Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Neerav Kharche

Entitled Atomistic Modeling of Electronic Structure and Transport in Disordered Nanostructures

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

G. Klimeck, Co-Chair
Chair

S. Datta

T. B. Boykin, Co-Chair

M. S. Lundstrom

R. G. Reifenberger

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): G. Klimeck, Co-Chair

Approved by: V. Balakrishnan 1/27/10
Head of the Graduate Program Date

Graduate School Form 20
(Revised 10/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation: Atomistic Modeling of Electronic Structure and Transport in Disordered Nanostructures

For the degree of Doctor of Philosophy

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Neerav Kharche
Signature of Candidate

1/26/10
Date

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

ATOMISTIC MODELING OF ELECTRONIC STRUCTURE AND TRANSPORT IN DISORDERED NANOSTRUCTURES

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Neerav Kharche

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2010

Purdue University

West Lafayette, Indiana

To my family …

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Professor Klimeck, for valuable insights he provided for this research. I would like to thank him for providing all the necessary computing resources. He introduced me to high performance computing and its applications to the problem solving in Physics and Engineering which will be invaluable experience for my future career pursuits.

I would also like to express my deep appreciation to my co-advisor, Professor Timothy Boykin, for providing invaluable guidance on the tight-binding theory. His theoretical approaches are always mathematically rigorous and developed from the point of view of efficient implementation into computer programs. I consider myself fortunate to have an opportunity to work with him.

I would like to thank Dr. Mathieu Luisier for providing quantum transport simulators which were used to model electron transport in AlGaAs and SiGe nanowires and InAs quantum well field effect transistors. I am especially thankful to him for the guidance he provided for InAs quantum well field effect transistors work.

I would also like to thank Professor Supriyo Datta, Professor Mark Lundstrom, and Professor Ronald Reifenberger for serving on my advisory committee.

I am thankful to Professor Shaikh Ahmed for his valuable advice. I am grateful to Dr. Rajib Rahman, Samarth Agarwal, Abhijeet Paul, Srikant Srinivasan, Dr. Neophytos Neophytou, and Dr. Bhaskaran Muralidharan for valuable scientific discussions. I would also like to thank Hoon Ryu and Sunhee Lee for their advice regarding computational algorithms. I would like to thank all other members of “Klimeck group”, specifically Seung Hyun Park, Seongmin Kim, Zhengping Jiang, Parijat Sengupta, Saumitra Mehrotra, Mehdi Salmani, and Muhammad Usman. I would also like to thank all my colleagues in EE350 specifically Dr. Ninad Pimparkar, Dr. Pradeep Nair, and Yang Liu. I

consider myself fortunate to have an opportunity to work with such talented and dedicated people. I would like to thank Cheryl Haines for her help in scheduling meetings and making conference travel arrangements.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

ABSTRACT

Kharche, Neerav. Ph.D., Purdue University, May 2010. Atomistic Modeling of Electronic Structure and Transport in Disordered Nanostructures. Major Professors: Gerhard Klimeck and Timothy Boykin.

As the Si-CMOS technology approaches the end of the International Technology Roadmap for Semiconductors (ITRS), the semiconductor industry faces a formidable challenge to continue the transistor scaling according to Moore's law. To continue the scaling of classical devices, alternative channel materials such as SiGe, carbon nanotubes, nanowires, and III-V based materials are being investigated along with novel 3D device geometries. Researchers are also investigating radically new quantum computing devices, which are expected to perform calculations faster than the existing classical Si-CMOS based structures. Atomic scale disorders such as interface roughness, alloy randomness, non-uniform strain, and dopant fluctuations are routinely present in the experimental realization of such devices. These disorders now play an increasingly important role in determining the electronic structure and transport properties as device sizes enter the nanometer regime. This work employs the atomistic tight-binding technique, which is ideally suited for modeling systems with local disorders on an atomic scale.

High-precision multi-million atom electronic structure calculations of (111) Si surface quantum wells and (100) SiGe/Si/SiGe heterostructure quantum wells are performed to investigate the modulation of valley splitting induced by atomic scale disorders. The calculations presented here resolve the existing discrepancies between theoretically predicted and experimentally measured valley splitting, which is an important design parameter in quantum computing devices. Supercell calculations and the zone-unfolding method are used to compute the bandstructures of inhomogeneous nanowires made of

AlGaAs and SiGe and their connection with the transmission coefficients computed using non-equilibrium Green's function method is established. A unified picture of alloy nanowires emerges, in which the nanodevice (transmission) and nanomaterials (bandstructure) viewpoints complement each other and illuminate the interesting physics of these disordered nanostructures that otherwise can not be explained using the traditional averaging methods such as the virtual crystal approximation. Finally, a multi-scale modeling approach is employed, which combines the atomistic tight-binding method to compute the electronic structure and the real-space effective mass based quantum transport model including gate leakage to simulate the three terminal characteristics of III-V quantum well field effect transistors (QWFETs). The simulation methodology has been benchmarked against experimental data and it is then applied to investigate the logic performance of ultra-scaled III-V QWFETs with high mobility InAs channels.

1. INTRODUCTION

1.1. Background: Device scaling and disorders in nanoscale devices

The shrinking of the dimensions of individual semiconductor devices has enabled the exponential increase in the density of integrated circuits predicted by Moore's law [1]. The Si based Metal Oxide Semiconductor Field Effect Transistor (MOSFET) has been the basic building block of integrated circuits over the past three decades. [1]. However, recently many researchers have pointed out that the conventional MOSFET scaling is already facing its limits and will slow down in near future. The physical gate length of Si-transistors that are utilized in the current 45-nm node is about 30 nm [1]. As the critical device dimensions such as gate length and oxide thickness reach physical limitations, it is extremely challenging from the fabrication point of view to maintain the dimensional integrity of devices. Besides reducing the area per transistor on the die, scaling down of the gate length reduces the time it takes to move charge (electrons) from the source to the drain, increasing the switching speed of the transistor and hence the switching frequency at which logic circuits operate. As the device sizes scale and switching frequencies increase the power consumption per chip increases and the need to develop low power devices emerges [2, 3].

To continue device scaling, the semiconductor industry is facing the challenge to invent fundamentally new approaches to device design [1]. The classical approaches being pursued include migration to 3D device geometries from 2D planar geometries to improve the electrostatic control of the channel [3-5] and the use of the high mobility materials such as III-V compounds [6-13] and carbon nanotubes [14-17], that can operate at lower power supply voltage. Fundamentally new ways of information storing, processing and transporting such as quantum computing architectures are also being actively researched [18-20].

Regardless of the specific computing architecture, it is clear that the future integrated circuits will have critical device dimensions of few tens of nanometers or even smaller. In this regime, the atomistic granularity of constituent materials cannot be neglected: Effects of atomistic strain, surface roughness, dopant fluctuations, the underlying crystal symmetries, or distortions of the crystal lattice can have a dramatic impact on the device performance [21-26]. Such effects make characteristics of a device different from other devices on the same chip. These fluctuations in device characteristics deteriorate the circuit performance. However, such fluctuations if calculated precisely can be taken into account in variation-aware circuit design techniques [27]. This work employs the atomistic tight-binding technique in the supercell framework, which is ideally suited for modeling systems with local disorders on an atomic scale.

1.2. Supercell approach to model disorder

The knowledge of the electronic structure and transport characteristics of nanoscale semiconductor devices is the first and essential step toward the interpretation and understanding of the experimental data and reliable device design at the nanometer scale. The effective mass approximation has been extensively used to investigate the electron transport in semiconductor devices. The most basic effective mass model uses the bulk effective masses and the band gaps, which work well for “large” devices. However, in nanoscale devices band gap and effective mass values differ from their bulk values because of strong physical and electrostatic confinement. In more advanced models, the effective masses and the band gaps in nanoscale devices are derived from the bandstructure that takes into account the effects such as finite dimensions, electrostatics, and local material variations on the atomic scale [28].

The bandstructure calculations are typically performed within a plane-wave basis in an infinitely periodic system. However, as devices are scaled to nanoscale, sharp variations on atomic scale play an increasingly important role in determining the electronic structure and transport properties. At the nanoscale, methods that employ local basis sets are theoretically more intuitive and practically more convenient compared to plane-wave basis set methods [29, 30]. Local basis set methods such as semi-empirical tight-binding

have smaller computational requirements than corresponding pseudopotential methods for the calculation of electronic structure in multimillion-atom systems [31-40].

Traditional tight-binding bandstructure calculations in periodic systems are performed by choosing the smallest repeating unit cell along the periodic directions of the semiconductor crystal. However, nanostructures are seldom perfect. Examples include quantum dots or wires made of alloys or having rough surfaces [21, 26, 40], nanoscale FETs with inhomogeneous strain distributions [7, 8, 41], or simply three-dimensional random alloys [22]. In disordered structures translational symmetry is broken and the small cells differ from each other. The traditional averaging methods such as virtual crystal approximation (VCA) can not account for bandstructure effects of local atomic variations. This work uses the supercell framework to model local variations on the atomic scale. A supercell contains many small cells which are different from each other in terms of atom/strain distribution and thus the supercell itself has a random disorder in it. If the supercell is sufficiently large, it will represent the true disorder in the device. However, due to Brillouin zone-folding such supercell calculations only deliver the energy spectrum which can only be used to extract absolute band-minima of the disordered nanostructures. Supercell calculations alone can not deliver the transport parameters such as effective masses and relative locations of valley-minima. The zone-unfolding method is used to unfold the supercell bandstructure, which is then used to extract the transport parameters such as effective mass and relative locations of valley-minima [22, 42]. The transport parameters thus calculated include the effect of local atomic scale variations in the nanostructure.

1.3. Outline of the dissertation

In chapters 2 to 4 the electronic structure calculations of (111) and (001) Si quantum wells are presented. These chapters are based on references [40, 43, 44]. (111) and (001) Si surfaces typically reconstruct to minimize the surface energy, which results in the formation of staircase like features as shown in Figure 1.1(a) [45-48]. The electronic structure calculations that do not take such staircase like features into account predict the valley degeneracy of (111) Si quantum wells to be 6 [49]. However, experiments show

lower 2-fold and higher 4-fold valley degeneracies [45]. The supercell formalism can model atomic scale variations shown in Figure 1.1(a). In chapter 2, it is shown that the surface miscut breaks the 6-fold valley degeneracy of (111) Si quantum well into lower 2-fold and higher 4-fold valley degeneracies. It is also pointed out that, in the effective mass picture, the lifting of 6-fold valley degeneracy in (111) Si quantum well is caused by different effective masses of valleys in the confinement direction.

Degenerate valleys can couple and split in the presence of sharp potential variations on an atomic scale [50-52]. Such splitting is absent in (111) Si quantum wells because valley coupling in the presence of [111] directed confinement potential is negligible [51]. However, certain valleys strongly couple in the presence of [001] directed confinement, which causes valley splitting in (001) Si quantum wells [51]. Valley splitting in (001) Si quantum wells is being actively researched because recently several research groups have proposed quantum computing architectures based on (001) Si quantum wells [53-55]. In quantum computing devices, valley splitting is a critical design parameter, which controls the qubit decoherence time [55-57]. Chapter 4 of this report focuses on the valley splitting in SiGe/Si/SiGe quantum wells, which is the fabrication platform for the spin based quantum computing architecture proposed in [54, 55]. SiGe/Si/SiGe quantum wells are typically grown on miscut substrates to ensure the uniform growth of the heterostructure. Miscut (001) Si surfaces have irregular step-like morphology as shown in Figure 1.1(a) [48]. SiGe/Si interface is irregular as shown in Figure 1.1(b) due to alloy randomness of SiGe buffer. Chapter 3 develops the understanding of the origin of valley splitting in flat and miscut (001) quantum wells, where the oscillations in valley splitting

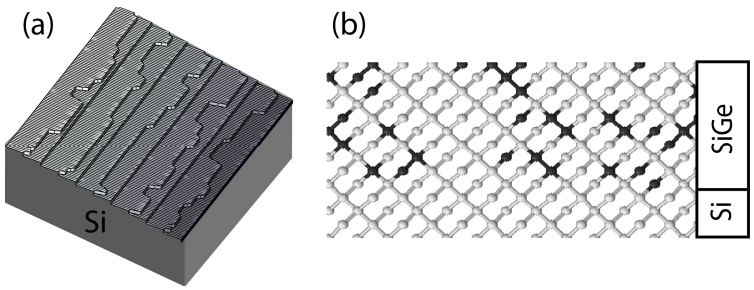

Figure 1.1. Schematic representation of (a) the surface roughness disorder on (111) and miscut (001) Si surface and (b) the alloy disorder in SiGe buffer layer in SiGe/Si/SiGe heterostructure.

with confinement dimensions are explained qualitatively and quantitatively. Computationally more intensive calculations that include step and alloy disorders (Figure 1.1) are presented and compared with the experimental data in chapter 4. It is found that, models that do not include miscuts [58, 59] over-predict the value of the valley splitting by an order of magnitude compared with the experimental data [60] while perfect miscut quantum well models [24, 61] under-predict the observed valley splitting by an order of magnitude. The supercell calculations that include step and alloy disorders are found to predict the experimentally observed valley splitting within variations induced by the presence of the disorders [24]. The atomistic representation of the disorders is found to be an essential ingredient in high precision calculations of valley splitting in heterostructures.

Chapters 5 and 6 present a methodology to analyze the electronic structure and transport characteristics of disordered materials/devices. These chapters are based on references [22, 23, 25, 42]. To continue the device scaling several industry and academic research groups are investigating Field Effect Transistors (FETs) with novel-channel materials which are expected to perform faster and dissipate less power than conventional Si MOSFETs. In the nanometer regime, the variations on the atomic scale of the constituent material (Figure 1.2) have a dramatic impact on the device performance and such variations should be included in the simulation models [21, 23, 25, 26, 62]. In chapter 5, the zone-unfolding method [22, 42], which can be used to compute the "approximate" bandstructure of disordered nanowires is presented. The zone-unfolding method is an efficient way to extract the bandstructure from the folded supercell eigenspectrum. This method can be used to compute bandstructure of disordered

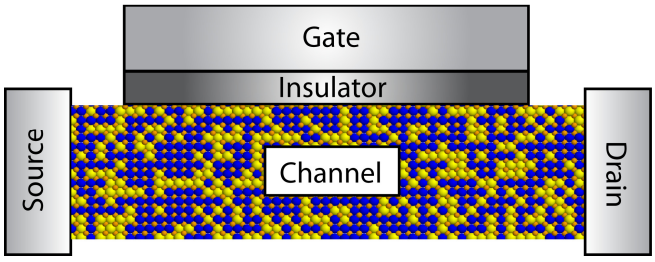

Figure 1.2. Schematic of a FET device with the channel made up of a disordered material. At the nanoscale, disorder in the channel has pronounced effects on the device characteristics.

nanostructures including nanowires, quantum wells and bulk alloys. This method can also treat variety of disorders such as alloy disorder, inhomogeneous strain, and surface roughness.

