant varies from 3.9 to 30 (Table I). Physical oxide thickness (T_{OX})

changes with the k value to keep EOT fixed, according to the following equation for the cases without any interface layer:

$$T_{OX} = EOT \times \frac{k_{OX}}{k_{SiO_2}} \tag{5.1}$$

However, for the gate stacks with SiO_2 interface layer and fixed EOT, $T_{OX} = \mathrm{T_{SiO_2}}+\mathrm{T_{high-k}}$ and $\mathrm{T_{high-k}}$ is calculated as:

$$T_{high-k} = (EOT - T_{SiO_2}) \times \frac{k_{high-k}}{k_{SiO_2}} \tag{5.2}$$

Two gate-dielectric configurations with five different oxide materials are examined in this chapter. (I) All oxide materials are directly deposited on the Si channel, without any interface oxide, with the same EOT, but different k and T_{OX}. TiN metal gate contacts are characterized by work function of 4.25 eV, and their electron effective mass ($m^* = m_0$). (II) SiO_2-high-k gate stack with TiN contacts. Effective masses are assumed isotropic (Table I). Relative dielectric constants and band gap for each material are described in Table I.

5.2 Results and Discussion

The ballistic $\mathrm{I_D - V_{GS}}$ characteristics of DG MOSFET presented in Fig. 1 are simulated for $\mathrm{V_{DS} = V_{DD}} = 0.83$ V. Fig. 2-A shows the results for the configuration I, without interface layer and with no gate leakage assumption. As expected, the thicker oxide would degrade the device performance by increasing the sub-threshold swing (SS) and lowering ON-current for the fixed OFF-current [77]. However, when gate tunneling is included, in SiO_2 case, the OFF-current rises above 100 nA/um. OFF-current below 100 nA/um is considered as the OFF state in ITRS guideline [9]. For Si_3N_4, in the configuration I, ON-current is around 3430 uA/um, which is a bit less than ON-current for the device with Al_2O_3 oxide. Although, $\mathrm{T_{OX_{Si_3N_4}}}$ is thinner than $\mathrm{T_{OX_{Al_2O_3}}}$, it provides less ON-current for the fixed OFF-current. This is a result of high-gate leakage at OFF state in Si_3N_4, which drops SS as well. $\mathrm{I_{Gate_{Si_3N_4}}}$ is 90 times more than $\mathrm{I_{Gate_{Al_2O_3}}}$ and contributes as 22% of $\mathrm{I_{OFF}}$. As the gate leakage cannot

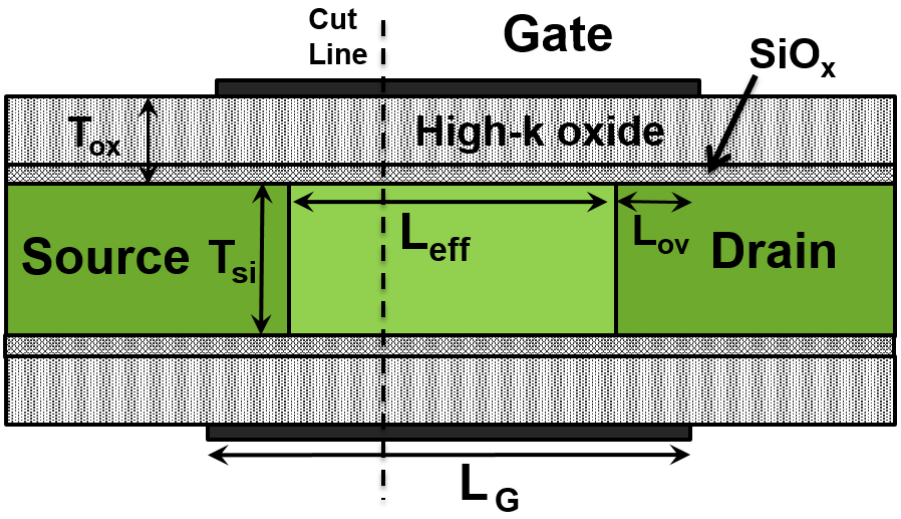

Fig. 5.1. Schematic view of the double gate MOSFETs considered in this chapter,. The geometry parameters are the same node 2015 from ITRS tables [9]. The gate length is set to 16.7 nm. There is gate overlap equal to 10% of the gate length (i.e. 1.6 nm). The thickness of Si channel is 5.3 nm. Currents are normalized by the channel width, which means simulated current divides by 2. supply voltage (V_{DD}) is set to 0.83 V.

be controlled by gate voltage, it degrades subthreshold swing, and consequently the ON-current.

Figs. 2-B to 2-D show the effect of using interface layer and its thickness on $I_D - V_{GS}$. Fig. 2-B depicts the transfer characteristics of a device with Al_2O_3 as the gate stack. As it is shown in Table I, T_{OX} ($T_{OX} = T_{SiO_2} + T_{Al_2O_3}$) with different interface layer thickness up to 5 Å, is still thick enough to keep the gate leakage very low. 4 ÅAl_2O_3 on top of 7 Å SiO_2 can not keep the leakage current low enough to turn off the device. The electrostatics and the gate leakage in the required voltage range does not vary drastically, which keeps the $I_D - V_{GS}$ characteristics similar in the interested range. In Fig. 2-C high-k material is HfO_2. T_{OX} reduces from 4.9 nm (no interface layer) to 1.5 nm (7 Å interface layer) impacts on SCEs. All cases can turn off the device ($I_{OFF} < 100$ nA/um) except the case with 7 Å interface layer. The case with 5 Å interface layer provides the best SS and respectively; better ON

Table 5.1.
Dielectric constants k, band gap, conduction band (CB) offset to Si, and tunneling effective masses (m^*) for multiple oxide materials used in this work are shown. tunneling effective masses are from multiple experiments and theoretically calculated references [85, 86]. P values are calculated as equation 4 for each specific oxide material. Starred (*) numbers are used values for these materials in our simulations. Physical thickness of each oxide material is shown for both configurations I and II.

Material	SiO_2	Si_3N_4	Al_2O_3	HfO_2	La_2O_3
K	3.9	7*-8	9*-10	20*-25	27-30*
Band gap [eV]	9	5.3	8.8	5.8	6
CB offset [eV]	3.5	2.4	2.8	1.5	2.3
m^*/m_0	0.4*-0.5	0.4	0.35	0.11*-0.17	0.27
P$[nm^{-1}]$	6.06	5.02	5.07	2.08	4.04
T_{OX}(No Interface)	0.9	1.6	2.0	4.9	6.6
$T_{OX}(T_{SiO_2} = 3Å)$	-	1.3	1.6	3.5	4.6
$T_{OX}(T_{SiO_2} = 5Å)$	-	-	1.4	2.5	3.3
$T_{OX}(T_{SiO_2} = 7Å)$	-	-	1.1	1.5	1.9

current at fixed I_{OFF} of 100 nA/um. In the case of La_2O_3 as high-k material in Fig. 2-D, T_{OX} reduces from 6.6 nm (no interface layer) to 1.9 nm (7 Å interface layer), which impacts SCEs drastically. Each case turn off the device ($I_{OFF} < 100$ nA/um), however the case with 7 Å interface layer provides the best ION at fixed I_{OFF}, as it has the thinnest oxide.

In Figs. 3-A to 3-D, performance metrics, including I_{ON}, gate leakage current in the OFF-state, subthreshold swing (SS) and DIBL for fixed I_{OFF} of 100 nA/um are depicted. I_{ON} increases by reduction of the high-k thickness for the same EOT in the gate stack (i.e. increase in SiO_2 interface layer thickness). Optimum physical oxide thickness helps the device to turn off, as well as provides stronger electrostatics. The optimum oxide thickness for each gate stack can be achieved with a layer of SiO_2 and high-k material. Al_2O_3 gate stack shows optimum ON current at $T_{SiO_2} = 3$

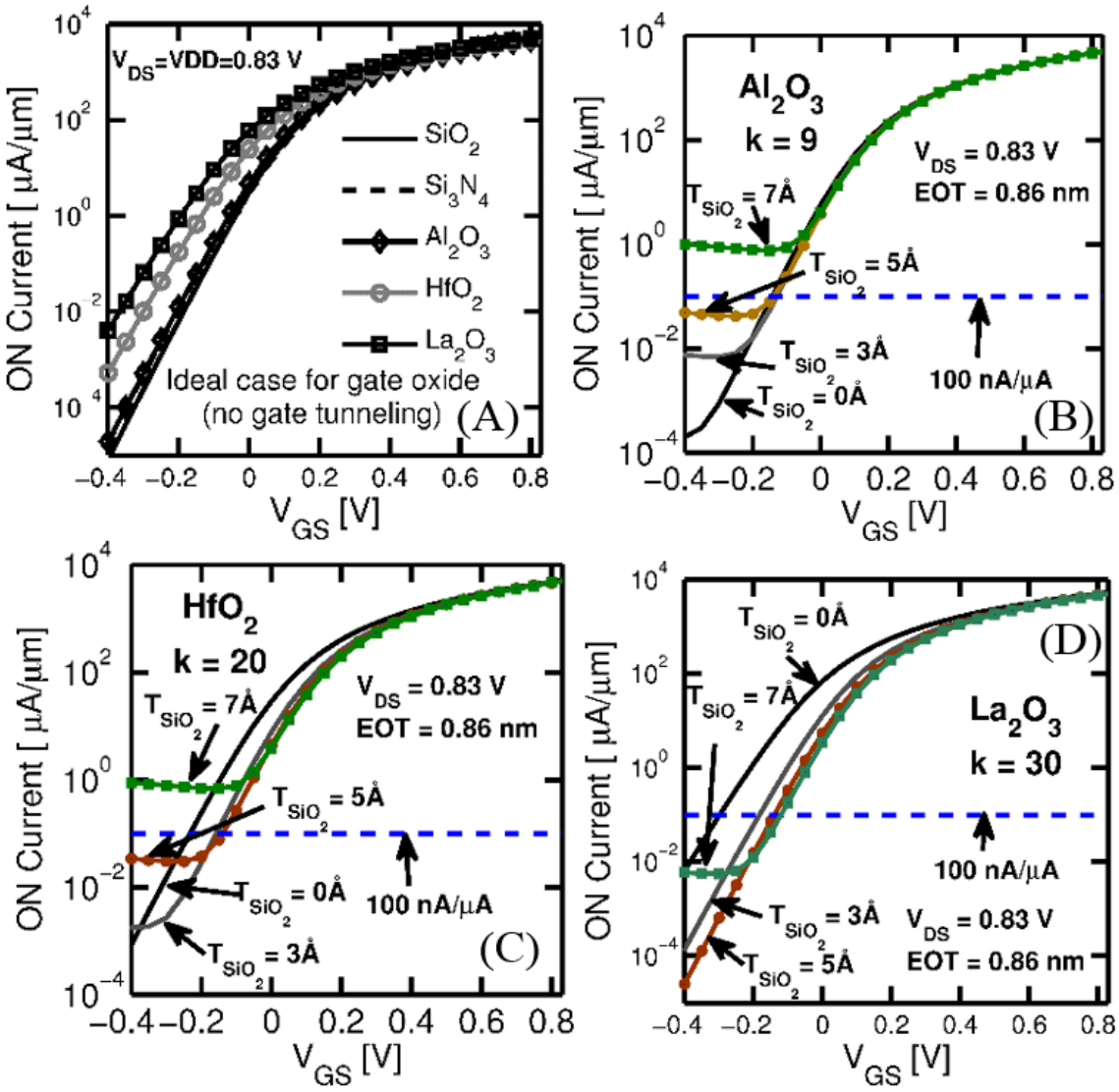

Fig. 5.2. Typical $I_D - V_{GS}$ transfer characteristics at $V_{DS} = 0.83$ V of DG MOSFETs with fixed EOT of 0.86 nm, and various dielectric materials (different k). (A) With the assumption, that ideal oxides which do not have any leakage. (B) with Al_2O_3 and SiO_2. (C) with HfO_2 and SiO_2. (D) with La_2O_3 and SiO_2.

Å. However, for both HfO_2 and La_2O_3 oxide materials thicker SiO_2 interface layer improves the device I_{ON} (Fig. 3-A). In Fig. 3-B, the gate leakage current at the OFF-state is shown, which is part of the drain current. The drain current has two components; the source to drain current and the gate to drain current (gate leakage).