In chapter 6, the zone-unfolding method of chapter 5 is used to compute bandstructures of AlGaAs and SiGe nanowires and compared with the transmission coefficients. This chapter is based on references [23, 25]. AlGaAs nanowires have atom disorder due to random placement of Al atoms in GaAs lattice but no strain disorder because Al-As and Ga-As bond lengths are almost same. SiGe nanowires have both random atom disorder and inhomogeneous strain disorder due to random placement of Ge atoms in Si lattice and different Si-Si and Ge-Ge bond lengths. The "approximate" bandstructures obtained from supercell calculations and zone-unfolding are compared with the transmission coefficients through the same nanowires. As opposed to smooth step-like transmission through ideal nanowires, the transmission coefficients in disordered nanowires are noisy and smaller in magnitude. Transmission coefficients of disordered nanowires show peculiar peaks that can be identified as the localized states in the nanowire. The supercell approach provides a unified picture of alloy nanowires, in which the nanodevice (transmission) and nanomaterials (bandstructure) viewpoints complement each other and illuminate the interesting physics of these disordered nanostructures that otherwise can not be explained using the traditional averaging methods such as the virtual crystal approximation.

In chapter 7, the supercell approach to the bandstructure calculation is coupled with the effective mass based quantum transport model to develop a multiscale device modeling approach and it is applied to investigate the scaling performance of composite InGaAs/InAs/InGaAs multi-quantum-well channel Quantum Well Field Effect Transistors (QWFETs). This chapter is based on reference [63]. In the multiscale approach, the bandstructure of the multi-quantum-well channel is computed using atomistic tight-binding method. The tight-binding calculations are performed on large supercells to accurately include the effect of the inhomogeneous strain and atomistic disorder in the InGaAs buffer layer. The effective masses are extracted from the tight-binding bandstructure and used in an effective mass based quantum transport simulator to

obtain current-voltage characteristics of QWFETs. The effective mass approach is chosen to simulate the transport because it is computationally too intensive to include the gate leakage in the atomistic approach. A very good quantitative agreement of the simulations with the experimental data [8] on QWFETs with gate lengths ranging from 30 to 50 nm is demonstrated. The simulation methodology is then used to investigate the scaling performance of InAs QWFETs. It is shown that such devices can be successfully scaled to 20 nm gate lengths and that the best performance can be achieved in thin InAs channel devices by reducing the insulator thickness to improve the gate control, increasing the gate work function to suppress the gate leakage, and maintaining a flat gate geometry.

Finally, summary of this work is presented in chapter 8.

2. VALLEY DEGENERACIES IN (111) SI QUANTUM WELLS

2.1. Introduction

Silicon nanostructures exhibit a plethora of interesting physical phenomena due to the 6 fold valley degeneracy of the bulk conduction band. Silicon devices are being pursued for spin based quantum computing [53, 64] and spintronics [65] due to their scaling potential and integratability within the industrial nanoelectronic infrastructure. Relative energies and degeneracies of spin and valley states are critical for device operation in these novel computing architectures [53, 64, 65], and conventional metal–oxide–semiconductor field–effect transistors (MOSFETs) that often involve the formation of a 2 dimensional electron gas (2DEG) at the semiconductor-insulator interface. Valley degeneracy of the 2DEG is highly dependent on the interface orientation. (100) Si quantum wells show lower 2 fold and raised 4 fold valley degeneracy while (110) Si quantum wells show lower 4 fold and raised 2 fold valley degeneracy. The origin of these valley degeneracies is well understood and the experimental observations are in agreement with the effective mass based theoretical predictions [51].

Valley degeneracy in (111) Si quantum wells should be 6 according to standard effective mass theory. Experimental measurements on (111) Si/SiO_2 MOSFETs, however, show a conflicting valley degeneracy of 2 and 4 [51, 66, 67]. Recently 2-4 valley splitting has also been observed in magneto-transport measurements performed on hydrogen terminated (111) Si/vacuum field effect transistors [45]. Previously proposed theory of local strain domains [66] can not explain this splitting since the Si-vacuum interface is stress free. The splitting is also unlikely to be a many-body phenomenon [68, 69].

Careful imaging of the surface morphology shows the presence of mono-atomic steps (miscut) on the (111) Si surface [45, 46] as well as at the Si/SiO_2 interface in (111) Si MOSFETs [46, 47]. Atomistic models such as tight-binding are needed to accurately model the electronic structure of miscut quantum wells. Through systematic tight-binding calculations of flat and miscut (111) Si quantum wells we show that the surface miscut leads to the 2-4 degeneracy breaking and resolve the conflict between theory and experimental observations. To reduce the computational burden associated with searching the whole Brillouin-zone for valley minima, an effective mass based valley-projection model [49] tailored to miscut (111) surfaces is used. Electronic structure calculations are performed using the general purpose NEMO-3D Code [39].

2.2. (111) surface model

Hydrogen terminated (111) Si surfaces as well as the (111) Si/SiO_2 interfaces have mono-atomic steps (Figure 2.1(a,b)). For simplicity, the steps are assumed to run perpendicular to the $[1\bar{1}2]$ direction. This surface morphology can be implemented by repeating the rectangular unit cell of Figure 2.1(c) in the miscut $[1\bar{1}2]$ direction. The in-

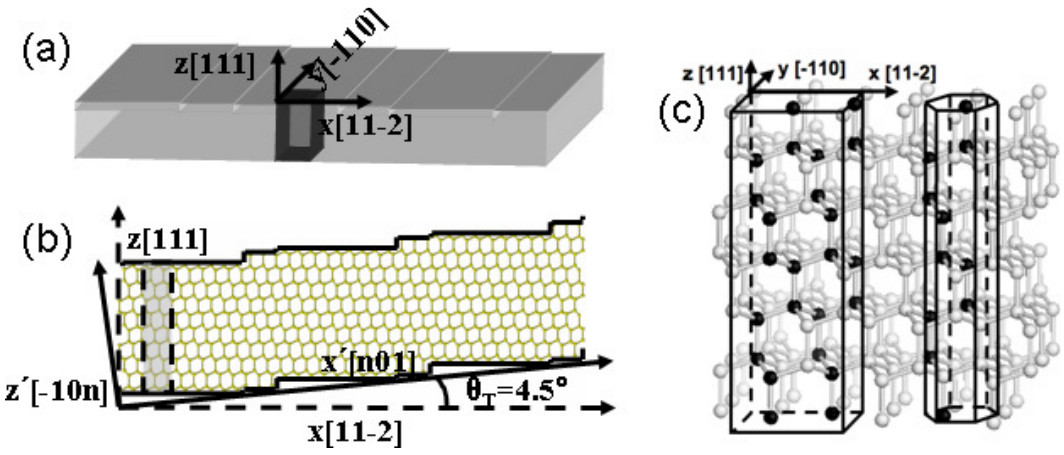

Figure 2.1. (a) Schematic of a miscut (111) quantum well. A rectangular unit cell is repeated in space to build the miscut quantum well. (b) Atomistic view of a unit cell of a 4.5° miscut (111) Si quantum well. The reduced symmetry along the direction perpendicular to steps results in a larger unit cell. The smallest repeated miscut quantum well unit cell along $[1\bar{1}2]$ direction has 6 steps. (c) Atomistic view of a flat (111) Si quantum well. The rectangular unit cell is larger than the hexagonal primitive unit cell.

plane unit cell dimensions are $a_x = \sqrt{3/2}a_{Si}$ and $a_y = \sqrt{1/2}a_{Si}$, where a_{Si} is Si lattice constant [70]. The advantage of using a rectangular unit cell is two-fold. First, the rectangular geometry simplifies the underlying mathematics and implementation of the periodic boundary conditions in the bandstructure calculation. Second, the surface miscut can be easily implemented by applying shifted boundary condition to the rectangular supercell.

The Brillouin Zone and 6 degenerate valleys in the bandstructure of the hexagonal primitive unit cell of a flat (111) Si quantum well (Figure 2.1(c)) are shown in Figure 2.2(a). The 2 bulk valleys along $[001]$ direction are projected along $[11\bar{2}]$ direction while the remaining 4 bulk valleys along $[100]$ and $[010]$ are symmetrically projected in each of the four quadrants [49]. The 2 valleys along $[11\bar{2}]$ are labeled as *A* and remaining 4 valleys are labeled as *B*. The rectangular unit cell is larger than the hexagonal primitive unit cell (Figure 2.1(c)). The Brillouin Zone of the rectangular unit cell of Figure 2.1(b) is given by $\{(k_x,k_y): -\sqrt{2/3}\,\pi/a_{Si} \le k_x \le \sqrt{2/3}\,\pi/a_{Si},\ -\sqrt{2}\,\pi/a_{Si} \le k_y \le \sqrt{2}\,\pi/a_{Si}\}$. It is smaller than the hexagonal Brillouin Zone in Figure 2.1(a) and 2 *A* type valleys are folded.

2.3. Miscut (111) surface

The miscut surface morphology can be conveniently implemented by extending the rectangular unit cell in one direction (Figure 2.1(b,c)). The unit cell shown in Figure 2.1(b) has a miscut angle of 4.5°. In the (*x,y,z*) co-ordinate system of Figure 2.1(a) the surface normal of this quantum well is along $[\bar{1}\,0n]$ direction. This direction is related to the miscut angle by $\theta_T = \tan^{-1}(a_z / na_x)$, where $a_x = \sqrt{3/2}a_{Si}$ and $a_z = \sqrt{3}a_{Si}$.

Typical miscuts range from 0.1-8°. Before going to the experimental 0.2° miscut [45] we illustrate the essential physics and reduce the computational burden significantly by studying the effect the miscut of 13°, which can be investigated in a system extending six unit cells along the miscut direction and has a smaller supercell compared to 0.2° miscut. These unit cells are schematically shown along with energy contours of their lowest conduction bands in Figure 2.2(c,d). Only the positive quadrant is shown. The flat

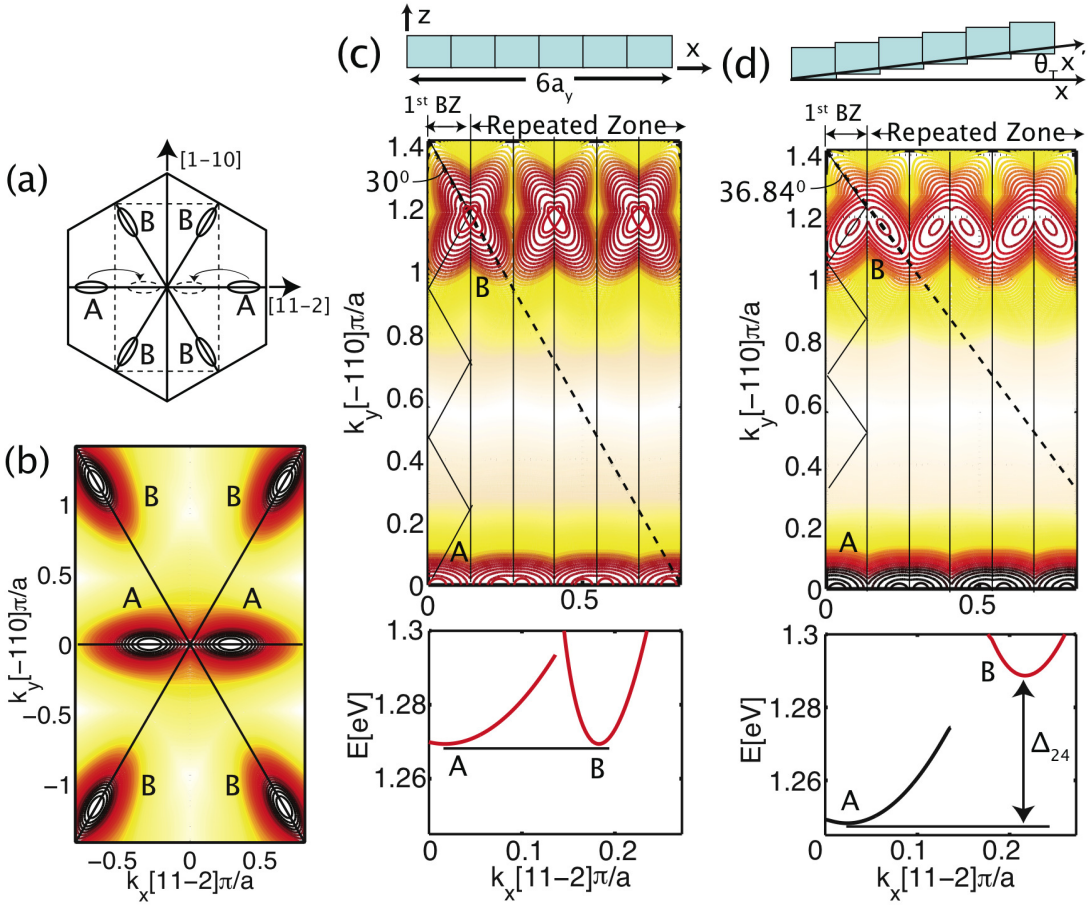

Figure 2.2. (a) Brillouin zone of the primitive unit cell of a flat (111) Si quantum well. 6 degenerate valleys are shown. (b) Brillouin zone and bandstructure of a (111) Si quantum well plotted using the rectangular unit cell of Figure 2.1(b). The bandstructure is folded as shown schematically in (a). (c) Bandstructure of a flat (111) Si quantum well in repeated zone scheme plotted using a supercell which contains 6 small cells in *x*-direction. The bandstructure of Figure 2.1(b) is folded in the 1st Brillouin zone. A type valleys along k_x-direction and B type valleys along the dotted line are degenerate. (d) Bandstructure of a 13° miscut (111) Si quantum well in repeated zone scheme. A-type valleys along k_x-direction are lower in energy than B-type valleys along the dotted line. In the flat quantum well the dotted line subtends an angle of 30° with the negative *y*-axis while this angle is 36.84° in a miscut quantum well.

quantum well supercell of Figure 2.2(c) is 6 times longer in the *x*-direction compared to the rectangular unit cell used to plot the bandstructure in Figure 2.2 (b) resulting in zone-folding of this supercell to 1/6th. In the first Brillouin zone *A*-type valleys lie along the k_x-direction and the *B* type valleys lie along the solid zig-zag line. In the repeated-zone scheme the bandstructures along the solid zig-zag line is the same as that along the dotted straight line. The bands corresponding to *A* and *B* type valleys clearly show that in a flat (111) Si quantum well both valleys are degenerate while in a miscut quantum well two *A*-type valleys have lower energy compared to four *B*-type valleys (Figure 2.3(d)). This degeneracy breaking is the effect of different confinement effective masses of *A*- and *B*-type valleys in a miscut quantum well.

2.4. Valley-projection model

Experimentally relevant quantum wells have typical miscuts ranging from 0.1-8°. As the miscut angle becomes smaller the size of the unit cell increases. For example, the unit cell of 23 nm thick 13° miscut quantum well contains 2,100 atoms while a 0.2° miscut quantum well requires 119,100 atoms. An effective mass theory can be used to determine the directions of valley minima thereby reducing the computational burden associated with searching the whole Brillouin zone. Here we outline the valley projection model [49] as applied to determine the directions of valley minima of miscut (111) Si quantum wells.

Consider the miscut quantum well unit cell of Figure 2.1(c), the rotation matrix *R* from bulk valley co-ordinate system to (*x´,y,z´*) co-ordinate system is given by

$$R = \frac{1}{\sqrt{n^2+1}} \begin{bmatrix} n & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & n \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -\sqrt{2/3} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \tag{2.1}$$

Wavefunction of a quantum well with the growth direction along $k_{z'}$ can be written as $\psi(x', y, z') = \xi(z')e^{ik_{x'}x'+ik_y y}$. In the effective mass formalism the sub-bands of this quantum well are given by [49]

$$E_i(k_{x'},k_y)=E_i^0+\frac{1}{2}\hbar^2\left[\left(w_{x'x'}-\frac{w_{x'z'}^{\ 2}}{w_{z'z'}}\right)k_{x'}^{\ 2}+2\left(w_{x'y}-\frac{w_{x'z'}w_{yz'}}{w_{z'z'}}\right)k_{x'}k_y+\left(w_{yy}-\frac{w_{yz'}^{\ 2}}{w_{z'z'}}\right)k_y^{\ 2}\right] \quad (2.2)$$

where, the reciprocal effective mass matrix $[W]$ is given by

$$[W]_{ij}=\sum_{\alpha}R_{i\alpha}R_{j\alpha}[M_0]^{-1}{}_{\alpha\alpha}\ ;\ \ i,j,\alpha\in\{x',y,z'\} \quad (2.3)$$

here $[M_0]$ is the effective mass matrix in the bulk valley co-ordinate system. The position of the sub-band minimum, E_i^0, is determined by the confinement effective mass $m_{z'}=1/w_{z'z'}$ and the confinement potential in the direction perpendicular to the quantum well surface. In this formalism two *A*-type valleys which lie along [100] direction are projected along $k_{x'}$ while the remaining four *B*-type valleys are projected along the directions which subtend angles ±φ with $\pm k_y$ axis. The angle φ can be determined by rotating the co-ordinate system ($k_{x'},k_y$) such that the cross term in equation (2.2) vanishes. This angle is given by $\varphi=\frac{1}{2}\tan^{-1}\left(\frac{c}{a-b}\right)$ where *a*, *b* and *c* denote coefficients of terms $k_{x'}^2$, k_y^2 and $k_{x'}k_y$ respectively. One of these four directions which lie in the positive quadrant is shown for flat and miscut quantum wells in Figure 2.2.

While in a flat (111) quantum well confinement effective masses, $m_{z'}$, are the same for *A*- and *B*-type valleys, a miscut alters these masses such that $m_A>m_B$ resulting in a broken degeneracy of lower 2 fold and raised 4 fold. Although the effective mass theory can explain the origin of $\Delta_{2\text{-}4}$ splitting more sophisticated methods such as tight-binding are needed to accurately model the effect of mono-atomic surface steps on the electronic structure. The $\Delta_{2\text{-}4}$ splitting increases with the miscut angle due to increasing difference between confinement effective masses of *A*- and *B*-type valleys. Both effective mass and tight-binding models show this trend, the effective mass model, however, gives smaller splitting compared to the tight-binding model. As shown in Figure 2.3 the step morphology of the quantum well surface modulates the wavefunction which in turn influence energy levels to give rise to $\Delta_{2\text{-}4}$ splitting.

2.5. Comparison with experiments

A 0.2° miscut resembles closely to that of the experiment [45]. The unit cell of a quantum well of this miscut is L_x=264.07 nm and L_y=0.38 nm long in x- and y-directions, respectively. A constant z-directed electric field of 10 MV/m which corresponds to the electron density of 6.5×10^{11} cm^{-2} is assumed [45]. To avoid any truncation effects of electronic domain on eigenvalues a quantum well thickness of L_z=23 nm is simulated. This is the smallest unit cell which can be repeated in xy-plane to generate a 23 nm thick 0.2° miscut quantum well. This unit cell contains around 0.1 million atoms which makes it computationally expensive to search the whole 2 dimensional Brillouin zone for valley minima. To reduce the computational burden, the valley-projection method described above is used to identify the directions of valley minima. The valley minima of A-type valleys occur along the k_x-direction while valley minima of B-type valleys occur at an angle $\varphi=\pm30.13°$ to k_y axis. The confinement effective masses, m_z, for A- and B-type valleys are $0.2608m_0$ and $0.2593m_0$ (where m_0 is the mass of the free electron) respectively. The $\Delta_{2\text{-}4}$ valley splitting in a 0.2° miscut quantum well calculated using an effective mass and the $sp^3d^5s^*$ tight-binding models are 1.25 (108 μeV) and 3.98 K (343 μeV) respectively. An analytical formula for energy levels in triangular potential wells is used to estimate the splitting in the effective mass model [49]. The splitting reported from the temperature dependence of the longitudinal

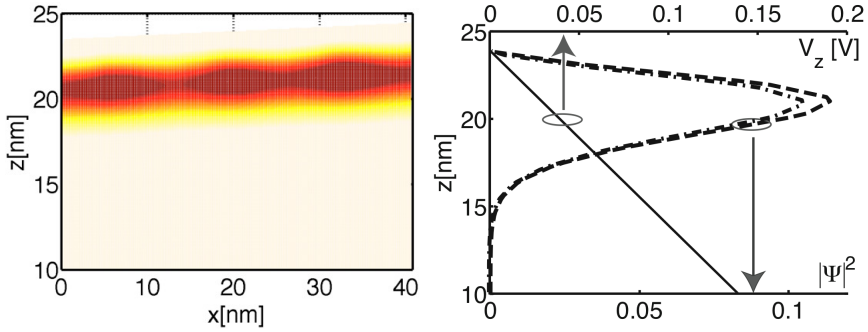

Figure 2.3. (a) A typical wavefunction of a 4.5° miscut (111) Si quantum well of Figure 2.1(b) at the valley minimum. Thickness of the quantum well is 23 nm, however, only the near surface portion having appreciable wavefunction magnitude is shown. (b) Electrostatic potential and wavefunction cut along two dotted lines in Fig (a). Confining potential due to constant electric field (10 MV/m) pulls the wavefunction to the surface. Surface steps modulate the wavefunction amplitude and modify energy levels which give rise to the $\Delta_{2\text{-}4}$ splitting.

resistance [45] is 7K (604 μeV) which shows that the tight-binding calculation matches closer to experiments compared to the effective mass calculation.

In conclusion, the miscut morphology of the (111) Si surface is shown to be the origin of breaking of 6 fold valley degeneracy into lower 2 and raised 4 fold valley degeneracies. Atomistic basis representation such as tight-binding is needed to capture the effect of wavefunction modulation at mono-atomic steps on the electronic structure. Compared to effective mass the tight-binding calculations are found to match closer to experimentally measured $\Delta_{2\text{-}4}$ splitting. Additional surface phenomena could be responsible for enhanced $\Delta_{2\text{-}4}$ splitting reported in experiments [45].

2.6. Conclusion

The miscut morphology of the (111) Si surface is shown to be the origin of breaking of 6 fold valley degeneracy into lower 2 and raised 4 fold valley degeneracies. Atomistic basis representation such as tight-binding is needed to capture the effect of wavefunction modulation at mono-atomic steps on the electronic structure. Compared to effective mass the tight-binding calculations are found to match closer to experimentally measured $\Delta_{2\text{-}4}$ splitting. Additional surface phenomena could be responsible for enhanced $\Delta_{2\text{-}4}$ splitting reported in experiments [45].

3. VALLEY SPLITTING IN NANOSTRUCTURES GROWN ON MISCUT SUBSTRATES

3.1. Introduction

Valley splitting in (001) Si quantum wells is being actively researched and recently several research groups have proposed quantum computing architectures based on (001) Si quantum wells embedded in SiGe buffer layers [24, 53-55, 60, 61, 71, 72]. Si is desirable for quantum computing due to its long spin-decoherence times, scaling potential and integrability within the present microelectronic infrastructure [56, 57]. Bulk Silicon has 6 degenerate X-valleys. In strained Si 6-fold valley-degeneracy of Si is broken into lower 2-fold and raised 4-fold valley-degeneracies. The presence of 2-fold valley-degeneracy is a potential source of decoherence which leads to leakage of quantum information outside qubit Hilbert space [57]. Therefore, it is of great interest to study the lifting of remaining 2-fold valley degeneracy in strained Si due to sharp confinement potentials in recently proposed [55] SiGe/Si/SiGe quantum well heterostructures based quantum computing architectures.

Miscut surfaces as opposed to flat surfaces are often used to ensure uniform growth of Si/SiGe heterostructures. The most commonly used miscut substrate has its surface oriented at 2^{o} to the (001) crystallographic plane. Existing theories of valley splitting, which do not include miscut surface geometries, predict valley splitting in SiGe/Si/SiGe quantum wells higher by at least an order of magnitude than experimental measurements [60]. This over-prediction has been associated with the experimental condition that the quantum wells are grown on a 2^{o} tilted (001) substrate. The tight-binding treatments in small atomistic representations support this claim [61]. Effective mass treatments [71, 72] can explain the behavior of valley splitting in miscut quantum wells, however they rely on perturbation theories, ad-hoc structure dependent fitting parameters, and they cannot

address the atomistic alloy disorder at the interface. This chapter develops qualitative and quantitative understanding of valley splitting in miscut quantum wells with no disorders such as surface roughness or alloy randomness. The valley splitting theory is extended to include disorders and magnetic fields and the results are compared with the experimental data in chapter 4.

The valley splitting in flat (001) Si quantum wells is reviewed in section 3.2. Section 3.3 describes the bandstructure of the miscut quantum well from effective mass and tight-binding point of view. Sections 3.4 and 3.5 develop the understanding of valley splitting in miscut nanowires (which resemble quantum wells in the presence of electrostatic confinements) from bandstructures of miscut quantum wells. Wavefunction parities and analytical calculations of valley splitting are discussed. Finally the results are summarized in section 3.6.

3.2. Valley splitting in a flat (001) quantum well

In the effective mass picture, electrostatic confinement due to finite thickness of the (001) Si quantum well splits 6-fold valley degeneracy of Si into lower 2 and higher 4-fold degeneracy. The lower 2-fold degeneracy consists of the 2 degenerate *z*-directed valleys while higher 4-fold degeneracy consists of 2 *x*-directed and 2 *y*-directed valleys [51]. Here, *x*, *y*, and *z* directions respectively correspond to [100], [010], and [001] crystallographic axes. This splitting is a result of different effective mass of valleys in the confinement (growth) direction of the quantum well. Similar lifting of 6-fold valley degeneracy of bulk Si into lower 2-fold and higher 4-fold degeneracies occurs in the presence of biaxial strain applied in *xy*-plane [51]. Lower 2-fold degenerate valleys further split in the presence of the sharp confinement potential at the edges of the quantum well. Because the valley minima lie completely within the Brillouin zone, there are four propagating states (with *z* wave vectors $\pm k_z^1$, $\pm k_z^2$ near the minima at $\pm k_z^0$) at each energy within the valleys (Figure 3.1 (a)). In the simplest approximation, the quantum well bound states are linear combinations of all four propagating states, and as a result, the bound states occur in doublets. The splitting between the states comprising the

lowest doublet is referred to as the valley splitting, which is plotted as a function of the quantum well thickness in Figure 3.1 (b). In a flat [001] quantum well the valley minima of split valleys occur at $(k_x, k_y) = (0,0)$. The valley splitting can be included in effective mass theory by using perturbation theories and ad-hoc structure dependent parameters. However, tight binding naturally gives such splitting because the finite size effects are automatically included due to atomistic nature of the Hamiltonian and the finite size of the semiconductor crystal. The physics of valley splitting in a flat [001] quantum well in tight-binding formalism is discussed in detail in [58, 59]. The valley splitting oscillates as a function of quantum well thickness, with a period that is determined by the location of the valley minimum (k_z^0) in the bulk Brillouin zone. The envelope of the splitting decays as S^{-3}, where S, the number of atomic layers in the quantum well. The analytical expression describing valley splitting in (001) Si quantum wells is derived in [58, 59] and it is given by

$$\Delta_v \approx \frac{16\pi^2 u}{(S+2)^3}\left|\sin(S+2)k_z^0 a\right|\sin\left(k_z^0\right) \tag{3.1}$$

where, k_z^0 is the location of z-valley minima, a is the lattice constant of Si. The fitting parameter, u depends on the valley effective mass and the location of valley minima. As

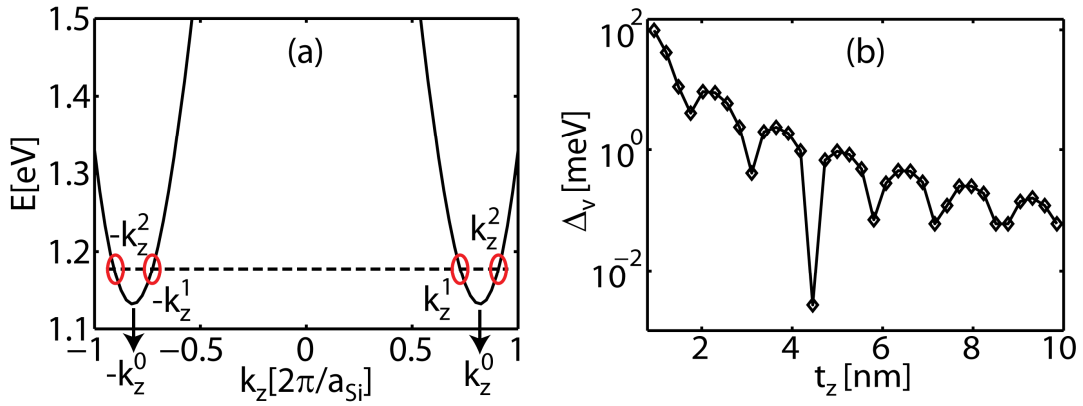

Figure 3.1. (a) z-directed valleys of bulk silicon. Quartet of states near the valley-minima $\pm k_0^z$ interact and split in the presence of [001] directed confinement present in (001) quantum well. (b) Valley splitting in a (001) Si quantum well oscillates with the quantum well thickness and decays with the envelope of S^{-3}, where S, the number of atomic layers in the quantum well.

discussed in the following sections, the theory of valley splitting in miscut quantum wells is closely related to the flat quantum well and it can be estimated by the similar analytical formula.