Gate leakage results from tunneling of electrons through the potential barrier between the gate and the channel. I_{Gate} is exponentially related to the oxide thickness (T_{OX} = T_{SiO_2}+T_{high-k}) and oxide effective mass [**?**,74]. Al_2O_3 and HfO_2 have alike I_{Gate}. This similarity in I_{Gate} is a result of the light effective mass of HfO_2 (0.11 m_0) compared to Al_2O_3 (0.35 m_0), while their dielectric constants are different. Replacing oxide materials and varying the interface layer thickness impacts SS values (Fig. 3-C). SS depends on the electrostatics and the gate leakage, which can not be controlled by gate voltage [**?**]. Thinning T_{OX} enhances the gate control, which improves SS until I_{Gate} becomes comparable to I_{OFF}. I_{Gate} does not decay exponentially with the gate bias, which leads to higher SS value. As it is depicted in Fig. 3-D DIBL improves by a reduction in the oxide thickness [77], but for very thin oxides, higher gate leakage slows down the DIBL value reduction [9]. Thin T_{OX} improves the gate control over the channel (or top of the barrier), which results in weaker drain control over the channel.

Gate leakage occurs due to the carrier tunneling through the oxide layer between the gate contact and the channel. The potential barrier for a gate stack is depicted in Fig. 4-A. Using tunneling transmission equation, we can estimate the tunneling gate leakage. In Fig. 4-B, the tunneling transmission is calculated using equation 3, which shows a strong correlation with gate leakage in Fig. 2-B. tunneling transmission for the gate stack is calculated as:

$$Transmission \approx e^{-2(T_1.P_1+T_2.P_2+...)} \tag{5.3}$$

$$P = \sqrt{\frac{2m^*U}{\hbar^2}} \tag{5.4}$$

where T_i and P_i are thickness and decaying wave-vector of carrier in i^{th} oxide within the gate stack, accordingly. The decaying wave-vector, P_i, in each oxide layer is calculated from the effective mass, m^*, and the potential barrier height , U (the CB offset in Table I) and are listed in Table I.

In Figs. 5-A to 5-D, potential difference of OFF and ON states are shown overlapped with the electron flow in the OFF state for SiO_2 (Fig. 5-A) and different

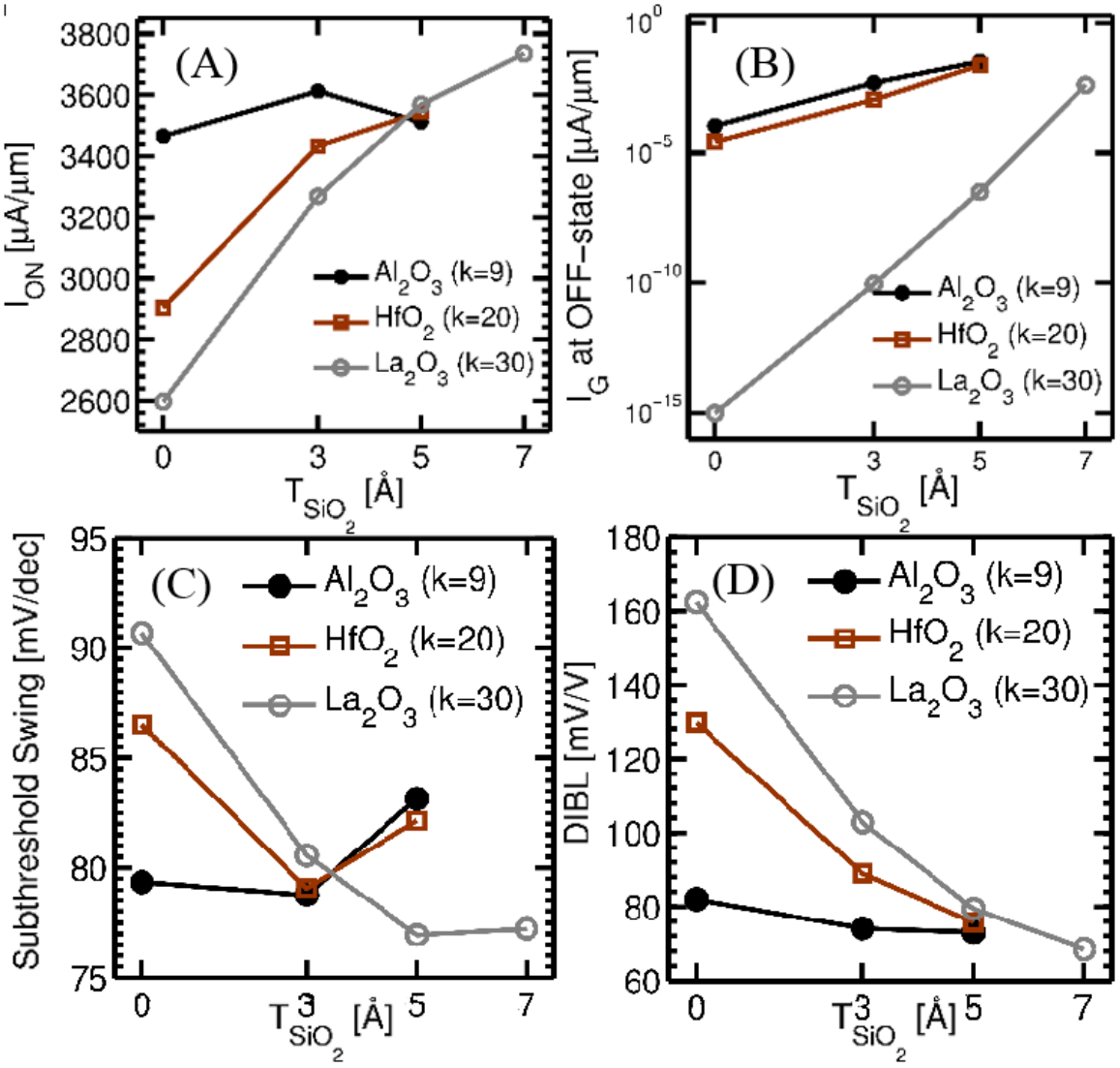

Fig. 5.3. Performance metrics of DG MOSFETs with fixed EOT of 0.86 nm, and various dielectric materials (different k) at fixed I_{OFF} (100 nA/um). If $I_D - V_{GS}$ could not go below 100 nA/um at off state were dropped out of these figures. (A) ON current I_{ON}, (B) Gate leakage (I_{Gate}), (C) Sub-threshold swing (SS) and (D) Drain induced barrier lowering (DIBL).

HfO_2 thicknesses over interface layers. Potential spread over the source and the drain are stronger for thicker oxides (Figs. 5-A and 5-B). These potential spreads show increasing fringing capacitance by using thicker oxides, which degrades the device

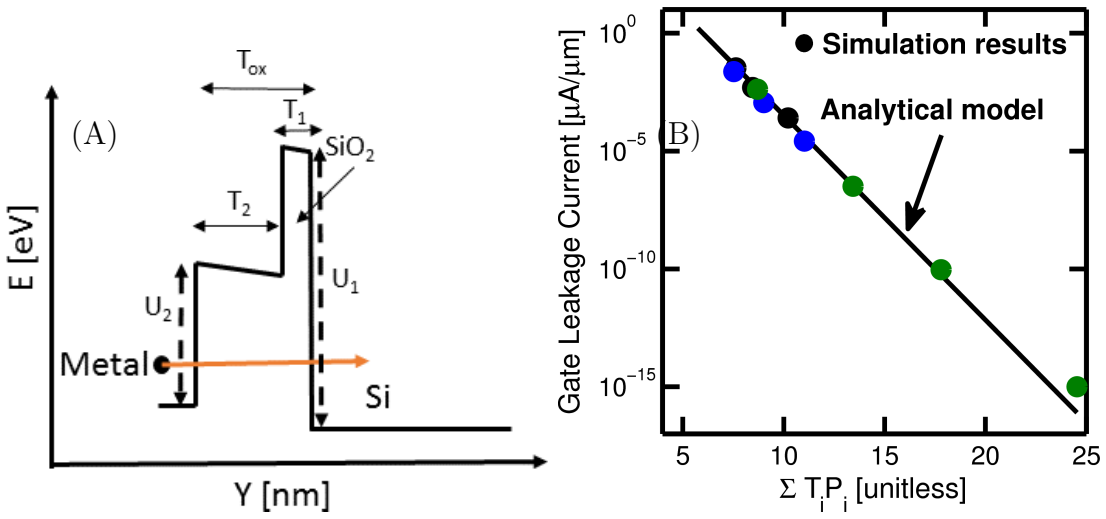

Fig. 5.4. (A) tunneling through the potential barrier for a 2-material gate stack. (B) Gate tunneling current calculated by the presented extensive computational model, QTBM, and the line is calculated by the analytical model in equation 3. Different colored circles are for different gate stacks in the simulations.

performance. By T_{OX} reduction the potential spread reduces (Figs. 5-B to 5-E). However, in the very thin oxide case (Fig. 5-E) gate tunneling increases and drastically drops the device performance.

Fig. 6-A shows the capacitance network, including the fringing capacitance between gate and source/drain, which will be added to the gate capacitance. From the potential profile and stringing field according to Figs 5-A to 5-D it can be seen that the total oxide capacitance can be estimated as $\frac{C_{OX}}{A} \approx \frac{\epsilon_{OX}}{T_{OX}}(1 + \frac{2T_{OX}}{L_G})$, where A is the gate area and L_G is the gate length and ϵ_{OX} is equivalent gate stack dielectric constant. Fig. 6-B shows the effect of the fringing capacitance on the total gate oxide capacitance from our simple model $1 + \frac{2T_{OX}}{L_G}$ on the SS values. When SiO_2 interface layer is thin, the physical thickness and fringing capacitance are large and the fringing capacitance deteriorate the performance for thick gate stacks. This simple model can be used to **estimate** the effect of physical oxide thickness and gate length on the performance of DG MOSFETs. It is not accurate because current in sub-threshold region in ultra-scaled MOSFETs have multiple components, including source to drain

thermionic current, source to drain tunneling current and gate leakage. Here, we used the SS equation [87]:

$$SS \approx ln(10)\frac{kT}{q}(1 + \frac{C_d}{C_{OX}}) \tag{5.5}$$

Where C_d is the depletion capacitance and C_{OX} is oxide capacitance. In Fig. 6-B, C_d is extracted from calculated SS for the DG MOSFET with only La_2O_3 as oxide. Then, the SS for gate stack with different SiO_2 interface layer is estimated using equation 5. In gate stack with $T_{SiO_2} = 7\ \mathring{A}$ gate leakage current increases, which shows the equation 5 will not work well at that point.

5.3 Conclusion

Scaling MOSFETs below 20 nm using thicker high-k oxides drastically degrades the device performance for a fixed EOT. Therefore, introducing higher-k oxide should be examined carefully for penalty in electrostatics. Using very thin oxides also causes gate leakage. An optimum combination of k value and oxide effective mass and thickness, or engineered gate stack, should be used to provide strong electrostatics and acceptable gate leakage. We provided a simple analytical model to estimate the gate leakage for each gate stack, as well as a model for the electrostatic degradation due to the fringing capacitance. Additionally, we showed there is an optimum gate stack thickness for any high-k material. Gate overlaps, source/drain doping and scattering can impact on the quantitative results and for a specific device, it will be more accurate to do further investigation to find the optimum geometry and design for the best performance.