3.3. Bandstructure of a miscut quantum well

The valley splitting physics of a miscut quantum well is slightly different than a flat quantum well. The first step in understanding the valley splitting in a miscut quantum well is to start with the bandstructure. Following sub-sections describe the bandstructure in effective mass and tight-binding formalisms.

3.3.1. Effective mass

The lowest conduction band of a miscut quantum well in the effective mass approximation can be obtained by projecting the *z*-valleys onto the quantum well growth surface ($\bar{1}0n$) as shown in Figure 3.2. n=28 for a 2° miscut quantum well. Projected *z*-valleys are shifted away from the origin along the tilt-direction [$n01$]. The states in these

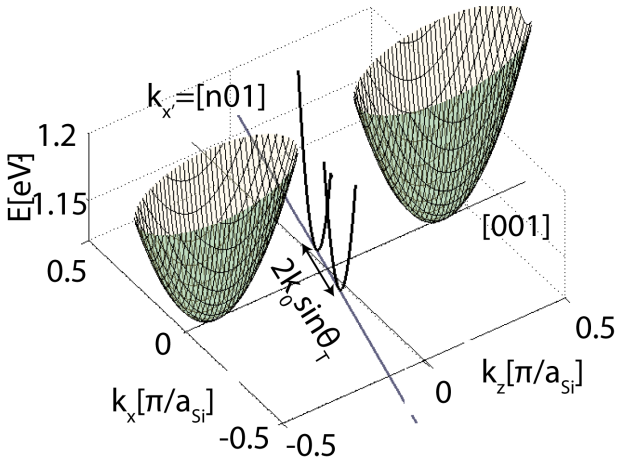

Figure 3.2. Bandstructure of a miscut quantum well in effective mass approximation: (a) Bandstructure of a quantum well grown on a miscut surface can be obtained by projecting bulk conduction band valleys at (0, 0, ±k_0) onto the quantum well growth surface [73]. Two degenerate valleys are located at $k_{x'}^0 = \pm k_0 \sin\theta_T$ along [$n01$] direction. The quantum well growth direction is [$\bar{1}0n$]. n=28 for 2° miscut.

valleys only weakly interact in the presence of confinement. Thus a miscut quantum well shows the presence of two degenerate valleys located at $k_{x'}^0 = \pm k_0 \sin\theta_T$. Valley-valley interaction, however, is stronger at $k_{x'} = 0$ which causes the splitting called the minigap (Δ_m) [74, 75].

In this analytical view where m_t and m_l are the bulk Si transverse and longitudinal effective masses, the effective mass of the projected z-valleys along [n01] direction is $1/\mathrm{m}_{\mathrm{eff}} = \cos^2(\theta_T)/m_t + \sin^2(\theta_T)/m_l$. Here, the tilt angle $\theta_T = \tan^{-1}(1/n)$ is the angle between [100] and [n01] crystal directions. At $k_{x'}^0 = 0$ a small minigap forms due to stronger valley-valley interaction [73]. This gap is, however, significantly higher in energy than the two degenerate valleys at $\pm k_{x'}^0$. Thus atomic scale modulation of surface topology leads to very different electronic structures in flat and miscut quantum wells. In a flat (001) quantum well, valleys are split while in a miscut quantum well they are degenerate. Additional lateral confinement is needed in miscut quantum wells to split the degeneracy of these lowest valleys which can be provided by lateral electrostatic gates or a magnetic field [60].

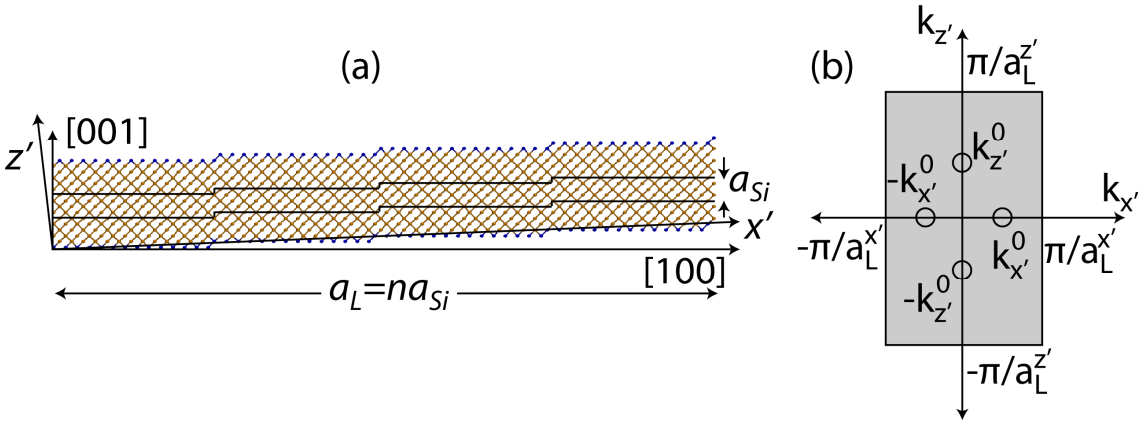

Figure 3.3. (a) A minimal unit cell of a 2° miscut quantum well consists of 4 steps. x' and z' directions are along [n01] and [$\bar{1}0n$] crystallographic axes respectively. n=28 for 2° tilt. (b) The $x'z'$-cut through the Brillouin zone. $a_L^{x'} = n a_{Si}$, $a_L^{z'} = \frac{n}{\sqrt{n^2+1}} a_{Si}$. $\pm k_{x'}^0$ and $\pm k_{z'}^0$ are the positions of the conduction-band valley-mimima.

3.3.2. Tight-binding

A quantum well grown on top of a miscut substrates has steps resulting from the misorientation between crystallographic direction [001] and the quantum well growth direction $[\bar{1}0n]$. Typically miscut produces steps of mono-atomic height $a_{Si}/4$ [48], where a_{Si} is the lattice constant of Silicon. The detailed atomistic view of the unit cell of a miscut quantum well grown on a $(\bar{1}0n)$ surface is shown in the xz-plane in Figure 3.3(a). The structure shown in the figure can be repeated in x' and y-directions to form the whole quantum well.

The length of the steps (L_s) is related to the miscut angle of the substrate, by $L_s = h/\tan(\theta_T)$. The periodicity is along x' and y-directions while z' direction is confined. In NEMO-3D, periodic boundary conditions are generally applied along x, y and/or z directions i.e. along the crystallographic axes [100], [010], and [001]. However the unit cell of Figure 3.3(a) is periodic along x' and z' directions i.e. $[n01]$ and $[\bar{1}0n]$ directions

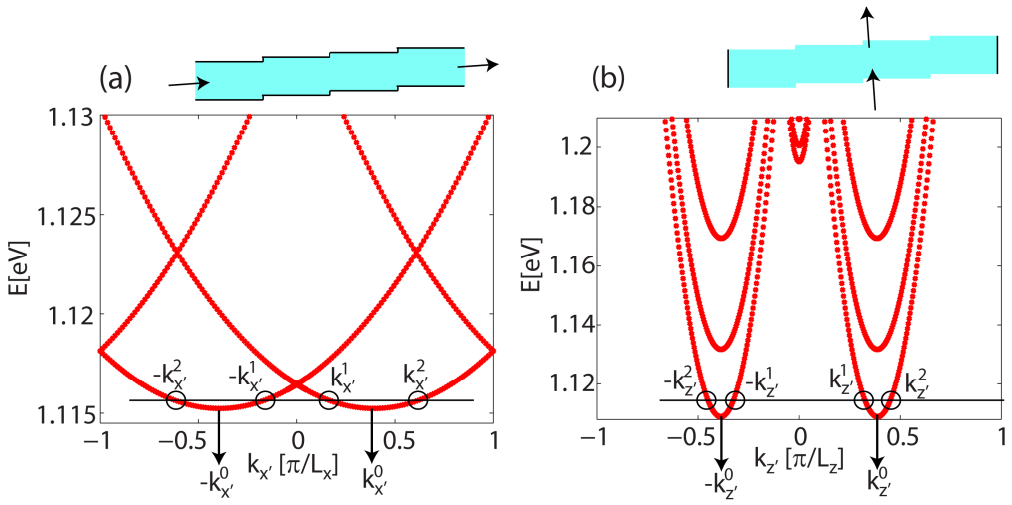

Figure 3.4. Tight-binding bandstructures of a 2° miscut quantum well. (a) Bandstructure of a 2° miscut unit cell confined in z′ direction shows two degenerate valleys centered at $k^0_{x'} = \pm k_0 \sin\theta_T$. Lateral extensions of the unit cell are L_x=15.33nm, L_y=0.5432nm, and thickness is t_z=5.26nm. Confinement in x'-direction causes quadrate of states in these two degenerate valleys to interact and split. (b) Bandstructure of a 2° miscut unit cell confined in x'direction shows two degenerate valleys centered at $k^0_{z'} = \pm k_0 \cos\theta_T$. Confinement in z'-direction causes quadrate of states in these two degenerate valleys to interact and split. All bandstructures are calculated in $sp^3d^5s^*$ tight-binding model with spin-orbit coupling. Lateral extensions of the unit cell are L_x=15.33nm, L_y=0.5432nm, and thickness is t_z=5.26 nm.

respectively. In the NEMO-3D framework the periodic boundary conditions are applied after integer number of unit cells. Since the step height is equal to the quarter of unit cell dimension this unit cell can be made periodic after four steps. The Brillouin zone of this unit cell is

$$\left\{\left(k_{x'},k_y,k_{z'}\right):-\frac{\pi}{na}\le k_{x'}\le\frac{\pi}{na},-\frac{\pi}{a}\le k_y\le\frac{\pi}{a},-\frac{\pi}{na}\sqrt{n^2+1}\le k_{z'}\le\frac{\pi}{na}\sqrt{n^2+1}\right\} \tag{3.2}$$

where a is the lattice constant of Silicon. The $x'z'$-cut through the Brillouin zone is shown in Figure 3.3(b). The positions of valley-minima are marked by $\pm k_{x'}^0$ and $\pm k_{z'}^0$.

Figure 3.4(a) shows the bandstructure of the unit cell in Figure 3.3(a) plotted by applying periodic boundary conditions in x' and y-directions and closed boundary conditions in z'-direction. The lowest conduction band shows the two degenerate valleys similar to those observed in effective mass picture of Figure 3.2. Additional confinement in x' direction will cause the quartet of states shown in Figure 3.4(a) to interact and form two valley-split states. However, with x'-confinement in addition to z-confinement a miscut quantum well becomes a miscut nanowire. Thus miscut quantum well has two degenerate valleys while in a miscut nanowire this valley-degeneracy is broken to form valley-split states.

Figure 3.4(b) shows the bandstructure of the unit cell for Figure 3.3(a) plotted by applying periodic boundary conditions in y and z'-directions and closed boundary conditions in x' direction. Similar to Figure 3.4(a) bandstructure under these boundary conditions shows two degenerate valleys along z′-direction. These valleys interact and split in the presence of additional z′ confinement to give rise to valley splitting in a miscut nanowire.

3.4. Wavefunction parities in a miscut nanowire

Each of the two valley-split wavefunctions of a miscut nanowire can be roughly expressed as the product of x' and z' components as

$$\left|\Psi\right\rangle=\left|\psi_{x'}\right\rangle\left|\psi_{z'}\right\rangle \tag{3.3}$$

As shown in Figure 3.5 and Figure 3.6, $|\psi_{z'}\rangle$ and $|\psi_{x'}\rangle$ coefficients of the ground state on anions/cations have odd/even parities. While $|\psi_{z'}\rangle$ and $|\psi_{x'}\rangle$ coefficients of the valley-split state on anions/cations have even/odd parities. The wavefunction components exhibit such definite parity relations in x' and z' directions because they are the result of

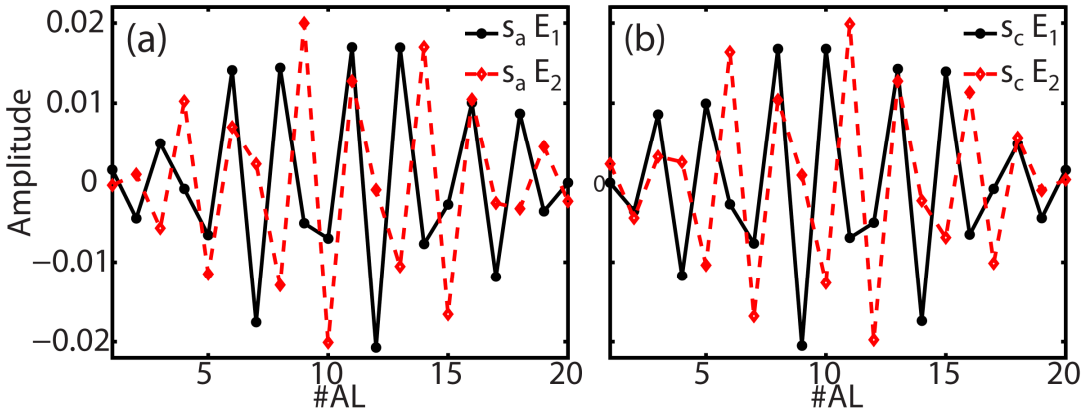

Figure 3.5. (a) Anion s-orbital contributions of the two valley-split states along thickness of a 2° miscut nanowire. Low energy state has even parity while the valley-split state has odd parity. (b) Cation s-orbital contributions have parities opposite to the anion s-orbital contributions of Fig. (a). Dimensions of the simulated nanowire unit cell are L_x=15.33nm, t_z=5.26nm in the confinement directions and L_y=0.55nm in the periodic direction.

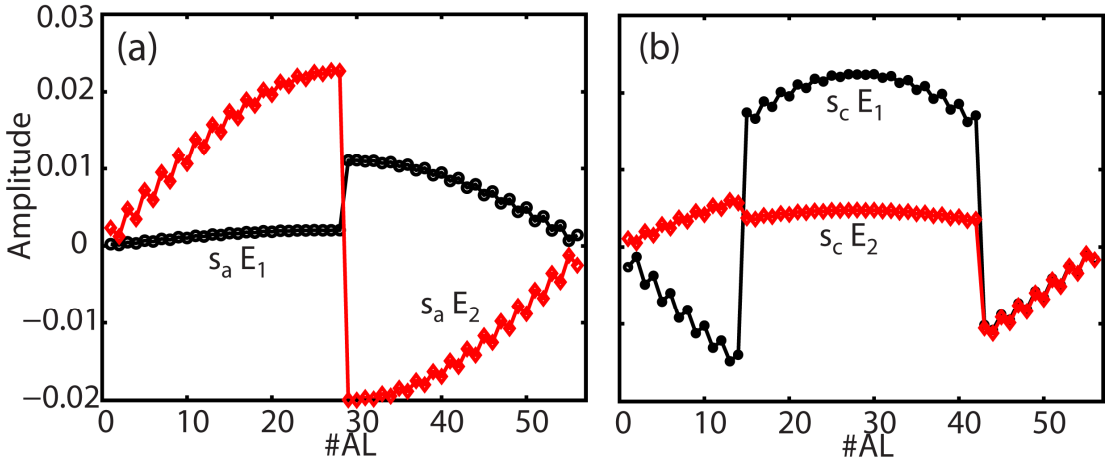

Figure 3.6. s-orbital contributions of the valley-split wavefunctions on anion (a) and cation (b) sites along lateral (x') direction. Anion s-orbital contributions of the ground state have even parity while those of the valley-split state have odd parity. Parities are reversed for cation s-orbital contributions. Dimensions of the simulated structure are the same as in Figure 3.3.

interaction of a quartet of states shown in Figure 3.4. Similar wavefuntion parities are observed in a flat [001] quantum well [76]. In [76] an approximate expression for valley splitting is derived using even/odd parities of wavefunction coefficients. Valley splitting in a miscut quantum well essentially follows similar physics as a flat quantum well with additional complication that the valley splitting in a miscut quantum well (nanowire) is a result of two confinements. Therefore similar expressions as those derived in [76] can be used to estimate the valley splitting in a miscut nanowire.

3.5. Dependence of valley splitting on confinement dimensions

The valley splitting as a function of the thickness (z' confinement) of a miscut nanowire can be approximated as

$$\Delta_v^{t_z} \approx \frac{16\pi^2 u_z^{L_x}}{(S_z+2)^3}\left|\sin(S_z+2)k_{z'}^0 a\right|\sin\left(k_{z'}^0\right) \tag{3.4}$$

where, $k_{z'}^0 = k_0\cos\theta_T$ is position of the valley-minimum as shown in Figure 3.4(b). S_z is the number of atomic layers in the thickness direction of the miscut nanowire. The fitting parameter $u_z^{L_x}$ depends on the lateral electrostatic confinement L_x of the quantum well and the position of valley-minimum ($k_{x'}^0 = \pm k_0\sin\theta_T$) in x'-direction. The thickness

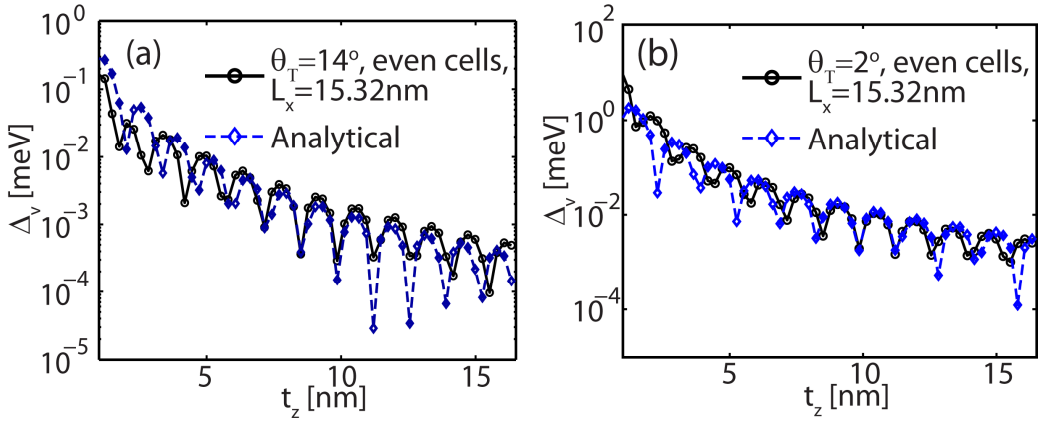

Figure 3.7. Valley splitting as a function of miscut nanowire thickness: (a) θ_T =14°. The solid black line shows the valley splitting calculated in $sp^3d^5s^*$ tight-binding model and dotted blue line shows valley splitting calculated using equation (3.4). (b) valley splitting in a 2° miscut nanowire. Confinement in x′-direction is 15.32 nm.

dimension (t_z) of a miscut nanowire is related to S_z by $t_z = S_z a/2$.

Figure 3.7(a) shows the valley splitting versus thickness of the 14^o ($=\theta_T$) miscut nanowire calculated with full nearest-neighbor $sp^3d^5s^*$ tight-binding model and using equation (3.4). The fitting parameter $u_z^{L_x} = 3\times10^{-3} eV$ is used. Similar plot of the dependence of the valley splitting on the thickness of a 2^o ($=\theta_T$) miscut quantum well is shown in Figure 3.7(b). The fitting parameter used here is $u_z^{L_x} = 1\times10^{-3} eV$. Valley splitting shows an oscillatory behavior with the decaying envelop of S_z^{-3}.

The valley splitting as a function of a lateral dimension L_x of a miscut nanowire can be approximated as

$$\Delta_v^{L_x} \approx \frac{16\pi^2 u_x^{t_z}}{(S_x+2)^3}\left|\sin(S_x+2)k_x^0 a\right|\sin\left(k_x^0 a\right) \tag{3.5}$$

where, $k_{x'}^0 = k_0 \sin\theta_T$ is position of the valley-minimum as shown in Figure 3.4(a). The smallest repeating unit in x' direction is comprised of 2 steps. Hence, $S_x = N_x/2$, where N_x is the number of steps in the lateral (x') direction of a miscut nanowire. In a miscut nanowire grown on ($\bar{1}0n$) surface, the lateral confinement length (L_x) is related to S_x by $L_x = S_x na/4$. The fitting parameter $u_x^{t_z}$ depends on the electrostatic confinement length in thickness direction (t_z) and the position of valley-minimum ($k_{z'}^0 = \pm k_0 \cos\theta_T$) in z′-direction.

The valley splitting in a 14^o miscut nanowire as a function of the lateral confinement is shown in Figure 3.8(a). 14^o tilt with respect to the [100] crystallographic axis corresponds to the crystal direction [4 0 1]. The step length of this quantum well is equal to the lattice constant in x-direction (a_x) which corresponds to 4 atomic layers. The valley splitting calculated using full $sp^3d^5s^*$ tight-binding model and using equation (3.5) show good agreement. The fitting parameter $u_x^{t_z} = 3.22\times10^{-4}$ is used. The valley splitting shows decaying oscillatory behavior when lateral confinement length is made commensurate with even number of steps.

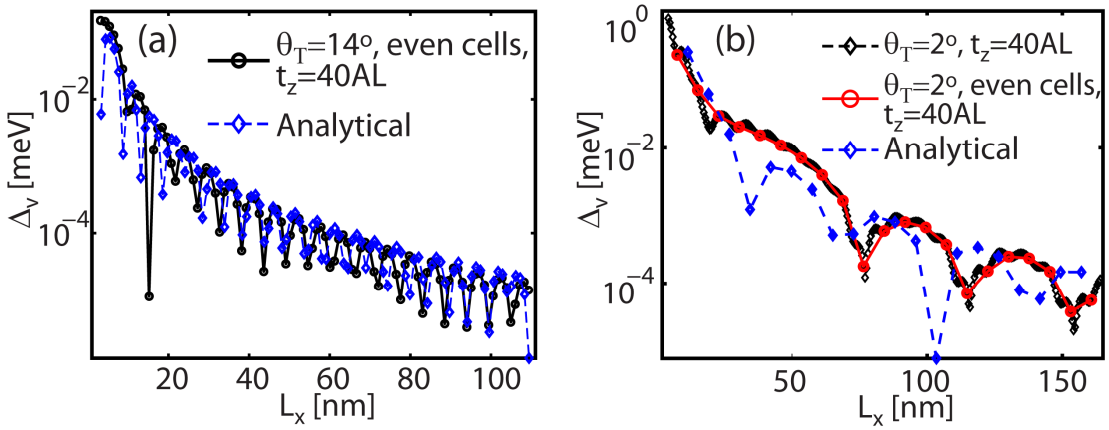

Figure 3.8. Valley splitting in a miscut nanowire as a function of the lateral confinement: (a) θ_T =14°. Valley splitting calculated in $sp^3d^5s^*$ tight-binding model (solid line) and from equation (3.5) (dotted line). (b) Valley splitting in a 2° miscut nanowire. To compare with the valley splitting calculated from equation (3.5) (blue line), only the points which are even multiples of step length are chosen (red line). Thickness of the miscut nanowire is t_z=5.26nm.

Figure 3.8(b) graphs the valley splitting in a 2° miscut nanowire as a function of the lateral confinement. The fitting parameter, $u_x^{t_z} = 2.12\times10^{-2} eV$ is used in equation (3.5). 2° tilt with respect to the [100] crystallographic axis corresponds to the crystal direction [28 0 1]. The step length of this quantum well is equal to $7a_x$ where a_x is the lattice constant in x-direction. Oscillations in 14° miscut quantum wells are more rapid than 2° miscut quantum wells because length of steps in 14° miscut quantum wells is smaller than 2° miscut quantum wells. The valley splitting calculated using full $sp^3d^5s^*$ tight-binding model and using equation (3.5) show same qualitative behavior and reasonable quantitative agreement.

3.6. Conclusion

A simple understanding of the origin of valley splitting in nanostructures grown on miscut substrates is developed. It is found that the lowest energy valleys in a miscut quantum well are degenerate. Additional lateral confinement in required to break this degeneracy. Wavefunctions of valley split states in a miscut nanowire can be expressed

as a product of x and z components. Based on this observation analytical formulae are proposed for calculating valley splitting in a miscut nanowire. Valley splitting calculated using full $sp^3d^5s^*$ tight-binding model are found to be in agreement with proposed formulae.

4. MAGNETIC FIELD DEPENDENCE OF VALLEY SPLITTING IN SI/SIGE HETEROSTRUCTURES

4.1. Introduction

The quantum well energy spectrum is often measured using traditional magnetic probe techniques such as Shubnikov de Haas oscillations. Recent experiments employ quantum point contact [60] techniques for higher resolution and to probe individual nanostructures. Valley and spin-splittings are determined by analyzing the energy spectrum data. In these measurements magnetic field is applied perpendicular to the plane of a quantum well. In the absence of lateral electrostatic gates, in plane (lateral) confinement is provided by the magnetic field, which results in the formation of Landau-levels. This chapter focuses on the effect of magnetic field confinement and substrate orientation on the valley splitting in the miscut (001) SiGe/Si/SiGe quantum wells, which is the fabrication platform for the spin based quantum computing architecture proposed in [54, 55].

Schematics of quantum wells grown on flat (001) and miscut (001) substrates are shown in Figure 4.1(a,b). Miscut modifies the energy spectrum of a quantum well. In a flat quantum well, valley splitting is higher that spin splitting while in a miscut quantum well valley splitting is suppressed and becomes smaller that spin splitting (Figure 4.1). As a result of this, first two conduction band states in a flat quantum well have same valley-index and opposite spins. However, first two conduction band states in a miscut quantum well have identical spins but different valley-index. Valley splitting and spin splitting are important design parameters for spin-based quantum computing and spintronic devices and thus need precise modeling techniques that include miscut orientations and disorders in buffer layers. In this chapter, a supercell tight-binding technique is employed to model valley splitting in the miscut (001) SiGe/Si/SiGe quantum wells. This chapter expands on the concepts of previous chapter, represents strain and interfaces atomistically, includes

disorder at the steps [48] and disorder in the confining alloy, without any ad-hoc fitting parameters. The general purpose code NEMO-3D [39, 40, 77] is utilized to perform multi-million atom simulations in domains that are large enough to be representative of realistic device geometries. Outline of this chapter is as follows. Section 4.2 describes models used to include electric and magnetic fields in the tight-binding Hamiltonian. Section 4.3 presents comparison between valley splitting in ideal flat and miscut (001) Si quantum wells. Effect of barrier height at the Si/SiGe interface is discussed in section 4.4. In section 4.5 valley splitting calculations on realistic SiGe/Si/SiGe quantum wells are presented. These calculations include inhomogeneous strain and step roughness disorder on the atomistic scale and match well with the experimental measurements. Finally, conclusions are presented in section 4.6.

4.2. Model

The general purpose electronic structure calculation code NEMO-3D is used to model strain distribution and electronic structure in the SiGe/Si/SiGe quantum wells. The valence force field (VFF) model with the Keating potential [78] is used to move the atoms to positions which minimize the total strain energy. A subsequent electronic structure calculation is based on the 20 band $sp^3d^5s^*$ tight-binding model [76, 79]. The strain is included in the tight-binding Hamiltonian following the prescription given in

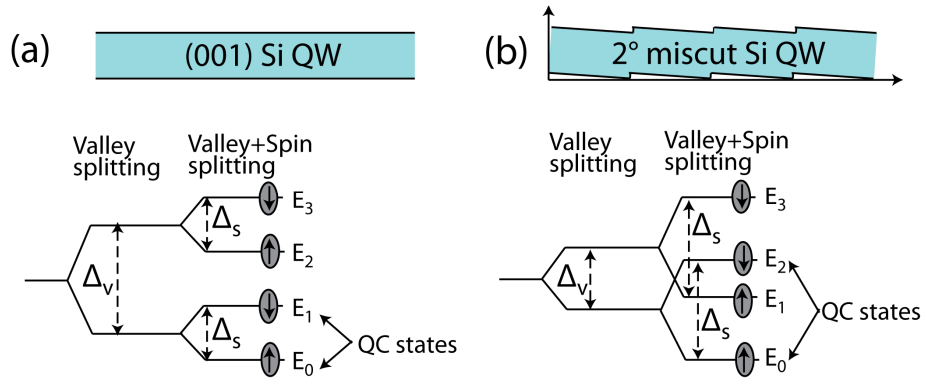

Figure 4.1. Fan diagrams of valley and spin-split states in flat (a) and miscut (b) quantum wells in the presence of the magnetic field. The two lowest energy states in a flat quantum well are spin-split states while in a miscut quantum well the states are valley-split and have identical spins.