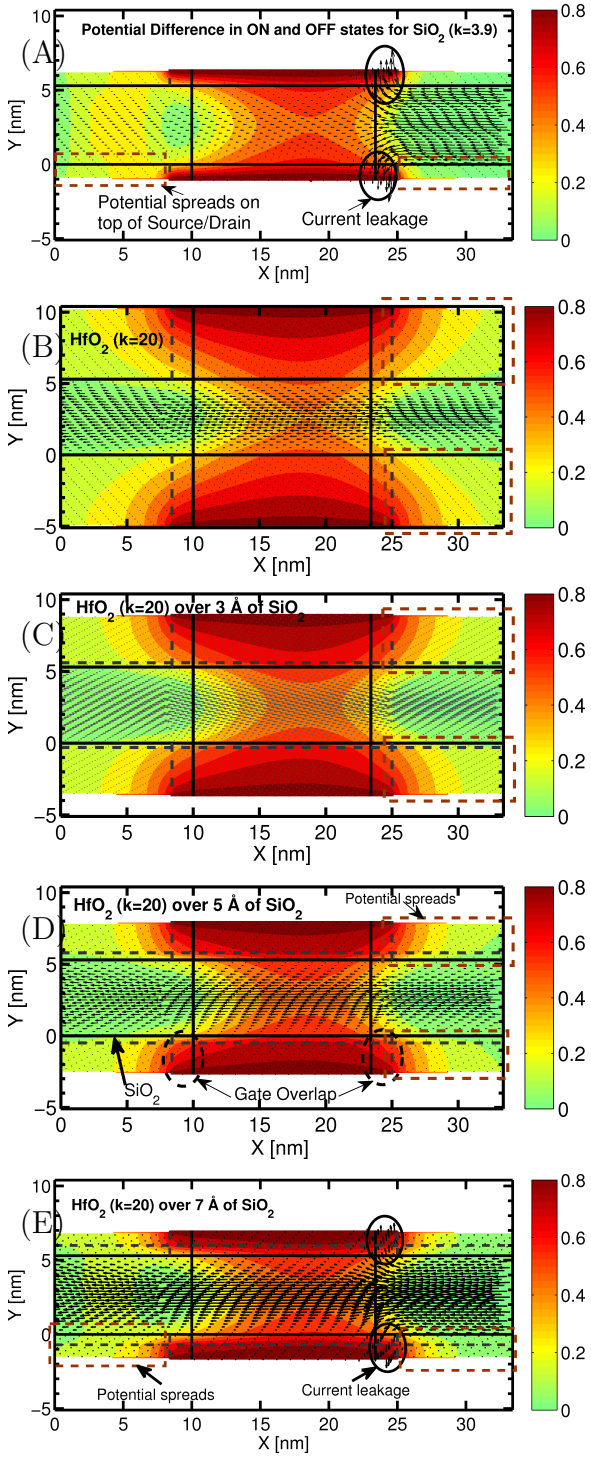

Fig. 5.5. Potential difference between OFF and ON states overlapped with the electron flow in the OFF-state for DG MOSFET with (A) SiO_2 oxide. (B) HfO_2 without interface layer. (C) HfO_2 over 3 Å interface layer. Potential spread over the source and the drain sections are weaker than only HfO_2 case, but stronger than SiO_2. (D) HfO_2 over 5 Å interface layer. Potential spread over the source and the drain sections are weaker than HfO_2 over 3 Å interface layer, but stronger than SiO_2. Still, gate tunneling is negligible. (E) HfO_2 over 7 Å interface layer. Gate leakage current is very strong which prevents the device from turning off.

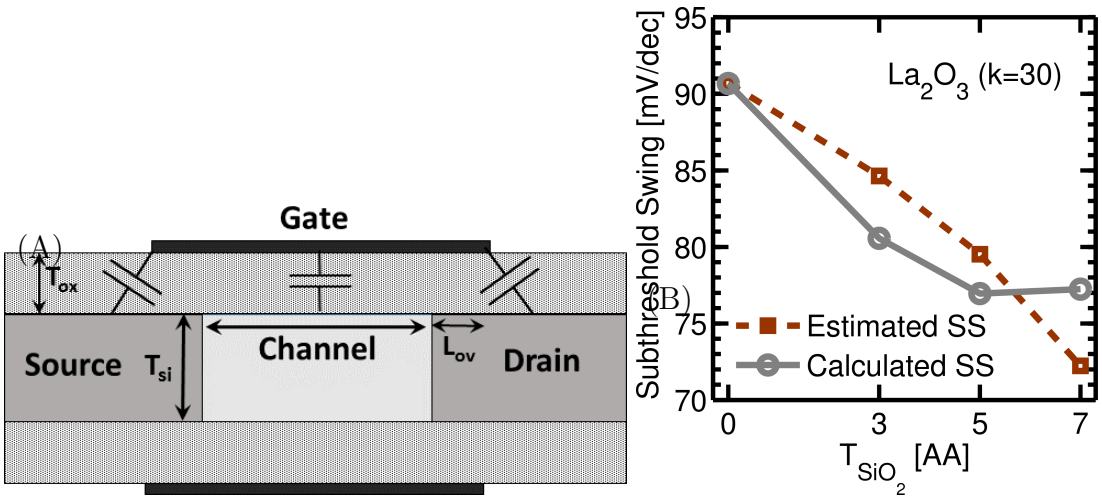

Fig. 5.6. (A) Capacitance network between gate and channel, source and drain is depicted. Fringing capacitance between the gate and the source/drain adds up to the gate capacitance, which degrades the device performance for thicker gate stack. (B) Electrostatics degradation factor, $1 + \frac{2T_{OX}}{L_G}$, reduces by thinning the gate stack. This impact on SS for different La_2O3 gate stack is depicted.

6. SIGE BASED MOSFETS

High mobility channel materials have been widely pursued for the continued miniaturization of transistors. Due to its superior hole mobility Silicon Germanium is a leading candidate as a replacement channel material for PMOS. High mobility materials including SiGe and III-V have also been incorporated in modified FinFET geometries to lower the leakage resulting from short channel electrostatics. For example, SiGe cladding FinFET employs a thin SiGe cladding layer as the channel around a non-conducting Si fin to provide stronger gate control. In this work, atomistic simulations are used to obtain band structure and transport characteristics in the ballistic regime for SiGe channels between 2 to 8 nm in the presence of strain for various Ge mole fractions between zero to one. We show that reducing thickness of the <110> SiGe channel improves ballistic ON-current and increases the band gap desirable for lowering leakage. These changes are quantified and are more pronounced for higher germanium percentage. Moreover, while strain has an important role in improving ballistic ON current, its effect diminishes for channel thicknesses below 3 nm. Finally, we compare a cladding FinFET composed of a 3 nm SiGe channel over a Si fin with a bulk FinFET of the same material composition and footprint. The simulations show that cladding FinFET outperforms regular FinFET; largely as a result of quantum confinement (QC) effects.

6.1 Introduction

FinFET metal oxide semiconductor field effect transistors (MOSFETs) have been commercialized as a replacement for conventional planar devices to extend scaling for near future technology nodes [88]. To enhance the performance of these devices in the ON-state, high mobility materials such as SiGe [89], Ge and III-V are potential

replacements for Si. Among these materials, SiGe is already in use as an embedded source/drain stressor in Si based PMOS devices to enhance hole mobility [90]. In addition to improvement in channel mobility, modified FinFET geometries have been investigated to enhance the device electrostatics to reduce the leakage current. Recently, forming channel SiGe layer over a Si fin for MOSFETs has been proposed and studied [91, 92]. The recent experimental reports demonstrate that such devices exhibit higher hole mobility, as compared to Si counterparts [92]. Additionally, it is believed that the Si fin in these cladding devices (Fig.1-a) should be non-conductive with the intent of reducing the OFF-state leakage [**?**]. However, some critical questions remain about these structures, including: 1) How large is the band offset between the SiGe cladding layer and the Si fin (Fig.1-a) to hold the carriers inside the channel? 2) What are the effects of QC, germanium percentage and strain on band gap, which is a critical parameter for the OFF-state leakage? 3) How the aforementioned factors vary the ballistic performance?

We show that QC and Ge mole fraction provide large enough (at least 300 meV) band edge offset between the channel SiGe layer and the Si fin (Fig.1-a). This keeps the carriers inside the high mobility channel in the ON-state, which is of critical importance for such devices to function (see Figs. 2-a and -b). This is backed up both by atomistic as well as TCAD simulations. In addition, QC causes band gap widening that reduces leakage current in the OFF-state, which is of special concern, as bulk SiGe has lower band gap than Si. We quantify the band gap widening for SiGe channels with thicknesses between 2 nm to 8 nm for various SiGe alloys (see Figs. 3-a and -b). Strain has a positive effect on performance and band edge offset but degrades band gap. The extent of these effects are also quantified.

As a gauge for potential performance of these devices, we show ballistic transport calculations for SiGe channel with Ge mole fractions ranging from 0.5 to 1 in the 110 transport direction in the presence of strain and QC. Interestingly, QC significantly enhances ballistic ON current (see Figs. 4-a and -b). This enhancement is more pronounced for higher concentrations of Ge. Strain markedly increases ballistic

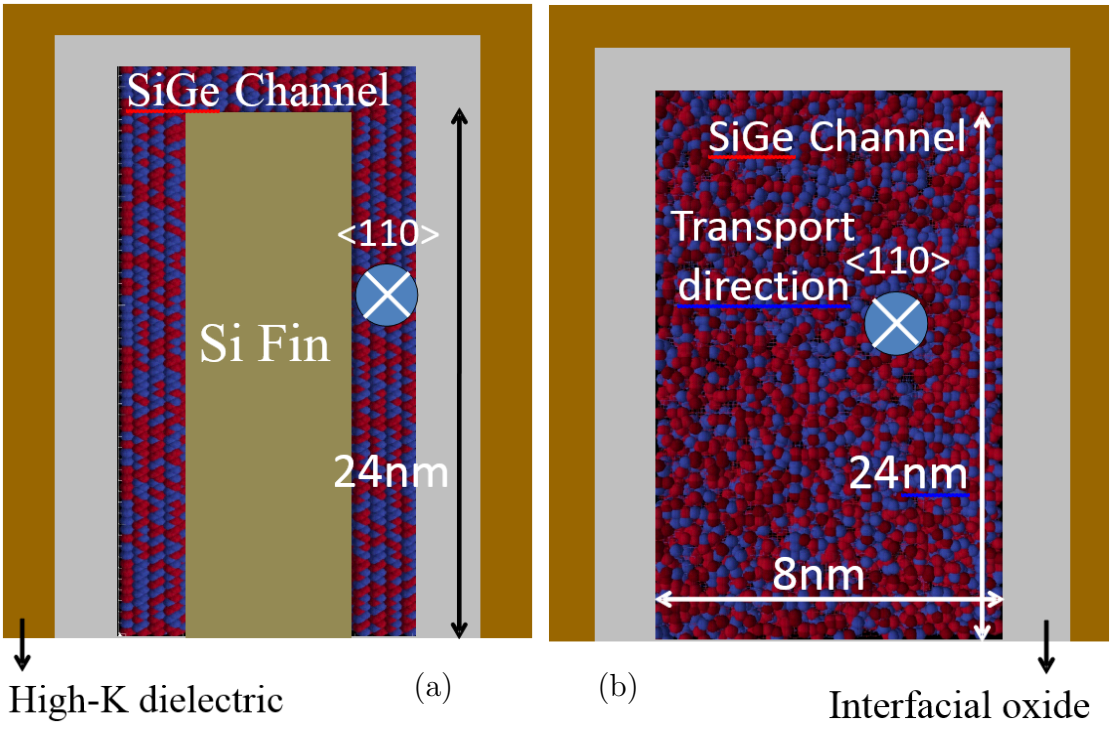

Fig. 6.1. FinFET cross section for PMOS devices: a) A cladding SiGe <110>/[110]/[100] PMOS FinFET as the channel (cladding layer) is over the Si fin. b) A 24x8 nm square cross section regular SiGe FinFET <110>/[110]/[100].

ON current in the SiGe channel but this effect is somewhat diminished for channel thicknesses below 3 nm, where ON current has already been boosted by QC. We compare a cladding FinFET composed of a 3 nm SiGe cladding layer over a Si fin with a bulk FinFET of the same conditions using similar material composition, equal footprint and the same stress. The simulation results show that the cladding FinFET outperforms the regular FinFET. Our observations are explained based on the changes to the band structure due to QC, different germanium concentrations and strain (see Fig. 5).

6.2 Methodology

Confinement and strain alter band structures. Also, percentage of different materials in alloys can change band structure. In cladding devices, strain on the channel is induced by external stresses from contacts, as well as lattice mismatch between cladding layer and the Si fin. Sentaurus TCAD is used for extensive process simulations to provide realistic strain component values of cladding layer and Si fin. Strain values are provided as input to calculate the band structure. The electronic band structure is calculated by the atomistic 20-band $sp^3d^5s^*$ with spin-orbit coupling tight-binding (TB) model [13,93] incorporated in NEMO5 [**?**,18]. The TB parameters for the different compositions of $Si_{1x}Ge_x$ are calculated based on the virtual crystal approximation (VCA) method [94]. Computationally simulating the structure as a whole is very expensive. Therefore, we split the structure into 4 parts (Si fin, two sides and one top part), which are then individually computed to speed up calculations. The results of the individual simulations show very close agreement to simulation of the entire device simultaneously. Valance band offset (VBO) is determined as the difference between the top of the valance bands of Si fin and SiGe cladding layer. These strain values are used as input for band structure calculations in NEMO5. Once the electronic structure is obtained, the performance metrics, including overdrive current are derived using the top-of-the-barrier transport model [95].