[80].

4.2.1. Electric field

Modulation doping used to generate two dimensional electron gas (2DEG) induces a built-in electric field in SiGe/Si/SiGe heterostructures. Here, we do not solve the electrostatic potential self-consistently but focus on the essential physics and assume a constant electric field of E_z=9 MV/m along the quantum well growth direction which resembles the experiments in [60]. The electric field is incorporated in the Hamiltonian by adding the scalar potential term $\Phi(R) = -zE_z$ to the diagonal elements of the Hamiltonian.

4.2.2. Magnetic field

The external magnetic field used to probe energy levels in the quantum well is included by using gauge-invariant Peierl's substitution [81, 82]. Consider a tight-binding Hamiltonian that is characterized by the on-site and coupling Hamiltonian matrix elements

$$\begin{aligned} \varepsilon^0_{\alpha,R} &= \langle \alpha, R | \hat{H}^0 | \alpha, R \rangle, \\ t^0_{\alpha',R',\alpha,R} &= \langle \alpha', R' | \hat{H}^0 | \alpha, R \rangle \end{aligned} \tag{4.1}$$

Here, the orthonormal basis states $|\alpha, R\rangle$ are labeled by a site index R and an orbital index α that labels an orbital at a given site. Under Peierl's substitution off-diagonal (coupling) matrix elements of the tight-binding Hamiltonian are scaled as

$$t_{\alpha',R',\alpha,R} = t^0_{\alpha',R',\alpha,R} \exp\left\{ -\frac{ie}{\hbar} \int_R^{R'} A(r,t).dl \right\} \tag{4.2}$$

The spin contribution to the magnetic moment is included by adding $\mu_B \boldsymbol{\sigma} \cdot \boldsymbol{B}$ to the diagonal matrix elements, where μ_B is Bohr magneton and $\boldsymbol{\sigma}$ are Pauli spin matrices. The modification to the diagonal matrix element due to electric field and spin contribution of the magnetic field is given by

$$\varepsilon_{\alpha,R} = \varepsilon^{0}_{\alpha,R} - \Phi(R) + \mu_B \sigma.B \tag{4.3}$$

4.2.3. Magnetic gauge and lateral device geometry

As explained in chapter 3, confinement along the step direction (*y*-direction) does not modify valley splitting in flat or miscut quantum wells. Therefore, asymmetric (Landau) gauge in which the vector potential is given by $A = Bx\hat{y}$ is chosen. Magnetic field is applied in the direction perpendicular to the quantum well surface. In a low magnetic field regime of interest the vector potential is slowly varying on the scale of nearest neighbor distance. Under this approximation the line integral in equation (4.2) can be approximated as

$$\int_{R}^{R'} A(r,t).dl = \frac{1}{2}\left(A(R) + A(R')\right).(R' - R) \tag{4.4}$$

Here time independent version of Peierl's substitution is used. Hence the time dependence in equation (4.2) is not included. In asymmetric gauge, equation (4.4) simplifies to

$$t_{\alpha',R',\alpha,R} = t^{0}_{\alpha',R',\alpha,R} \exp\left\{\frac{1}{2} B(x + x').(y' - y)\right\} \tag{4.5}$$

Due to the x-coordinate dependence of asymmetric gauge the device geometry cannot be taken as periodic in *x*-direction. The system must therefore be closed in the *x*-direction, which will indeed introduce confinement effects. The artificial confinement induced by these spatial boundary conditions compete with the realistic magnetic confinement effects. The lateral x-dimension is set to 150nm, which is about 7 times larger than the maximum magnetic confinement length in a 2DEG at B=1.5T ($\simeq$21nm). For the magnetic field ranges of 1-5T no lateral *x*-confinement effects due to the closed boundary conditions are visible in our simulations. Confinement is dominated by the magnetic field.

4.3. Valley splitting in infinite barrier quantum wells

Evolution of first four conduction band states of a 10nm thick strained flat (001) Si quantum well in magnetic field is shown in Figure 4.2. Magnetic field is applied in the direction perpendicular to the quantum well surface. Same scheme as in Figure 4.1 is used to label the quantum well states. First two states E_0 and E_1 are spin split while E_0 and E_2 are valley-split. Valley split states are distinguished from the spin split states by observing the wavefunctions. Spin split states have the same special distribution. Valley-split states have the same wavefunction envelop but their rapidly varying components are out of phase [58]. Valley splitting ($\Delta_v=E_2-E_0$) and spin splitting ($\Delta_s=E_1-E_0$) in a flat quantum well are shown in Figure 4.3(a). Valley splitting is higher than spin splitting in the low magnetic field regime.

Valley splitting in 2^o miscut quantum wells is suppressed by at least 2 orders of magnitude as compared to flat quantum wells (Figure 4.3(a,b)). Spin splitting, however, is not affected by the change in substrate orientation since it is a result of the geometry independent spin-correction $\mu_B\boldsymbol{\sigma}\cdot\boldsymbol{B}$. The Lande g-factor extracted from the spin splitting plots in Figure 4.3(a) is 2.0028, which is very close to the ideal free-electron value as reported in [60]. The computed valley splitting in 2^o miscut quantum wells (Figure 4.3(b)) scales linearly with the magnetic field as observed in [60].

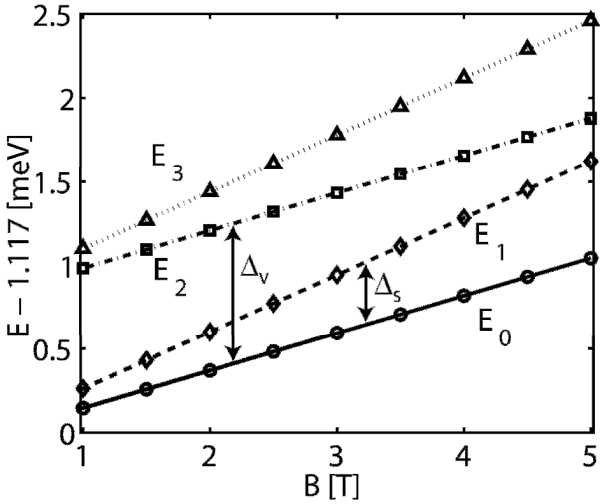

Figure 4.2. Evolution of first four conduction band states of a 10nm thick strained flat Si quantum well in magnetic field. Valley splitting is higher than spin splitting in the low magnetic field regime shown here. Inherent electric field of 9MV/m present in Si/SiGe heterostructure due to modulation doping is included.

4.4. Valley splitting in finite barrier quantum wells

Almost all fabricated quantum wells are realized in heterostructure configurations, few examples are Si quantum wells embedded in $Si_{1-x}Ge_x$ barriers, GaAs embedded in $Al_xGa_{1-x}As$ barriers and $In_xGa_{1-x}As$ embedded GaAs barriers. These barrier materials have higher bandgap than the quantum well material due to which electrons are confined in a quantum well. Although most of the electron wavefunction is localized in the quantum well it partly diffuses into the barriers. The extent of the penetration of the electron wavefunction into the barrier is controlled by the conduction band-offset between the quantum well and the barrier. Valley splitting in a finite barrier quantum well depends on the extent of the wavefunction penetration into the barrier. Therefore valley splitting in a Si quantum well depends on Si/SiGe barrier height. Disorders such as random placement of Ge atoms in SiGe and resulting inhomogeneous strain also affect valley splitting. First a simple case of smooth barriers is considered. Figure 4.4 shows the effect of the barrier height (or conduction band-offset) on valley splitting in a 2^o miscut quantum well. The smooth barriers are made up of a fictitious material formed by raising values of onsite tight-binding parameters of Si by energies equal to the conduction band-

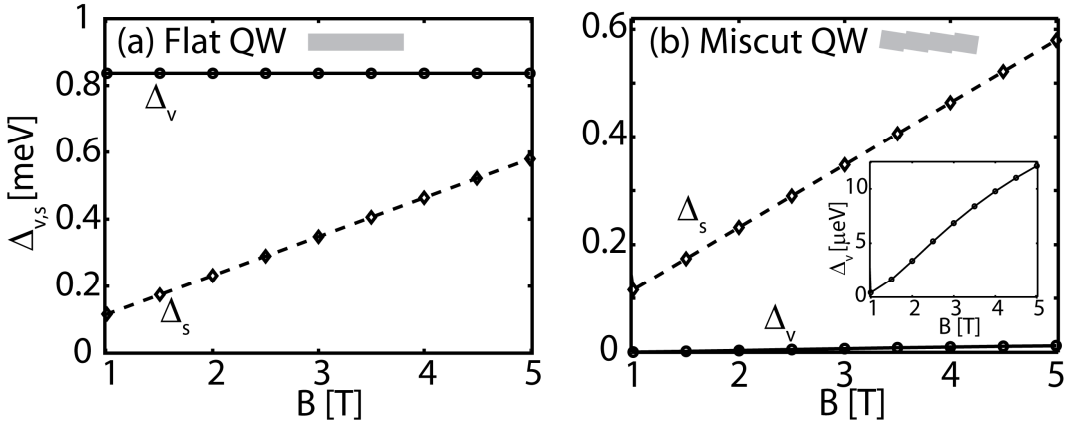

Figure 4.3. Valley and spin splittings versus magnetic field in a 10nm thick strained Si quantum wells grown on (a) flat and (b) 2^o miscut substrates. (a) Valley splitting in a flat quantum well is higher than spin-splitting in low magnetic field regime. (b) Valley splitting in a miscut quantum well is lower than that in a flat quantum well. Inset shows the zoomed-in view of the valley splitting in a miscut quantum well. Confinement in *x*-direction is 150nm to avoid interference of hard-wall boundary conditions with magnetic field confinement. All simulations include inherent electric field of 9MV/m present in Si/SiGe heterostructure due to modulation doping.

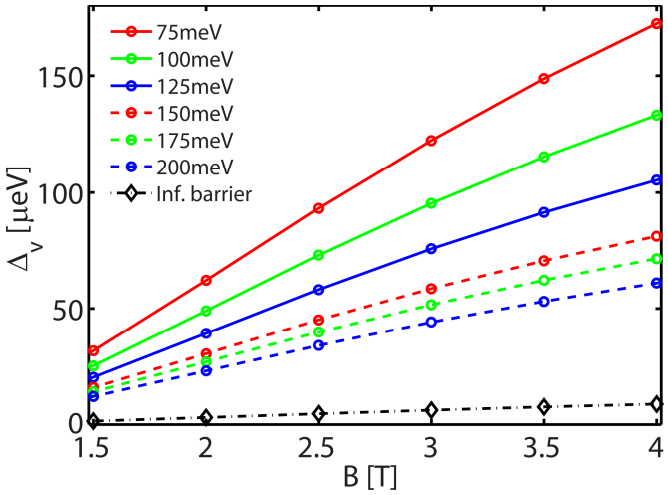

Figure 4.4. Valley splitting in 2° miscut Si quantum wells embedded in a smooth barrier. Valley splitting is higher for lower barrier height. Electric field = 9 MV/m. Thickness of Si quantum well is 10nm and it is embedded in a 3 nm thick barrier on both (top and bottom) sides. Confinement in *x*-direction is 150nm.

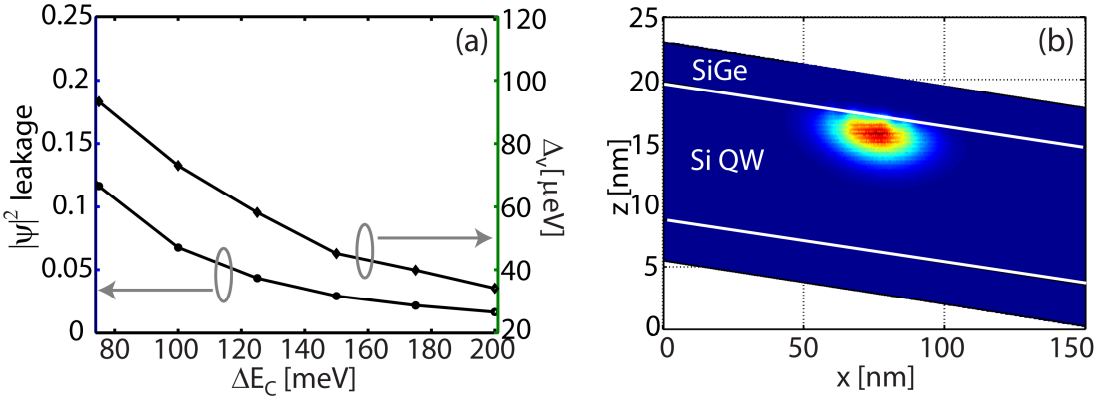

Figure 4.5. (a) Valley splitting in a 10nm thick 2° miscut Si quantum well as a function of the barrier height. Valley splitting follows the same trend as the wavefunction leakage in the barrier. (b) Ground state wavefunction of a 2° miscut Si quantum well embedded in a smooth barrier of height 150meV. Magnetic field: B = 2.5T, Electric field: E_z = 9 MV/m.

offset. Valley splitting increases as the barrier height is lowered.

Dependence of valley splitting on the barrier height can be qualitatively explained by comparing the wavefunction leakage into the barrier. Figure 4.5(a) shows valley splitting as a function of the barrier height at magnetic field of 2.5T. Valley splitting versus barrier height is plotted on the right hand side *y*-axis while the corresponding fraction of the wavefunction leakage into the barrier is plotted on the left hand side *y*-axis. Valley

splitting follows the trend of wavefunction leakage into the barrier. Same behavior is observed for small barrier heights (considered here) in finite barrier flat quantum wells. A typical ground state wavefunction is shown in Figure 3.5(b). The wavefunction is pulled towards top Si/SiGe interface because of the built-in electric field due to the modulation doping present in these heterostructures. The wavefunction leakage is the fraction of the total wavefuntion probability that resides in the barrier. A detailed study of valley splitting in finite barrier quantum wells is published in [83].

4.5. Valley splitting in SiGe/Si/SiGe quantum wells with disorder

This section presents valley splitting calculations in 2^{o} miscut strained SiGe/Si/SiGe quantum wells. The results are compared with experimental measurements of [60].