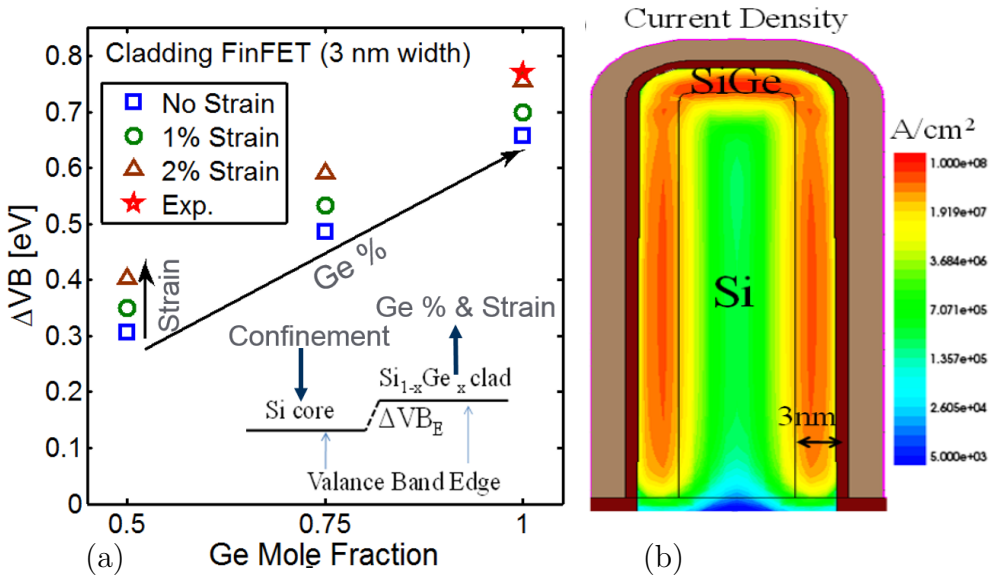

Fig. 6.2. a) Top of the VB (TVB) for cladding part moves up with confinement, Ge mole fraction and compressive uniaxial strain. The plot shows how much Ge mole fraction and strain could shifts the top of the VB for SiGe cladding part. The [red] star is experimental TVB value for Ge and the [blue] square is the calculated value for confined Ge with no strain, which is $\sim 110meV$ lower than bulk value. $Si_{0.5}Ge_{0.5}$ cladding layer with 3 nm thickness has the minimum VBO in respect to bulk Si, which is $\sim 300meV$ and Ge provides highest VBO, 900 meV. b) The TCAD output in the ON-state ($@V_{GS} - V_{DS} = 0.5V$ with EOT = 1 nm), which shows the carriers are confined in the cladding part.

Table 6.1.
These values are calculated energy level for TVB for the cladding part. TVB values for the Si fin is almost at 0 meV.

VB [eV]	$Si_{0.5}Ge_{0.5}$	$Si_{0.25}Ge_{0.75}$	Ge
$NoStrain$	0.31	0.49	0.66
-1%	0.35	0.53	0.70
-2%	0.40	0.59	0.76

6.3 Results and Discussion

Large valance band offset between Si fin and SiGe cladding layer is crucial to confine the carriers in the cladding region, i.e. the channel (see Fig. 1-a). The effects of confinement, uniaxial strain along the transport direction, and Ge mole fraction on VBO are depicted in Fig. 2-a. *Confinement downshifts the top of the valance band (TVB).* Confined Ge has $\sim$110 meV lower TVB than the experimental value for bulk Ge, i.e. 770 meV.

Compressive uniaxial strain, as well as higher germanium concentration move up the TVB of the cladding part compared to Si fin, resulting in increase in the VBO (Fig. 2-a and Table I). The induced upward shift in the VBO is helpful in order to keep the carriers in the channel region. Stress components values are simulated for multiple Si fin width, including 2, 3, 4 and 6 nm with 3 nm cladding layer width and different Ge percentage in both the cladding layer and contact. Germanium in contacts is used for external stresses over the channel. Stress over the channel contains multiple components. The major component is found to be compressive uniaxial along the transport while the other stress components are negligible. The stress component along the transport can change from tensile to compressive, depending on Ge mole fraction in the contacts, cladding layer width, and Si fin width. We only report no-strain to 2% compressive strain cases, since tensile stress on PMOS degrades the mobility and is not pertinent to this study. Compressive uniaxial strain up to 2% along the transport direction in the SiGe cladding layer elevates VBO by $\sim$100-150 meV.

Higher Ge mole fraction in SiGe raises the top of valance band. Ge percentage is increased in the channel up to 100% to gain better understanding of germanium effects (Fig. 1-a and Table I). Ge mole fraction (>0.5) generates $\sim$ 300-900 meV of VBO (see Fig. 2-a). These results are consistent with the TCAD outputs. Fig. 2-b is one of the outputs of the TCAD, which demonstrates how the cladding layer could hold the transport carriers in the ON-state. Reducing the width of the Si fin from 6

nm to 2 nm has negligible impact on its TVB. With or without strain, TVB of the Si fin moves up to 12 meV , at most, as a result of width reduction.

Leakage current is a concern with Ge based MOSFETs due to its low band gap (BG). BG is a detrimental indicator for the leakage in the MOSFET devices. Low BG leads to high band-to-band-tunneling, which increases the off state leakage current [96]. In addition to change in TVB, germanium and strain on alter the BG of cladding layer (Fig. 3-a). The compressive strain and higher germanium percentage reduce the BG. Ge has a lower BG compare to Si. By adding more Ge to the channel lower BG is expected. Fig. 3-b depicts, that confinement increases the BG due to pushing down the TVB. This effect is similar to particle in a box, which with stronger confinement, energy levels move up [97]. It is found that in highly quantum confined structures, the band gap can be as high as 960 meV for germanium percentage up to 100%. This number reduces in the presence of strain, but remains above 900 meV for 2% strain. Fig. 3-b shows that confinement leads to higher BG that can lower leakage current.

In order to quantify the effect of factors discussed previously on performance, we also calculated the ballistic ON-current. Ballistic ON-current provides insight about maximum possible performance of a device. Calculation results along with experiments [92] show <110> transport orientation outperforms <100>. Therefore, only the results for <110> are reported in Fig. 4. In the ballistic performance calculations, EOT is set to 1 nm, and over drive current is reported at $V_{GS} - V_{DS} = 0.5V$. Top-of-the-barrier model calculations show that 3 nm cladding devices outperform regular FinFETs in all cases of different Ge mole fraction and strains. This method shows that QC in cladding devices improves ballistic ON-current for the same footprint device (Fig. 4-a). This effect is significantly higher for increased Ge concentration. The effects of strain beyond 1% diminishes in heavily confined structures (Fig. 4-b).

Band structure analysis explains the aforementioned effects on the ballistic ON-current. Ge has lighter transport effective mass (m^*) compare to SiGe, meaning there is a sharper curvature on top of the VB. In both cases of Ge and $Si_{0.5}Ge_{0.5}$, **confinement sharpens the curvature**, but for Ge it is more pronounced (Fig.

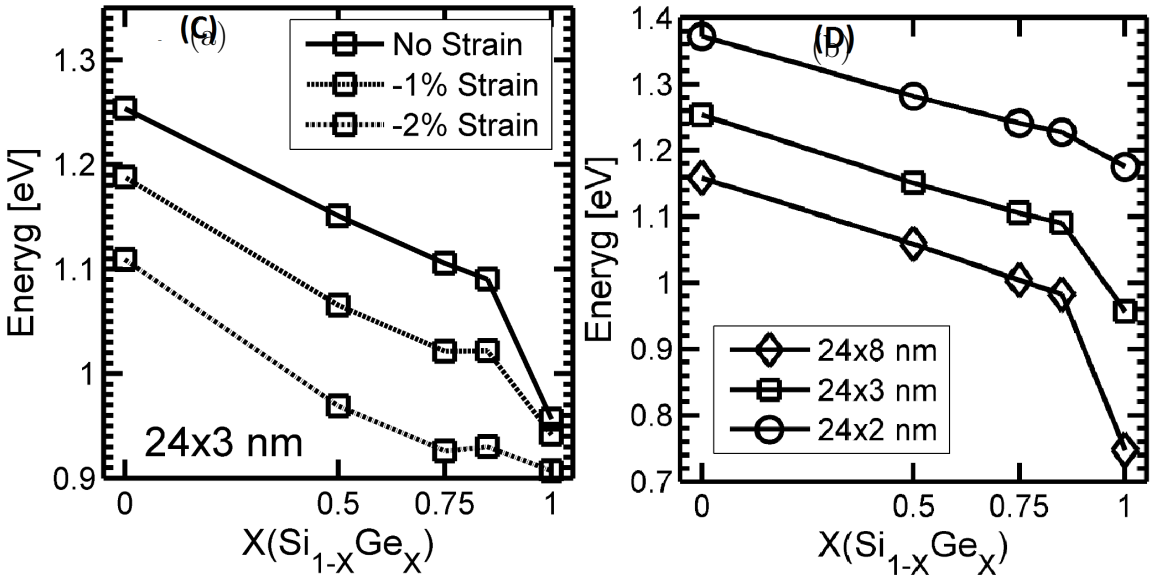

Fig. 6.3. (a) The band gap for confined cladding layer with 3 nm thickness. Ge mole fraction and compressive strain decrease the band gap. (b) Confinement can increase the band gap. The band gap variation in higher Ge with confinement is more pronounced.

5-a). m^*_{Ge} decreases from $0.107m_0$ to $0.065m_0$ (39% reduction), while for $Si_{0.5}Ge_{0.5}$ this value decreases from $0.155m_0$ to $0.114m_0$ (26% reduction). Strain also sharpens the bands curvature, which results in higher mobility. For cladding layer SiGe (24x3 nm), m^* for no strain is $0.114m_0$. By applying 1% compressive uniaxial strain, m^* will be $0.098m_0$ (14% reduction) and with 2% it declines to $0.092m_0$ (19% reduction in total). In regular SiGe FinFET (24x8 nm), m^* for no strain is $0.155m_0$. By applying 1% compressive strain m^* will be $0.102m_0$ (34% reduction) and with 2% compressive strain m^* decreases to $0.091m_0$ (41% reduction in total). This shows confinement diminishes the effects of strain, Fig. 5-b. The difference between 1% and 2% uniaxial strain at the range of first few k_BT is minimal, leading to reduced improvement in 2% strain cladding device performance (Fig. 5-b).

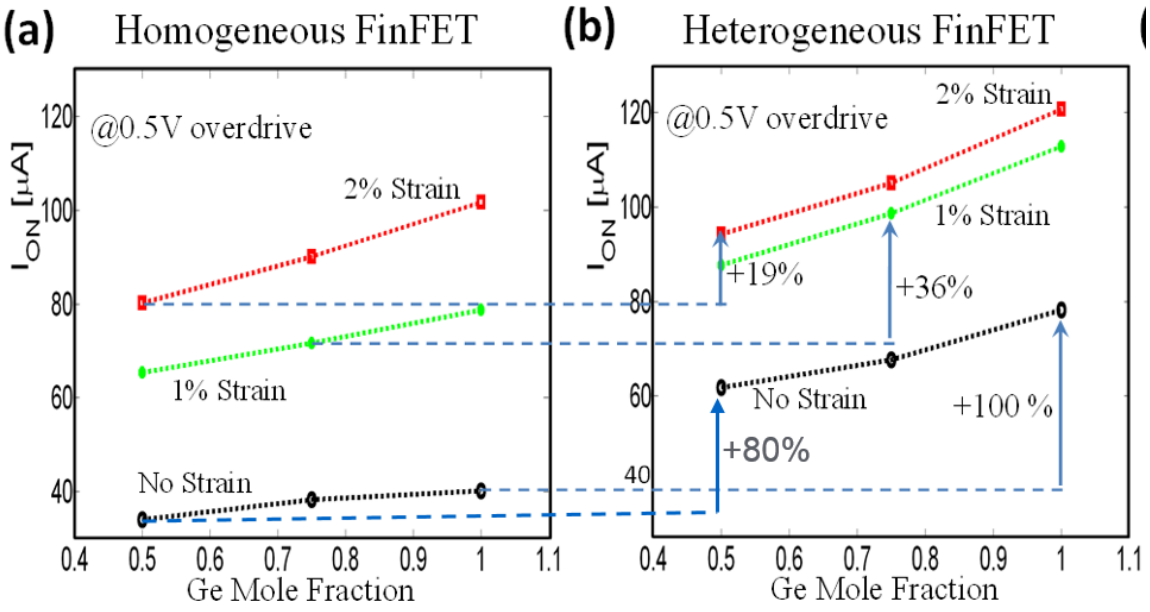

Fig. 6.4. Calculated overdrive current for (a) a regular FinFET and (b) a cladding FinFET with 3 nm cladding layer thickness with different Ge mole fraction and strain.

6.4 Conclusion

In this work we analyzed VBO, band gap, strain effects, confinement and used this analysis to asses feasibility and performance of SiGe cladding PMOS FinFETs. SiGe based cladding FinFET is compared with a regular FinFET. Five main conclusions of this work: 1) A thin SiGe layer on a Si fin can confine a significant number of holes and carry most of the current in cladding FinFET; 2) confinement raises the band gap, which helps to reduce the leakage current for SiGe based MOSFETs; 3) confinement and strain interplay on improving SiGe based device performance and with higher confinement, the effects of strain will be less pronounced; 4) the effects of confinement are more pronounced on higher germanium concentration; 5) the device overdrive current is higher for cladding structures than regular FinFETs. The computational models show SiGe cladding PMOS FinFET is an attractive alternative to Si channel PMOS FinFET.