4.5.1. Disorders in Si/SiGe heterostructures

Alloy and surface roughness disorders are inherently present in experimental realizations of devices fabricated in Si/SiGe heterostructures. Alloy disorder refers to the fact that Ge atoms are randomly distributed in SiGe. Figure 4.6(b) shows a zoomed view of a random interface configuration when the Si/SiGe interface is represented atomistically. This random distribution of Ge produces inhomogeneous strain due different Si-Si, Ge-Ge, and Si-Ge equilibrium bond lengths. Atom disorder is automatically taken into account because every atom is represented explicitly in the tight-binding calculations and Si and Ge have different tight-binding parameters. Inhomogeneous strain disorder is incorporated in the tight-binding Hamiltonian according to the prescription of [80].

Another type of disorder which is always present in the quantum wells grown on vicinal substrates is the step-roughness disorder. Miscut quantum wells grown on (001) vicinal substrates can be assumed in a first approximation, to be made up of repeated regular steps (Figure 4.6(a),(c)) creating a tilt angle (θ_T) to the [100] direction in the growth plane, causing the unit cell to be large compared to that of a flat quantum well. However, Si quantum wells grown on miscut substrates have an irregular step-like

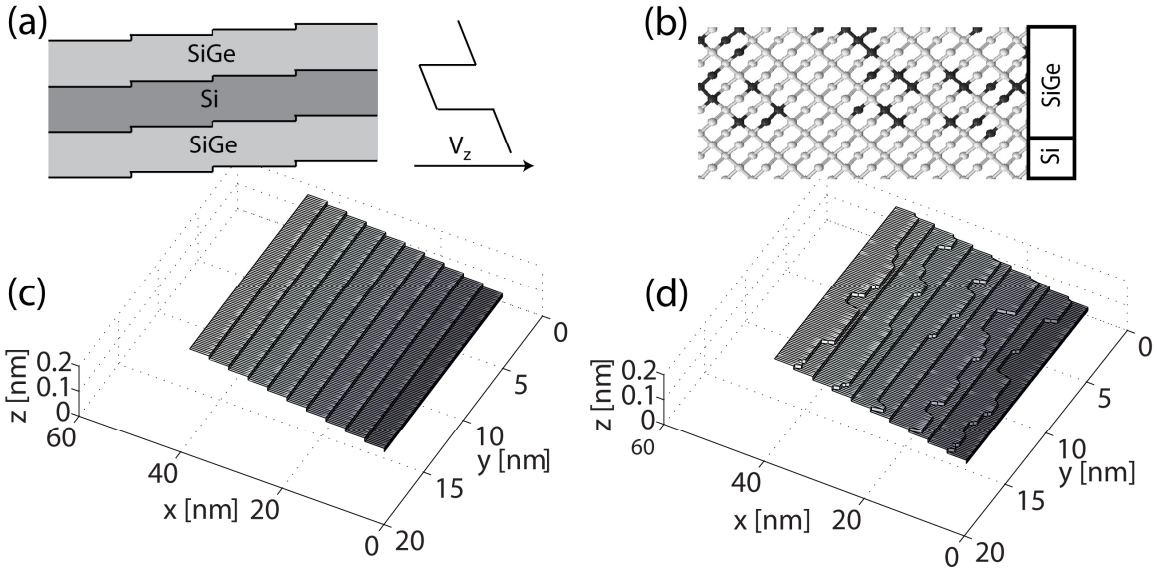

Figure 4.6. (a) Schematic of SiGe/Si/SiGe heterostructure and confining potential (V_z) along growth direction. (b) Atomic scale representation of a disordered Si/SiGe interface. (c) Ideal steps along [100] direction. (d) Schematic of the step disorder along y-direction in Si/SiGe heterostructures grown on 2^o miscut substrates. Flat and disordered steps alternate [48, 84].

structure along the [100] crystallographic axis (Figure 4.6(d)) [48, 84]. This irregular structure results from the fact that the mono-atomic steps on the miscut surface interact and settle into the minimum surface energy configuration. This minimum energy configuration has alternating regular and rough steps as shown in Figure 4.6(d). The deviation-deviation correlation function of the step edges, and the distribution of terrace lengths of miscut substrates are extensively studied using scanning tunneling microscopy (STM) and statistical physics [48, 84]. Step roughness is included in the tight-binding simulations by generating the step roughness profiles following the prescription of [84] and imposing these profiles on the atomistic grid.

4.5.2. Simulation domain

Schematics of the strain and the electronic structure domain for both flat and miscut quantum wells are shown in Figure 4.7. Strain disorder is known to have long range

nature [85]. 40nm of SiGe layers are included on the top and the bottom of 10nm Si quantum well for strain calculations. This SiGe thickness is sufficient to model the long-range strain disorder, where the detailed strain boundary conditions are not important and the SiGe volume exceeds the Si quantum well volume significantly. SiGe buffers provide electronic confinement of approximately 100meV due to which the electronic states of interest in this problem are spatially confined to the quantum well and only weakly penetrate into the SiGe buffer. Therefore one can safely reduce the electronic structure domain to 3nm of SiGe buffer around the Si quantum well. For this setup the strain calculation requires 10 million atoms and the electronic structure calculation requires 2 million atoms. For the idealized geometries without a SiGe buffer, a homogeneous lattice distortion of $\varepsilon_{\parallel} = 0.013$ is assumed throughout the Si quantum well, as approximated from the full SiGe buffer system calculation and hard wall boundary conditions are assumed in *z*-direction. Such electronic structure calculations require about 0.5 million atoms. Following the discussion of section 4.2.3, the lateral extension of the strain and the electronic structure domain is set to 150nm. The lateral confinement of the wavefunction is dominated by magnetic field and the closed boundary conditions applied along *x*-direction do not modify the energy spectrum of a quantum well. Modulation doping in Si/SiGe heterostructures induces built-in electric field. In the simulations performed here an electric field of 9 MV/m is assumed in the growth direction. Quantum wells extend 15 nm along *y*-direction to take into account the step roughness disorder shown in Figure 4.6(d). SiGe alloy disorder is also assumed to be quasi-periodic with the period of 15nm, which is sufficient to capture the effect of inhomogeneous strain disorder. The choice of asymmetric magnetic gauge ($A = Bx\hat{y}$) is still valid because the electronic domain is assumed to be quasiperiodic in y-direction.

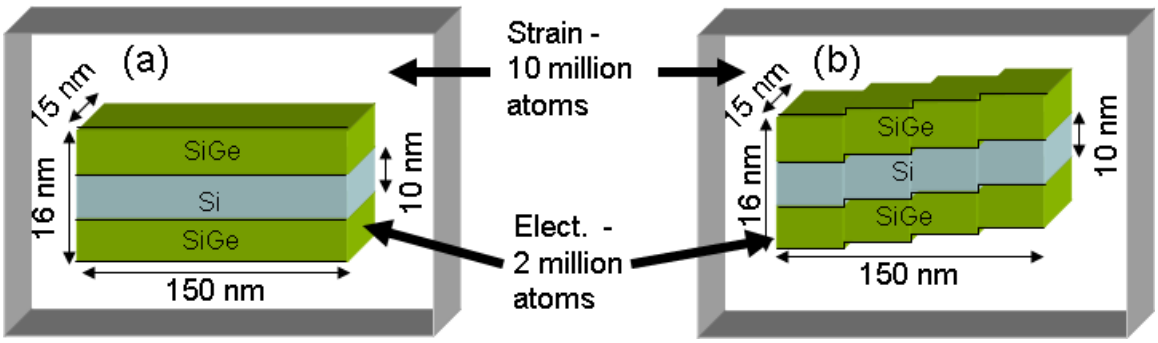

Figure 4.7. Schematics of simulated quantum wells: (a) Flat and (b) Miscut quantum wells embedded in SiGe barriers. Strain domain extends 40 nm on either (top and bottom) sides of a Si quantum well while electronic structures domain extends 3 nm on either sides. Quantum wells are 15nm long in *y*-direction to take into account the alloy and the step roughness disorder.

4.5.3. Comparison with experimental data

In the presence of the applied magnetic field the 2-DEG is quantized in Landau levels (LLs). The valley-degeneracy of LLs is broken in the presence of sharp confinement due to Si/SiGe interfaces in a quantum well. The valley splitting of the first LL in flat and 2^o miscut quantum wells is plotted as a function of the applied magnetic field in Figure 4.8. The valley splitting in ideal (no SiGe disorder) miscut quantum wells is 2 orders of magnitude less than that in flat quantum wells. The finite SiGe barriers enhance valley splitting in flat as well as miscut quantum wells. This is expected because as described in section 4.4 valley splitting is a decreasing function of the barrier height for smaller barrier heights. Addition of SiGe buffers to the electronic structure calculation domain in 2^o miscut quantum wells increases valley splitting, however, experimentally observed valley splitting is slightly higher.

If the step-roughness disorder alone is included in the simulations, the computed valley splitting is higher compared to that of an ideal miscut quantum well (Figure 4.9 (a)). In these calculations surface roughness model described in section 4.5.1 is used and the uniform biaxial strain of $\varepsilon_{\parallel} = 0.013$ which corresponds to $Si_{0.7}Ge_{0.3}$ buffer composition is assumed. This valley splitting, however, is slightly smaller than the experimentally measured valley splitting. This discrepancy can be answered by adding SiGe buffers in the electronic structure simulation domain. 3nm of SiGe buffer is

included on top and bottom of the Si quantum well to take into account the wavefunction penetration into the finite barrier quantum well buffers. Strain computation domain has same the *x* and *y* dimensions as the electronic structure domain. To take into account the long range nature of strain 40nm of SiGe buffer is included on both sides of Si quantum well. As shown in Figure 4.9(b), the valley splitting computed by taking both step and alloy disorders into account is found to match closely to the experimentally measured values. Figure 4.10 shows valley splitting data used to generate averaged valley splitting plots in Figure 4.8(b) and Figure 4.9. In general, valley splitting increases when SiGe buffer and step roughness are included in the simulation domain. Valley splitting values for each disorder sample are different because of local atomic scale variations at the

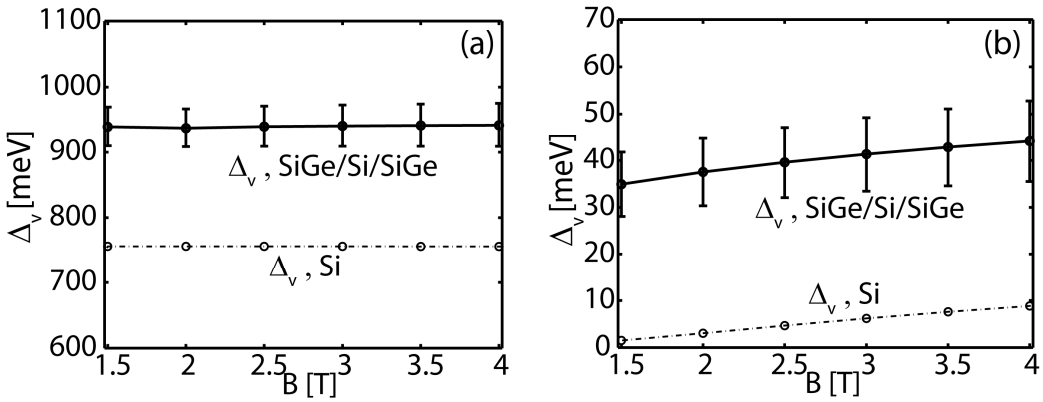

Figure 4.8. Enhancement in valley splitting due to SiGe buffers in flat (a) and 2° miscut (b) quantum wells. Valley splitting is higher due to the penetration of wavefunction in SiGe buffers.

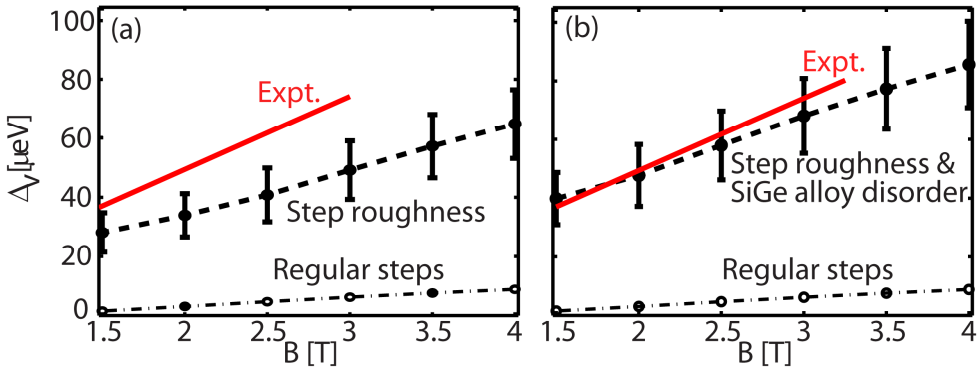

Figure 4.9. (a) Enhancement in valley splitting due to step disorder alone. SiGe buffers are not included in the electronic structure simulation domain. (b) Valley splitting calculated with both disorders matches with the experimentally observed valley splitting.

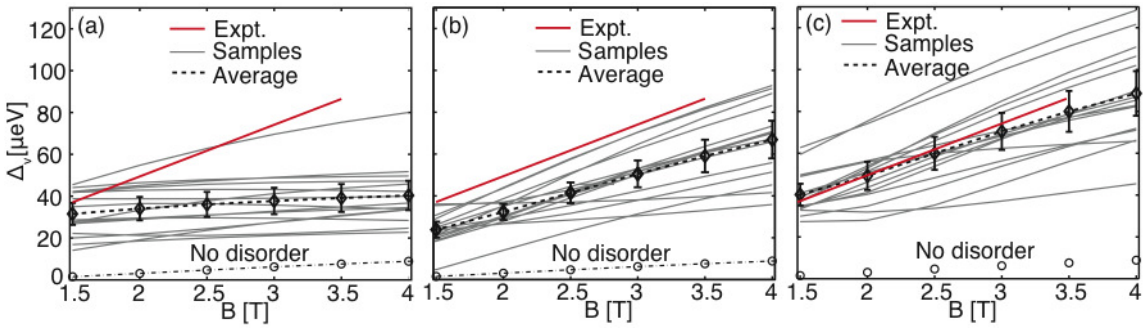

Figure 4.10. Valley splitting in the presence of various disorders. (a) Alloy disorders in SiGe buffer layer. (b) Step disorder at the Si/SiGe interface. (c) Both alloy and step disorder. Fifteen samples of each disorder are simulated.

Si/SiGe interface. The lateral extent of the experimental sample is much larger compared to the simulation domain of Figure 4.7. To mimic the large lateral extent of the experimental samples, several disorder samples are simulated. The average value gives better approximation to the experimental measurement.