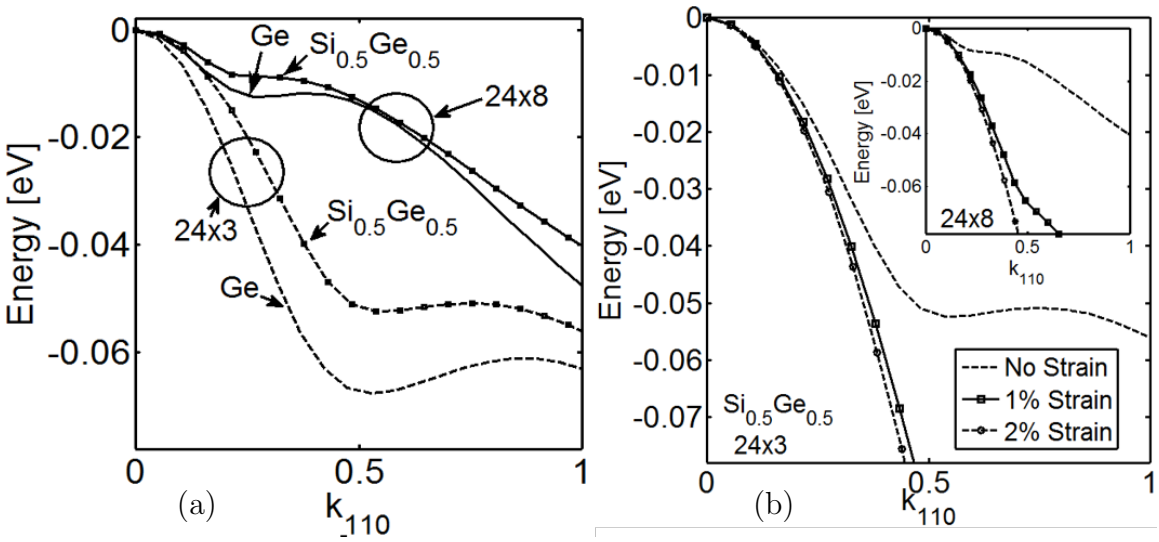

Fig. 6.5. (a) Band structured for Ge and SiGe with 24x8 nm cross section shows that Ge has narrower curvature, which means lighter carrier effective mass. With reduction in the width of the structure from 8 nm to 3 nm, the curvature for SiGe and Ge narrows. But for the case of Ge, this effect is more pronounced. (b) Strain narrows the curvature of the band structure around Γ point. This effect diminishes with confinement. The top valance band for the cladding layer (24x3 nm) is depicted and the inset shows the top valance band for the FinFET structure (24x8 nm).

7. SUMMARY AND FUTURE WORK

The present thesis explained the semi-empirical model development as a basis for computational and numerical modeling based on solid-state physics to explore the electron and lattice behavior at the nano-scale. Computational models are used for projecting CMOS scaling and finding issues and solutions for those issues. Source to drain tunneling is a critical scaling drawbacks. S/D tunneling is analyzed and two solutions are proposed. Thick high-k gate stack also degrades ultra-scaled MOSFETs. There should be an optimum thickness for each gate stack to provide low gate leakage as well as strong electrostatics. We provided an analytical model to estimate gate leakage for each gate stack. Lastly, confined SiGe based MOSFETs are analyzed as a potential solution for next technology nodes. Each part of the thesis is summarized in next sub-sections.

7.1 Semi-empirical Model Development

The process for a semi-empirical model development using parallel genetic algorithm was shown. The semi-empirical model for semiconductor lattice properties of GaAs was achieved by fitting its constant parameters to experiment. The model was verified to be predictive for other frequencies, as well as it is extendable for other materials like InAs. The model was used to predict thermo-electric properties of a square cross section nanowire as a practical application. The process for developing complex semi-empirical models are similar. We used the same method to fit parameters for strained Si [13] and MoS_2, GaSb and InSb, which we did not publish yet.

7.2 CMOS Transistor Scaling Roadmap

In this thesis, the need for revising the roadmap projection with more sophisticated device modeling tools is discussed (see Fig. 7.1). The methodology using a full band atomistic quantum transport tool is explained along with achieved results. The scaling effects on the device ON-current, SS, DIBL are also discussed and depicted in Fig. 7.2. The ON-current trend shows a decline with the gate length reduction, due to S/D tunneling, short channel effects and supply voltage scaling.

The major issue with scaling in ultra-scaled devices is S/D tunneling. Scaling leads to higher tunneling ratio and higher SS, lowering the device performance. Innovations to improve MOSFETs performance in further scaling such as using different channel orientation, better channel materials like SiGe or III-V or new device architectures such as vertical nanowire are vital for the future of the semiconductor industry.

The back-scattering model used in this work is simple and computationally inexpensive. The results can be more accurate by improving the approach to include scattering effects like full quantum treatment [98]. The next steps for more complete roadmap involve adding Si based nanowire and III-V and Ge based multi-gate devices.

The objectives for future work include the following:

- Exploring the design space for III-V/Ge based different device architectures and type (e.g., PMOS and NMOS).

- Extracting ION, SS, DIBL, required series resistance and finding out optimum design parameters for III-V/Ge based devices and nanowires.

7.3 Design Guidelines for ultra-scaled MOSFETs

A general understanding of MOSFET performance in the source-to-drain tunneling dominant regime (sub-12 nm channel length) is provided. In sub-12 nm, for each L_{Eff} there is an optimum m^*, which provides highest I_{ON}/I_{OFF} and lowest SS. It has been shown that optimum heavy effective mass leads to a better C_q, which means

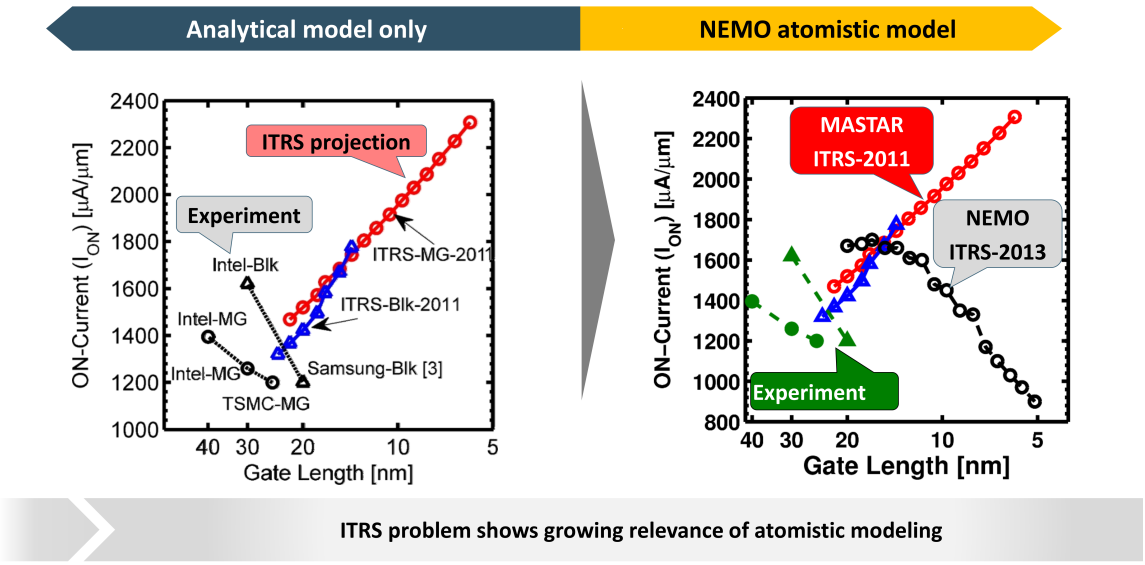

Fig. 7.1. On the left, the recent best experimental values are shown along with MASTAR's projections for DG and bulk devices. MASTAR's calculations are very dependent on the recently fabricated MOSFETs. Specifically, for future technology nodes cannot capture critical physics, such as quantum confinement and source to drain tunneling. Therefore, there the projection results can be far from practical MOSFETs performance. On the right side MASTAR's and NEMO's projections are depicted on top of recent experimental values.

improved DIBL and SS. For heavy effective mass, the ON-current is lower due to lower injection velocity. It is found that for nanowires, the optimum m^* lies between $0.2 - 1.0\, m_0$. A guideline to achieve the optimum effective mass for Si NMOS and PMOS at a given channel length is provided. It is known thinning Si can impact on its dielectric constant, which varies the potential profile shape and can change the tunneling ratio. These effects need to be investigated carefully. Different properties of Si MOSFETs, such as different transport orientations, stress, cross-section size and shape can be explored to obtain the optimum masses for different channel lengths. Other factors that affect S/D tunneling, include dielectric constant, high-k oxides, doping profile, band gap, non-parabolicity and scattering, which need further inves-

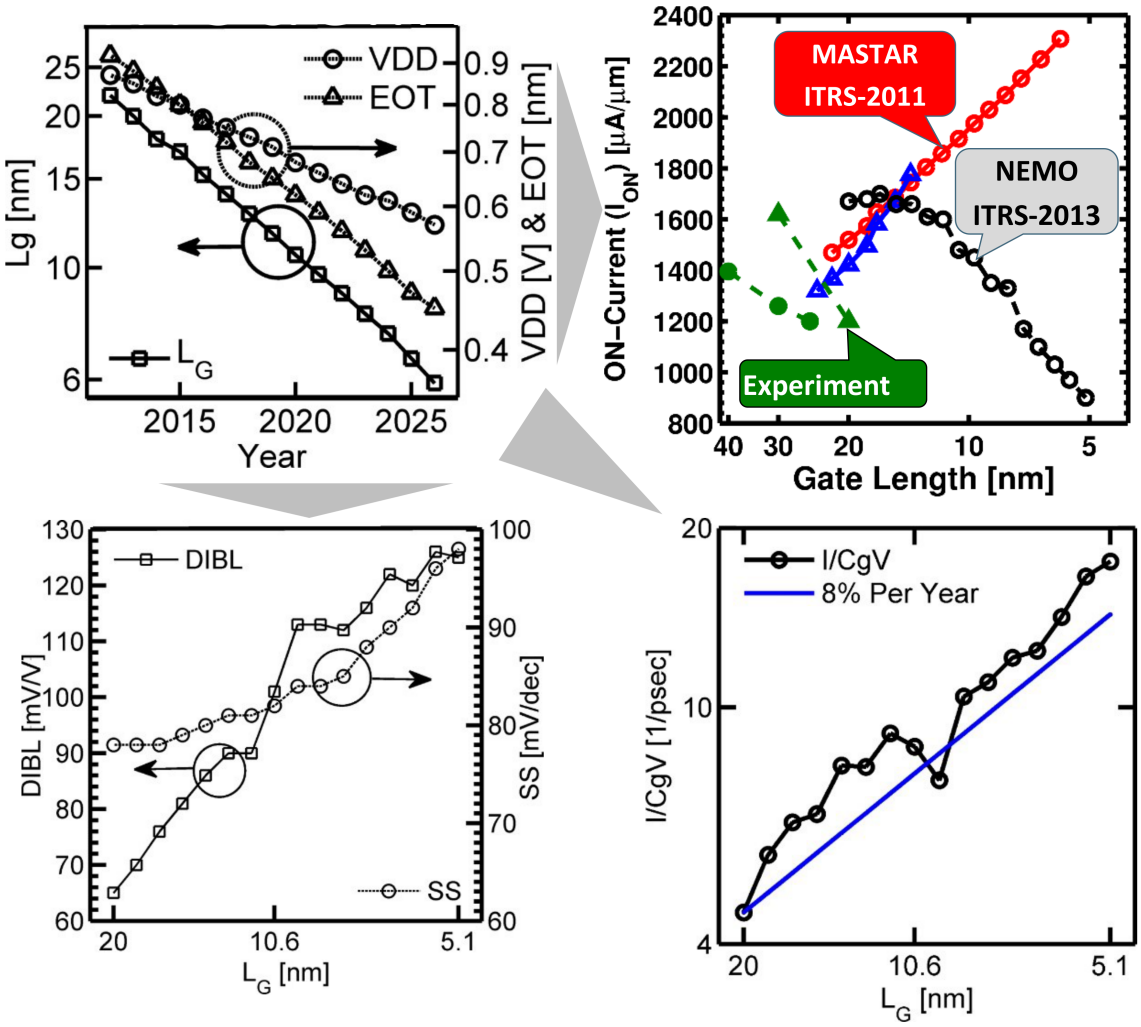

Fig. 7.2. (A) Scaling L_G, EOT, and V_{DD} from the ITRS PIDS-2012 table are depicted. (B) The calculated ON-current using MASTAR [17] for DG and bulk is compared with the recently reported best experimental data. (C) DIBL and SS value trends for geometry scaling are depicted. (D) The intrinsic speed ($I/(C_G V_{DD})$) keeps increasing by more than 8% per year.

tigation. Using heavy mass materials might impact device power consumption, f_{max} and gate leakage, which should be studied separately.

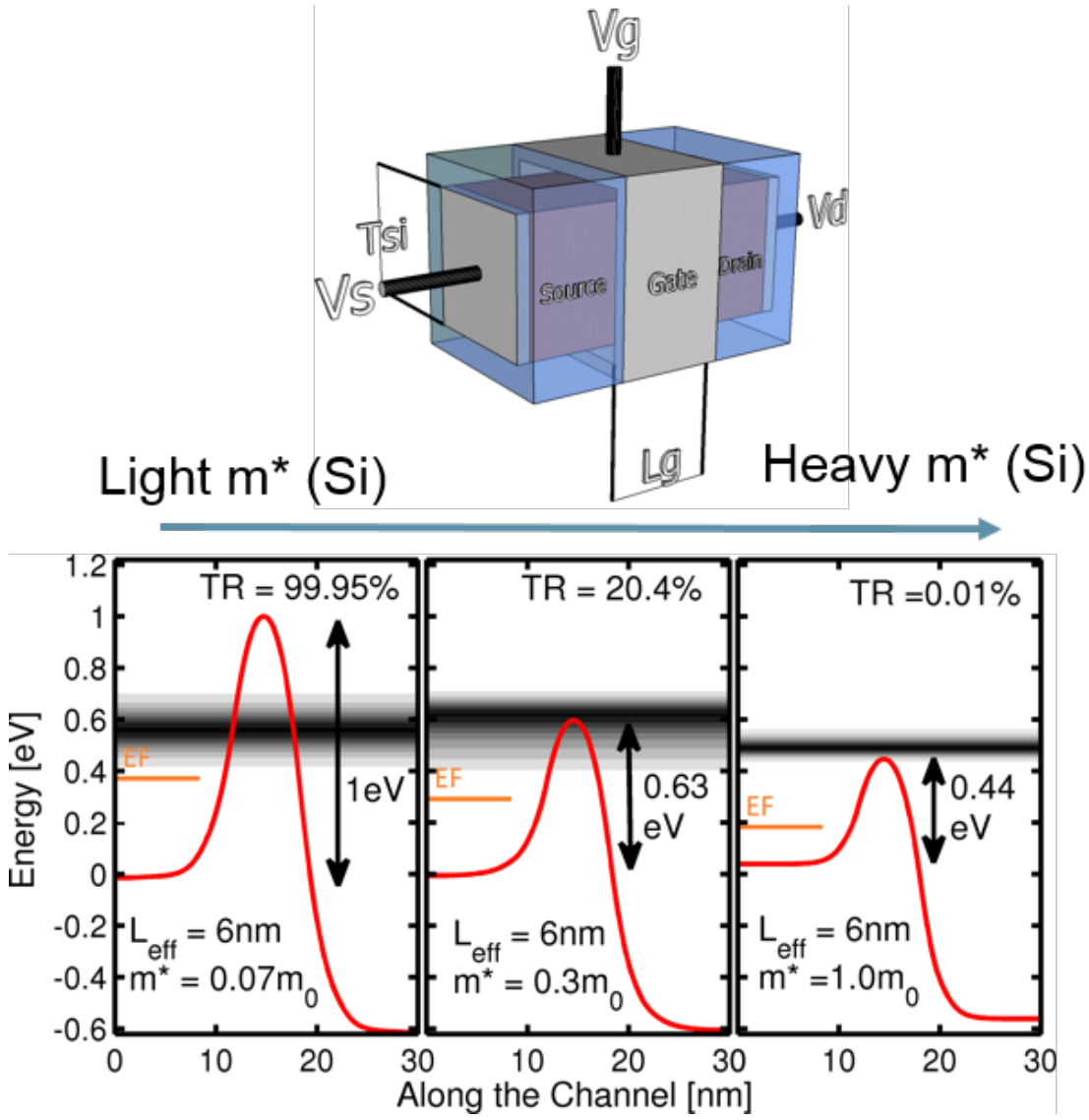

Fig. 7.3. (A) A 5x5 nm^2 square cross-section nanowire with different gate length is used as target device. (B) Current density (JEs) vs potential barrier. This shows the current above and below the potential barrier in the cases of $m* = 0.07$, 0.3 and 1.0 m_0 for $L_{Eff} = 6\,nm$. Tunneling rate (TR) is the ratio of tunneling current to total current. Heavier m^* materials have larger density of states which makes the distance between contacts Fermi-levels (EF) and bottom of conduction band (i.e. at 0 eV for the source contact) smaller in compare to light m^* materials for the same doping density. For lighter m^* the gate voltage needs to be higher negative to produce higher barrier to keep the OFF-current fixed at the 100 $nA/\mu m$.

7.4 Guidelines for High-K Gate Stack Design

Scaling MOSFETs below 20 nm using thicker high-k oxides drastically degrades the device performance for a fixed EOT. Therefore, introducing higher-k oxide should be examined carefully for degradation in electrostatics (see Fig. 7.4). Using very thin oxides also causes gate leakage (Fig. 7.2-A). An optimum combination of k value and oxide effective mass and thickness, or engineered gate stack, should be used to provide strong electrostatics and acceptable gate leakage, as Fig. 7.5 depicts. We provided a simple analytical model to estimate the gate leakage for each gate stack, as well as a model for the electrostatic degradation due to the fringing capacitance. Additionally, we showed there is an optimum gate stack thickness for any high-k material. Gate overlaps, source/drain doping and scattering can impact on the quantitative results and for a specific device, it is necessary to investigation in order to find the optimum geometry and design for the best performance.

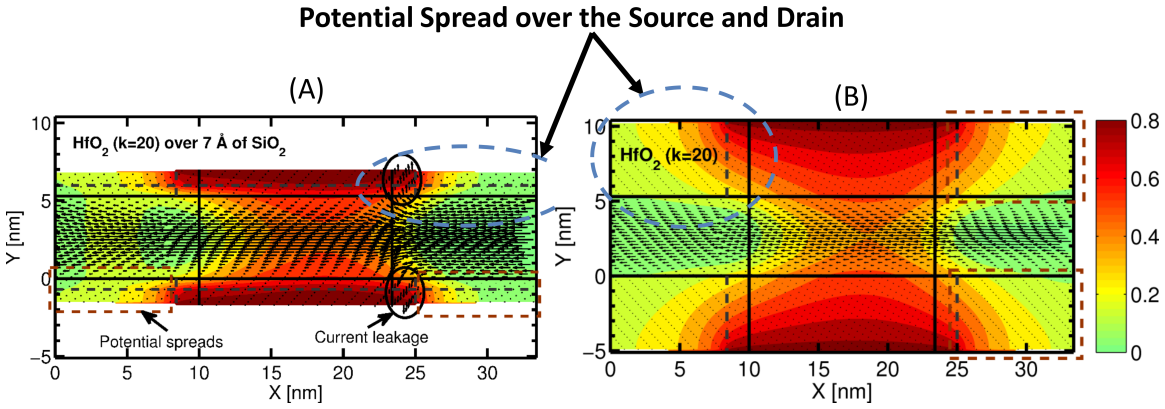

Fig. 7.4. Shows potential spread over the source and the drain in double gate MOSFET with HfO_2 gate stack (A) with 7 $\mathring{A}$ interface layer (SiO_2) and (B) no interface layer.

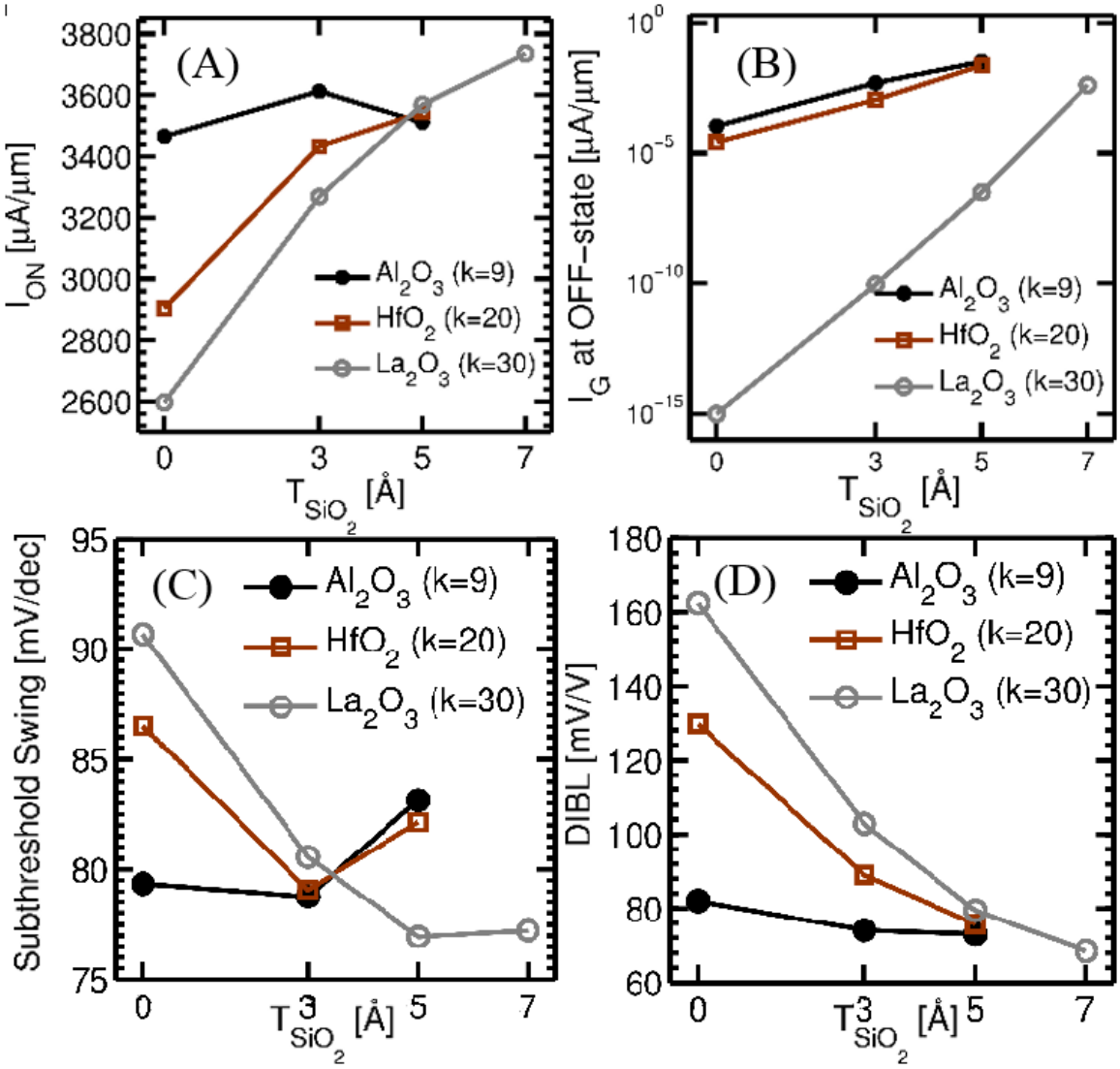

Fig. 7.5. Performance metrics of DG MOSFETs with fixed EOT of 0.86 nm, and various dielectric materials (different k) at fixed I_{OFF} (100 nA/um). If $I_D - V_{GS}$ could not go below 100 nA/um at off state were dropped out of these figures. (A) ON current I_{ON}, (B) Gate leakage (I_{Gate}), (C) Sub-threshold swing (SS) and (D) Drain induced barrier lowering (DIBL).

7.5 Quantum Confined SiGe-based MOSFETs

In this thesis, we analyzed VBO, band gap, strain effects and confinement and used this analysis to asses feasibility and performance of SiGe cladding PMOS FinFETs. SiGe based cladding FinFET is compared with a regular FinFET. The results can be summarized as: 1) a thin SiGe layer on a Si fin can confine a significant number of holes and carry most of the current in cladding FinFET; 2) confinement raises the band gap, which helps to reduce the leakage current for SiGe based MOSFETs; 3) confinement and strain interplay on improving SiGe based device performance and with higher confinement, the effects of strain will be less pronounced; 4) the effects of confinement are more pronounced on higher germanium concentration; 5) the device overdrive current is higher for cladding structures than regular FinFETs. The computational models show SiGe cladding PMOS FinFET is an attractive alternative to Si channel PMOS FinFET.

APPENDICES

REFERENCES

REFERENCES

[1] S. E. Thompson, M. Armstrong, C. Auth, M. Alavi, M. Buehler, R. Chau, S. Cea, T. Ghani, G. Glass, T. Hoffman *et al.*, "A 90-nm logic technology featuring strained-silicon," *Electron Devices, IEEE Transactions on*, vol. 51, no. 11, pp. 1790–1797, 2004.

[2] J. Kavalieros, B. Doyle, S. Datta, G. Dewey, M. Doczy, B. Jin, D. Lionberger, M. Metz, W. Rachmady, M. Radosavljevic *et al.*, "Tri-gate transistor architecture with high-k gate dielectrics, metal gates and strain engineering," in *VLSI Technology, 2006. Digest of Technical Papers. 2006 Symposium on.* IEEE, 2006, pp. 50–51.

[3] Y.-K. Choi, N. Lindert, P. Xuan, S. Tang, D. Ha, E. Anderson, T.-J. King, J. Bokor, and C. Hu, "Sub-20 nm cmos finfet technologies," in *Electron Devices Meeting, 2001. IEDM'01. Technical Digest. International.* IEEE, 2001, pp. 19–1.

[4] M. Salmani-Jelodar, S. Kim, K. Ng, and G. Klimeck, "Transistor roadmap projection using predictive full-band atomistic modeling," *Applied Physics Letters*, vol. 105, no. 8, p. 083508, 2014.

[5] M. Salmani Jelodar, S. Kim, K. Ng, and G. Klimeck, "Performance degradation due to thicker physical layer of high k oxide in ultra-scaled mosfets and mitigation through electrostatics design," in *Silicon Nanoelectronics Workshop (SNW), 2014 IEEE.* IEEE, 2014, pp. 16.1–16.2.

[6] S. Mehrotra, S. Kim, T. Kubis, M. Povolotskyi, M. Lundstrom, and G. Klimeck, "Engineering nanowire n-mosfets at lg¡ 8 nm," *IEEE TRANSACTIONS ON ELECTRON DEVICES*, vol. 60, no. 7, p. 2171, 2013.

[7] M. Salmani-Jelodar, S. Steiger, A. Paul, and G. Klimeck, "Model development for lattice properties of gallium arsenide using parallel genetic algorithm," in *Evolutionary Computation (CEC), 2011 IEEE Congress on.* IEEE, 2011, pp. 2429–2435.

[8] K. Mistry, "Transistor scaling: The age of innovation," Apr 2014. [Online]. Available: https://nanohub.org/resources/20880

[9] "International technology roadmap for semiconductors (itrs), howpublished = http://www.itrs.net, note = Accessed: 2013-11-30."

[10] R. M. Martin, *Electronic structure: basic theory and practical methods.* Cambridge university press, 2004.

[11] E. K. Gross and R. M. Dreizler, *Density functional theory.* Springer, 1995, vol. 337.

[12] M. Salmani-Jelodar, A. Paul, T. Boykin, and G. Klimeck, "Calculation of phonon spectrum and thermal properties in suspended <100> $in_xga_{1-X}as$ nanowires," *Journal of Computational Electronics*, vol. 11, no. 1, pp. 22–28, 2012.

[13] T. B. Boykin, M. Luisier, M. Salmani-Jelodar, and G. Klimeck, "Strain-induced, off-diagonal, same-atom parameters in empirical tight-binding theory suitable for [110] uniaxial strain applied to a silicon parametrization," *Physical Review B*, vol. 81, no. 12, p. 125202, 2010.

[14] S. Steiger, M. Salmani-Jelodar, D. Areshkin, A. Paul, T. Kubis, M. Povolotskyi, H.-H. Park, and G. Klimeck, "Enhanced valence force field model for the lattice properties of gallium arsenide," *Physical Review B*, vol. 84, no. 15, p. 155204, 2011.

[15] M. Salmani Jelodar, A. Paul, T. Boykin, and G. Klimeck, "Phonon spectrum and thermal properties of free standing¡ 100¿ and¡ 111¿ ingaas alloy nanowires," in *APS Meeting Abstracts*, vol. 1, 2012, p. 17003.

[16] M. Salmani-Jelodar, Y. Tan, and G. Klimeck, "Single layer mos2 band structure and transport," in *Semiconductor Device Research Symposium (ISDRS), 2011 International*. IEEE, 2011, pp. 1–2.

[17] T. Skotnicki, *MASTAR 4.0 user manual*.

[18] S. Steiger, M. Povolotskyi, H.-H. Park, T. Kubis, and G. Klimeck, "Nemo5: a parallel multiscale nanoelectronics modeling tool," *Nanotechnology, IEEE Transactions on*, vol. 10, no. 6, pp. 1464–1474, 2011.

[19] M. Salmani-Jelodar, S. Mehrotra, H. Ilatikhameneh, and G. Klimeck, "Tunneling paper for lg¡12 nm," *IEEE EDL*, vol. X, no. X, p. X, 2013.

[20] M. Salmani-Jelodar, S. Kim, K. Ng, and G. Klimeck, "Scaling issues and possible solutions for double gate mosfets at the end of itrs," in *International Semiconductor Device Research Symposium (ISDRS 2013)*. IEEE, 2013.

[21] S. Kim, M. Salmani-Jelodar, K. Ng, and G. Klimeck, "Quantum corrected drift-diffusion simulation for prediction of cmos scaling," in *Device Research Conference (DRC), 2013 71st Annual*. IEEE, 2013, pp. 119–120.

[22] M. Salmani-Jelodar, S. Kim, K. Ng, and G. Klimeck, "On the scaling issues and solutions for double gate mosfets at the end of itrs," in *TECHCON 2013*.

[23] S. H. Park, Y. Liu, N. Kharche, M. S. Jelodar, G. Klimeck, M. S. Lundstrom, and M. Luisier, "Performance comparisons of iii–v and strained-si in planar fets and nonplanar finfets at ultrashort gate length (12 nm)," *Electron Devices, IEEE Transactions on*, vol. 59, no. 8, pp. 2107–2114, 2012.

[24] M. Luisier, M. Lundstrom, D. A. Antoniadis, and J. Bokor, "Ultimate device scaling: Intrinsic performance comparisons of carbon-based, ingaas, and si field-effect transistors for 5 nm gate length," in *Electron Devices Meeting (IEDM), 2011 IEEE International*. IEEE, 2011, pp. 11–2.

[25] H. Kawaura, T. Sakamoto, and T. Baba, "Observation of source-to-drain direct tunneling current in 8 nm gate electrically variable shallow junction metal–oxide–semiconductor field-effect transistors," *Applied Physics Letters*, vol. 76, no. 25, pp. 3810–3812, 2000.

[26] J. Wang and M. Lundstrom, "Does source-to-drain tunneling limit the ultimate scaling of mosfets?" in *Electron Devices Meeting, 2002. IEDM'02. International*. IEEE, 2002, pp. 707–710.

[27] K. J. Kuhn, "Considerations for ultimate cmos scaling," *Electron Devices, IEEE Transactions on*, vol. 59, no. 7, pp. 1813–1828, 2012.

[28] N. Seoane and A. Martinez, "A detailed coupled-mode-space non-equilibrium green's function simulation study of source-to-drain tunnelling in gate-all-around si nanowire metal oxide semiconductor field effect transistors," *Journal of Applied Physics*, vol. 114, no. 10, p. 104307, 2013.

[29] S. S. Sylvia, H.-H. Park, M. Khayer, K. Alam, G. Klimeck, and R. K. Lake, "Material selection for minimizing direct tunneling in nanowire transistors," *Electron Devices, IEEE Transactions on*, vol. 59, no. 8, pp. 2064–2069, 2012.

[30] M. Salmani-Jelodar, S. Kim, K. Ng, and G. Klimeck, "On the scaling issues and solutions for double gate mosfets at the end of itrs roadmap," in *ISDRS2013*.

[31] M. Salmani-Jelodar, I. H. Mehrotra, Saumitra, and G. Klimeck, "Design guidelines for sub 12 nm mosfets," *Nanotechnology, IEEE Transactions on*, vol. accepted for publication, 2014.

[32] M. Salmani-Jelodar, H. Ilatikhameneh, S. Kim, K. Ng, and G. Klimeck, "Optimum high-k oxide for the best performance of ultra-scaled double-gate mosfets," *submitted to Nanotechnology, IEEE Transactions on*.

[33] M. Salmani-Jelodar, B. Behin-Aein, S. Mehrotra, G. Klimeck, and Z. Krivokapic, "Quantum confined p-type sige mosfets," *Submitted to Nanotechnology, IEEE Transactions on*.

[34] G. Klimeck, R. C. Bowen, T. B. Boykin, C. Salazar-Lazaro, T. A. Cwik, and A. Stoica, "Si tight-binding parameters from genetic algorithm fitting," *Superlattices and Microstructures*, vol. 27, no. 2, pp. 77–88, 2000.

[35] Y. Y. Peter and M. Cardona, *Fundamentals of semiconductors: physics and materials properties*. Springer, 2010.

[36] P. Keating, "Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure," *Physical Review*, vol. 145, no. 2, p. 637, 1966.

[37] A. Paul, M. Luisier, and G. Klimeck, "Modified valence force field approach for phonon dispersion: from zinc-blende bulk to nanowires," *Journal of computational electronics*, vol. 9, no. 3-4, pp. 160–172, 2010.

[38] M. S. Jelodar, S. N. Fakhraie, F. Montazeri, S. M. Fakhraie, and M. N. Ahmadabadi, "A representation for genetic-algorithm-based multiprocessor task scheduling," in *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*. IEEE, 2006, pp. 340–347.

[39] E. Kane, "Phonon spectra of diamond and zinc-blende semiconductors," *Physical Review B*, vol. 31, no. 12, p. 7865, 1985.

[40] R. M. Martin, "Elastic properties of zns structure semiconductors," *Physical Review B*, vol. 1, no. 10, p. 4005, 1970.

[41] O. L. Lazarenkova, P. von Allmen, F. Oyafuso, S. Lee, and G. Klimeck, "An atomistic model for the simulation of acoustic phonons, strain distribution, and grüneisen coefficients in zinc-blende semiconductors," *Superlattices and microstructures*, vol. 34, no. 3, pp. 553–556, 2003.

[42] M. Musgrave and J. Pople, "A general valence force field for diamond," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 268, no. 1335, pp. 474–484, 1962.

[43] Z. Sui and I. P. Herman, "Effect of strain on phonons in si, ge, and si/ge heterostructures," *Physical Review B*, vol. 48, no. 24, p. 17938, 1993.

[44] A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. Ipatova, *Theory of lattice dynamics in the harmonic approximation*. Academic press New York, 1963, vol. 3.

[45] D. Patrick, P. Green, and T. York, "A distributed genetic algorithm environment for unix workstation clusters," in *Genetic Algorithms in Engineering Systems: Innovations and Applications, 1997. GALESIA 97. Second International Conference On (Conf. Publ. No. 446)*. IET, 1997, pp. 69–74.

[46] T. M. Mitchell, *Machine learning*. McGraw-Hill Boston, MA:, 1997.

[47] T. Back, U. Hammel, and H.-P. Schwefel, "Evolutionary computation: Comments on the history and current state," *Evolutionary computation, IEEE Transactions on*, vol. 1, no. 1, pp. 3–17, 1997.

[48] C. C. Pettey and M. R. Leuze, "A theoretical investigation of a parallel genetic algorithm." in *ICGA*, vol. 3, 1989, pp. 398–405.

[49] D. Strauch and B. Dorner, "Phonon dispersion in gaas," *Journal of Physics: Condensed Matter*, vol. 2, no. 6, p. 1457, 1990.

[50] N. Orlova, "X-ray thermal diffuse scattering measurements of the [100] and [111] phonon dispersion curves of indium arsenide," *physica status solidi (b)*, vol. 93, no. 2, pp. 503–509, 1979.

[51] ioffe. (2013, Dec.). [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/GaInAs/

[52] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck, "Atomistic simulation of nanowires in the sp^{3} d^{5} s^{*} tight-binding formalism: From boundary conditions to strain calculations," *Physical Review B*, vol. 74, no. 20, p. 205323, 2006.

[53] Abinit. (2013, Dec.). [Online]. Available: http://http://www.abinit.org/

[54] R. Kim and M. S. Lundstrom, "Physics of carrier backscattering in one-and two-dimensional nanotransistors," *Electron Devices, IEEE Transactions on*, vol. 56, no. 1, pp. 132–139, 2009.

[55] C. Jeong, D. A. Antoniadis, and M. S. Lundstrom, "On backscattering and mobility in nanoscale silicon mosfets," *Electron Devices, IEEE Transactions on*, vol. 56, no. 11, pp. 2762–2769, 2009.

[56] M. Lundstrom, "Elementary scattering theory of the si mosfet," *Electron Device Letters, IEEE*, vol. 18, no. 7, pp. 361–363, 1997.

[57] P. Palestri, D. Esseni, S. Eminente, C. Fiegna, E. Sangiorgi, and L. Selmi, "Understanding quasi-ballistic transport in nano-mosfets: part i-scattering in the channel and in the drain," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2727–2735, 2005.

[58] K. Natori, "Ballistic mosfet reproduces current-voltage characteristics of an experimental device," *Electron Device Letters, IEEE*, vol. 23, no. 11, pp. 655–657, 2002.

[59] P. Palestri, R. Clerc, D. Esseni, L. Lucci, and L. Selmi, "Multi-subband-monte-carlo investigation of the mean free path and of the kt layer in degenerated quasi ballistic nanomosfets," in *Electron Devices Meeting, 2006. IEDM'06. International*. IEEE, 2006, pp. 1–4.

[60] A. Khakifirooz, K. Cheng, A. Reznicek, T. Adam, N. Loubet, H. He, J. Kuss, J. Li, P. Kulkarni, S. Ponoth *et al.*, "Scalability of extremely thin soi (etsoi) mosfets to sub-20-nm gate length," *Electron Device Letters, IEEE*, vol. 33, no. 2, pp. 149–151, 2012.

[61] T. Hiramoto, G. Tsutsui, K. Shimizu, and M. Kobayashi, "Transport in ultra-thin-body soi and silicon nanowire mosfets," in *Semiconductor Device Research Symposium, 2007 International*. IEEE, 2007, pp. 1–2.

[62] K. Uchida, J. Koga, and S.-i. Takagi, "Experimental study on electron mobility in ultrathin-body silicon-on-insulator metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 102, no. 7, pp. 074 510–074 510, 2007.

[63] K. Shimizu, G. Tsutsui, and T. Hiramoto, "Experimental study on mobility universality in (100) ultra thin body nmosfet with soi thickness of 5nm," in *International SOI Conference, 2006 IEEE*. IEEE, 2006, pp. 159–160.

[64] K. Cheng, A. Khakifirooz, P. Kulkarni, S. Ponoth, B. Haran, A. Kumar, T. Adam, A. Reznicek, N. Loubet, H. He *et al.*, "Etsoi cmos for system-on-chip applications featuring 22nm gate length, sub-100nm gate pitch, and 0.08μm 2 sram cell," in *VLSI Technology (VLSIT), 2011 Symposium on*. IEEE, 2011, pp. 128–129.

[65] D. J. Griffiths, *Introduction to quantum mechanics*. Pearson Education India, 2005.

[66] S. R. Mehrotra, M. Povolotskyi, D. C. Elias, T. Kubis, J. J. Law, M. J. Rodwell, and G. Klimeck, "Simulation study of thin-body ballistic n-mosfets involving transport in mixed γ-l valleys," 2013.

[67] M. Luisier, "Performance comparison of gasb, strained-si, and ingaas double-gate ultrathin-body n-fets," *Electron Device Letters, IEEE*, vol. 32, no. 12, pp. 1686–1688, 2011.

[68] K. Alam, S. Takagi, and M. Takenaka, "Analysis and comparison of l-valley transport in gaas, gasb, and ge ultrathin-body ballistic nmosfets," *Electron Devices, IEEE Transactions on*, vol. 60, no. 12, pp. 4213–4218, 2013.

[69] A. Rahman and M. S. Lundstrom, "A compact scattering model for the nanoscale double-gate mosfet," *Electron Devices, IEEE Transactions on*, vol. 49, no. 3, pp. 481–489, 2002.

[70] J. E. Fonseca, T. Kubis, M. Povolotskyi, B. Novakovic, A. Ajoy, G. Hegde, H. Ilatikhameneh, Z. Jiang, P. Sengupta, Y. Tan *et al.*, "Efficient and realistic device modeling from atomic detail to the nanoscale," *Journal of Computational Electronics*, vol. 12, no. 4, pp. 592–600, 2013.

[71] S. Migita, Y. Morita, M. Masahara, and H. Ota, "Electrical performances of junctionless-fets at the scaling limit (l ch= 3 nm)," in *Electron Devices Meeting (IEDM), 2012 IEEE International*. IEEE, 2012, pp. 8–6.

[72] ——, "Fabrication and demonstration of 3-nm-channel-length junctionless field-effect transistors on silicon-on-insulator substrates using anisotropic wet etching and lateral diffusion of dopants," *Japanese Journal of Applied Physics*, vol. 52, no. 4, p. 04CA01, 2013.

[73] C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman *et al.*, "A 22nm high performance and low-power cmos technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density mim capacitors," in *VLSI Technology (VLSIT), 2012 Symposium on*. IEEE, 2012, pp. 131–132.

[74] M. Salmani-Jelodar, S. Kim, K. Ng, and G. Klimeck, "Transistor roadmap projection using predictive full-band atomistic modeling," *Applied Physics Letters*, vol. 105, no. 8, pp. 083 508–083 511, 2014.

[75] Q. Xie, J. Xu, and Y. Taur, "Review and critique of analytic models of mosfet short-channel effects in subthreshold," *Electron Devices, IEEE Transactions on*, vol. 59, no. 6, pp. 1569–1579, 2012.

[76] Q. Xie, C.-J. Lee, J. Xu, C. Wann, J.-C. Sun, and Y. Taur, "Comprehensive analysis of short-channel effects in ultrathin soi mosfets," 2013.

[77] M. Salmani Jelodar, S. Kim, K. Ng, and G. Klimeck, "Performance degradation due to thicker physical layer of high k oxide in ultra-scaled mosfets and mitigation through electrostatics design," in *Silicon Nanoelectronics Workshop (SNW), 2014 IEEE*. IEEE, 2014, pp. 16.1–16.2.

[78] B. Cheng, M. Cao, R. Rao, A. Inani, P. Vande Voorde, W. M. Greene, J. M. Stork, Z. Yu, P. M. Zeitzoff, and J. C. Woo, "The impact of high-κ gate dielectrics and metal gate electrodes on sub-100 nm mosfets," *Electron Devices, IEEE Transactions on*, vol. 46, no. 7, pp. 1537–1544, 1999.

[79] D. J. Frank, Y. Taur, and H.-S. Wong, "Generalized scale length for two-dimensional effects in mosfets," *Electron Device Letters, IEEE*, vol. 19, no. 10, pp. 385–387, 1998.

[80] C. Manoj and V. Ramgopal Rao, "Impact of high-k gate dielectrics on the device and circuit performance of nanoscale finfets," *Electron Device Letters, IEEE*, vol. 28, no. 4, pp. 295–297, 2007.

[81] H.-J. Cho, K.-I. Seo, W. Jeong, Y.-H. Kim, Y. Lim, W. Jang, J. Hong, S. Suk, M. Li, C. Ryou *et al.*, "Bulk planar 20nm high-k/metal gate cmos technology platform for low power and high performance applications," in *Electron Devices Meeting (IEDM), 2011 IEEE International*. IEEE, 2011, pp. 15.1.1 – 15.1.4.

[82] M. Luisier, A. Schenk, and W. Fichtner, "Three-dimensional modeling of gate leakage in si nanowire transistors," in *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*. IEEE, 2007, pp. 733–736.

[83] M. Luisier and A. Schenk, "Two-dimensional tunneling effects on the leakage current of mosfets with single dielectric and high-gate stacks," *Electron Devices, IEEE Transactions on*, vol. 55, no. 6, pp. 1494–1501, 2008.

[84] S. Steiger, M. Povolotskyi, H.-H. Park, T. Kubis, and G. Klimeck, "Nemo5: a parallel multiscale nanoelectronics modeling tool," *Nanotechnology, IEEE Transactions on*, vol. 10, no. 6, pp. 1464–1474, 2011.

[85] Y.-C. Yeo, T.-J. King, and C. Hu, "Direct tunneling leakage current and scalability of alternative gate dielectrics," *Applied Physics Letters*, vol. 81, no. 11, pp. 2091–2093, 2002.

[86] J. Robertson, "Band offsets of wide-band-gap oxides and implications for future electronic devices," *Journal of Vacuum Science & Technology B*, vol. 18, no. 3, pp. 1785–1791, 2000.

[87] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. John Wiley & Sons, 2006.

[88] C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman *et al.*, "A 22nm high performance and low-power cmos technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density mim capacitors," in *VLSI Technology (VLSIT), 2012 Symposium on*. IEEE, 2012, pp. 131–132.

[89] L. Gomez, C. Ni Chleirigh, P. Hashemi, and J. Hoyt, "Enhanced hole mobility in high ge content asymmetrically strained-sige p-mosfets," *Electron Device Letters, IEEE*, vol. 31, no. 8, pp. 782–784, 2010.

[90] S. Takagi, T. Tezuka, T. Irisawa, S. Nakaharai, T. Numata, K. Usuda, N. Sugiyama, M. Shichijo, R. Nakane, and S. Sugahara, "Device structures and carrier transport properties of advanced cmos using high mobility channels," *Solid-State Electronics*, vol. 51, no. 4, pp. 526–536, 2007.

[91] A. Villalon, C. Le Royer, S. Cristoloveanu, M. Casse, D. Cooper, J. Mazurier, B. Previtali, C. Tabone, P. Perreau, J.-M. Hartmann *et al.*, "High-performance ultrathin body c-sige channel fdsoi pmosfets featuring sige source and drain: Tuning, variability, access resistance, and mobility issues," *Electron Devices, IEEE Transactions on*, vol. 60, no. 5, pp. 1568–1574, 2013.

[92] C. Smith, H. Adhikari, S.-H. Lee, B. Coss, S. Parthasarathy, C. Young, B. Sassman, M. Cruz, C. Hobbs, P. Majhi *et al.*, "Dual channel finfets as a single high-k/metal gate solution beyond 22nm node," in *Electron Devices Meeting (IEDM), 2009 IEEE International*. IEEE, 2009, pp. 1–4.

[93] T. B. Boykin, G. Klimeck, R. C. Bowen, and F. Oyafuso, "Diagonal parameter shifts due to nearest-neighbor displacements in empirical tight-binding theory," *Physical Review B*, vol. 66, no. 12, p. 125207, 2002.

[94] A. Paul, S. Mehrotra, M. Luisier, and G. Klimeck, "Performance prediction of ultrascaled sige/si core/shell electron and hole nanowire mosfets," *Electron Device Letters, IEEE*, vol. 31, no. 4, pp. 278–280, 2010.

[95] M. Lundstrom and Z. Ren, "Essential physics of carrier transport in nanoscale mosfets," *IEEE Transactions on Electron Devices*, vol. 49, no. 1, pp. 133–141, 2002.

[96] K. J. Kuhn, A. Murthy, R. Kotlyar, and M. Kuhn, "(invited) past, present and future: Sige and cmos transistor scaling," *ECS Transactions*, vol. 33, no. 6, pp. 3–17, 2010.

[97] X. Xiao, C. Liu, J. Sturm, L. Lenchyshyn, M. Thewalt, R. Gregory, and P. Fejes, "Quantum confinement effects in strained silicon-germanium alloy quantum wells," *Applied physics letters*, vol. 60, no. 17, pp. 2135–2137, 1992.

[98] S. Kim, M. Luisier, A. Paul, T. B. Boykin, and G. Klimeck, "Full three-dimensional quantum transport simulation of atomistic interface roughness in silicon nanowire fets," *Electron Devices, IEEE Transactions on*, vol. 58, no. 5, pp. 1371–1380, 2011.

VITA

VITA

Mehdi Salmani Jelodar is a PhD candidate in electrical engineering at Purdue University. His doctoral studies are in the field of computational nanotechnology. He is conducting research to predict and optimize the effects of novel materials and geometries in ultra-scaled transistors for future technology. Mehdi is an active member of ITRS and has delivered projections for SOI and multi-gate logic devices for the 2013-2028 ITRS. He has published more than 40 papers during his graduate studies and won 3 best paper/poster awards. He was elected as the outstanding graduate student of 2013 among +800 graduate students of the school of electrical and computer engineering. In summer 2013, Mehdi worked in semiconductor technology development at GLOBALFOUNDRIES in Silicon Valley. Mehdi obtained his M.Sc. degree in computer engineering from the University of Tehran, Iran. In Iran, he also co-founded Danesh Gostaran Sepanta, a technology startup, and served as its technical manager for five years.