4.6. Conclusion

Earlier theories of valley splitting, which do not include miscut surfaces [58, 59] over-predict the value of the valley splitting compared to experimental data [60] while perfect miscut quantum wells underpredict the observed valley splitting by an order of magnitude. Effective mass formalisms [71, 72] suggest that disorder in the miscuts enhances the valley splitting, their model however uses ad-hoc structure dependent fitting parameters. Unlike effective mass, tight-binding can predict valley splitting without any ad-hoc fitting parameters [71]. The atomistic representation of the confinement buffer and the local step disorder is therefore an essential ingredient in high precision calculations of valley splitting in heterostructures.

5. ZONE-UNFOLDING FORMALISM

5.1. Introduction

Bandstructure calculations are critical in evaluation of device properties. Such calculations are typically performed within a plane-wave basis in an infinitely periodic system. However, as devices are scaled to nanoscale sharp variations on atomic scale play an increasingly important role in determining the electronic structure and transport properties. At the nanoscale, methods that employ local basis sets are theoretically more intuitive and practically more convenient compared to plane-wave basis set methods. Local basis set methods such as semi-empirical tight-binding have smaller computational requirements than corresponding pseudopotential methods for the calculation of electronic structure in multimillion-atom systems [31-38]. The power of tight-binding method is fully utilized in the large-scale calculations performed on nanostructures with sharp interfaces and atomistic alloy-disorder. These methods have been successfully applied in treating SiGe/Si/SiGe quantum wells [24], impurities [86-89], quantum dots [90], and alloys [22].

Traditional tight-binding bandstructure calculations in periodic systems are performed by choosing the smallest repeating unit cell along the periodic directions of the semiconductor crystal. However, nanostructures are seldom perfect. Examples include quantum dots or wires made of alloys or having rough surfaces, nanoscale FETs with inhomogeneous strain distributions, or simply three-dimensional random alloys.

In disordered structures translational symmetry is broken and the small cells differ from each other. The traditional averaging methods such as the virtual crystal approximation (VCA) can not account for bandstructure effects of random alloying [22]. In nanostructures such effects due local disorders are more pronounced. Conventionally alloy disorders are treated in the supercell framework. Such a supercell contains many

small cells which are different from each other in terms of atom/strain distribution (Figure 5.1(a,b)). Thus the supercell itself has a random disorder in it. If the supercell is sufficiently large, one would expect it to represent the true disorder in the device. However, due to Brillouin zone-folding such supercell calculations only deliver the energy spectrum which can only be used to extract absolute band-minima of the disordered nanostructures (Figure 5.1(c)). Supercell calculations alone can not deliver the transport parameters such as effective masses and relative locations of valley-minima. This chapter outlines a zone-unfolding method which can be used to unfold the supercell bandstructure (Figure 5.1(c)). Unfolded bandstructure ((d)) can be used to calculate the transport parameters such as effective mass and relative locations of valley-minima, which can be used in the top-of-the-barrier-model to predict device performance.

The use of a supercell representation is associated with the problem of zone-folding as shown schematically in Figure 5.2. The advantage of using supercell representation is the fact that it can naturally account for disorders such as random alloying, impurities surface roughness. In the presence of disorders the very existence of the bandstructure becomes questionable. However, in the view that the experiments measure bandstructure properties such as band-gaps and effective masses one would expect that in such systems the bandstructure should exist in approximate sense.

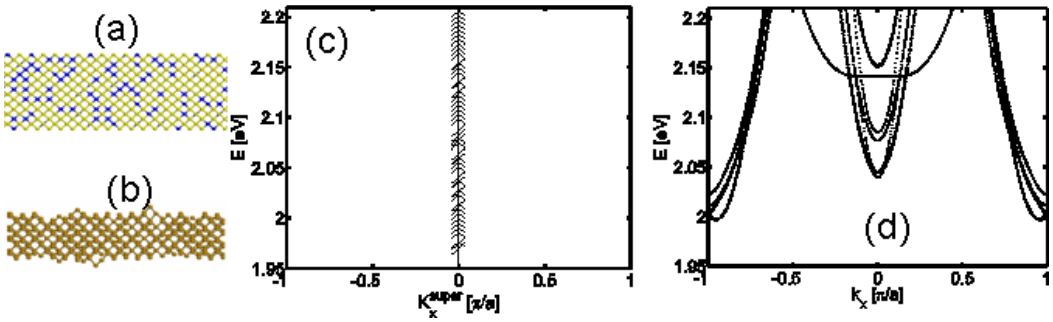

Figure 5.1. Supercell representation and its bandstructure: (a) Random alloy disorder and (b) surface roughness on atomic scale. Small cells (dotted rectangles) in the disordered supercell differ from each other. (c) Folded bandstructure of a supercell. Such bandstructure can only deliver the absolute band minima. (d) Unfolded bandstructure obtained by unfolding the supercell bandstructure. Unfolded bandstructure can be used to extract effective masses and relative positions of different valley minima.

This chapter presents the zone-unfolding method to obtain approximate bandstructures from the supercell eigenspectrum of the disordered nanostructures. In section 5.2 an analytical zone-unfolding calculation on a 1D toy nanowire is presented. Section 5.3 generalizes this method for realistic nanowires with finite cross-section while section 5.4 generalizes it to unfold irregular nanowire supercell bandstructures. The conclusion is presented in section 5.5.

5.2. 1D toy model

This section develops the zone-unfolding method for 1D toy nanowire. The method can be easily extended to realistically sized nanowires and to 2D and 3D nanostructures. We consider a 1D chain of atoms each having 2-orbitals, which are labeled as s and p. The onsite tight-binding parameters are denoted by ε_s and ε_p, while the two-center integrals are denoted by V_{ss}, V_{sp} and V_{pp}. Values of the tight-binding parameters are taken to be ε_s = 2.317 eV, ε_p = 3.683 eV, V_{ss} = 0.2705 eV, V_{sp} = 0.3415 eV, and V_{pp} = 0.9535 eV. Two different models are used to describe the same system, (i) Small cell model, where the smallest repeating unit contains only one atom. Since each atom has two orbitals this model leads to the two-band representation, (ii) Supercell model, where the

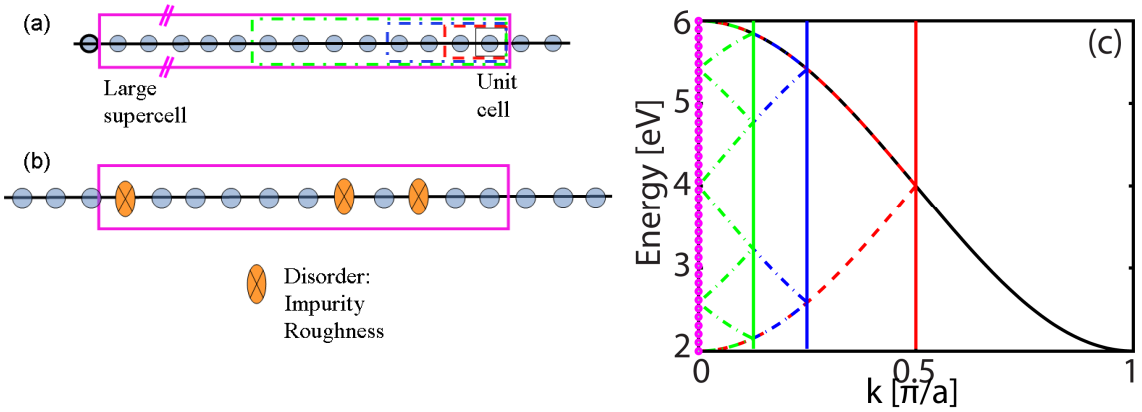

Figure 5.2. Zone-folding: (a) Unit cell (black) and subsequent larger supercells of a 1D toy nanowire. (b) Disorders such as impurities or roughness can be introduced in the supercell representation. (c) Bandstructures of a 1D toy nanowire plotted with repeating cells of Fig. (a). Brillouin zone of a supercell made up of N small cells spans from $-\pi/Na$ to π/Na, where a is length of a small cell. Thus the bandstructure of a small cell that spans from $-\pi/a$ to π/a is folded into a supercell Brillouin zone that spans from $-\pi/Na$ to π/Na.

smallest repeating unit contains three atoms. This model leads to the six-band representation. Both small cell and supercell models are physically identical and therefore, transformation of the bandstructure obtained from one model to the bandstructure obtained from another model is possible. The zone-unfolding algorithm gives us an efficient way to construct the small cell bandstructure from the supercell bandstructure.

5.2.1. Small cell bandstructure

In this representation the nanowire is assumed to be made up of repeating unit cell containing a single atom. Since each atom has two orbitals this model leads to the two band description of the electronic structure of the nanowire. The Hamiltonian for this unit cell can be written as

$$H = \begin{bmatrix} \varepsilon_s + 2V_{ss}\cos(ka) & -2iV_{sp}\sin(ka) \\ 2iV_{sp}\sin(ka) & \varepsilon_p + 2V_{pp}\cos(ka) \end{bmatrix} \tag{5.1}$$

The Brillouin zone of this representation extends over the range $-\pi/a < k \le \pi/a$, where a is length of the unit cell. This Hamiltonian can be easily diagonalized analytically. The eigenvalues are given by

$$\begin{aligned} E_1 &= \frac{\varepsilon_s + \varepsilon_p}{2} + (V_{ss} + V_{pp})\cos(ka) - \left[\left(\frac{\varepsilon_s - \varepsilon_p}{2} + (V_{ss} - V_{pp})\cos(ka) \right)^2 + 4{V_{sp}}^2 \sin^2(ka) \right]^{1/2} \\ E_2 &= \frac{\varepsilon_s + \varepsilon_p}{2} + (V_{ss} + V_{pp})\cos(ka) + \left[\left(\frac{\varepsilon_s - \varepsilon_p}{2} + (V_{ss} - V_{pp})\cos(ka) \right)^2 + 4{V_{sp}}^2 \sin^2(ka) \right]^{1/2} \end{aligned} \tag{5.2}$$

while corresponding eigenvectors of the Hamiltonian are

$$|\psi_{1,2}\rangle = \frac{1}{\sqrt{1 + (\alpha_{1,2})^2}} \begin{Bmatrix} 1 \\ \alpha_{1,2} \end{Bmatrix} \tag{5.3}$$

where $\alpha_{1,2} = \dfrac{\varepsilon_s + 2V_{ss}\cos(ka) - E_{1,2}}{2iV_{sp}\sin(ka)}$

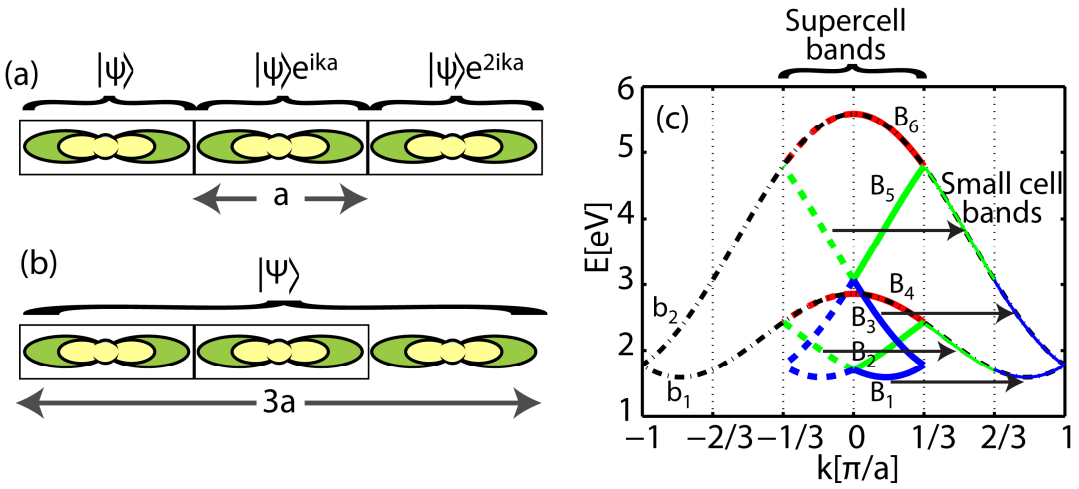

Figure 5.3. Illustration of zone-unfolding in a 2-band model: (a) Small cell contains one atom with 2 orbitals. (b) Supercell contains 3 small cells. (c) Bands in a supercell eigenspectrum are translated by appropriate reciprocal lattice vectors G_n to obtain small cell bandstructure.

The bandstructure (equation 5.2) of the small cell is plotted in Figure 5.3(c). In a two-band description the eigenvector of a 1-D chain has Bloch periodicity as shown in Figure 5.3(a). Since, the Bloch phase of the nearest unit cells differ by e^{ika} the eigenvector of the supercell (containing 3 small cells) in the small cell representation can be written as

$$\left|\psi_{1,2}\right\rangle = \frac{1}{\sqrt{3\left(1+\left(\alpha_{1,2}\right)^2\right)}}\begin{Bmatrix}\begin{Bmatrix}1\\ \alpha_{1,2}\end{Bmatrix}\\ \begin{Bmatrix}1\\ \alpha_{1,2}\end{Bmatrix}e^{ika}\\ \begin{Bmatrix}1\\ \alpha_{1,2}\end{Bmatrix}e^{2ika}\end{Bmatrix} \tag{5.4}$$

5.2.2. Supercell bandstructure

In this representation, we group three small unit cells to form a bigger supercell, which has three atoms or six orbitals per supercell. The length of the supercell is $A=3a$ where a is length of the small cell. The Hamiltonian matrix for the supercell representation is given by

$$H=\begin{bmatrix} \varepsilon_s & 0 & V_{ss} & -V_{sp} & V_{ss}e^{iKA} & V_{sp}e^{iKA} \\ 0 & \varepsilon_p & V_{sp} & V_{pp} & -V_{sp}e^{iKA} & V_{pp}e^{iKA} \\ V_{ss} & V_{sp} & \varepsilon_s & 0 & V_{ss} & -V_{sp} \\ -V_{sp} & V_{pp} & 0 & \varepsilon_p & V_{sp} & V_{pp} \\ V_{ss}e^{iKA} & -V_{sp}e^{iKA} & V_{ss} & V_{sp} & \varepsilon_s & 0 \\ V_{sp}e^{iKA} & V_{pp}e^{iKA} & -V_{sp} & V_{pp} & 0 & \varepsilon_p \end{bmatrix} \tag{5.5}$$

the Brillouin zone of this supercell extends over the range $-\pi/A < K \le \pi/A$ i.e. $-\pi/3a < K \le \pi/3a$, is one third of that of the small cell. Six eigenvalues of this Hamiltonian are given by

$$\begin{aligned} E_{2n-1} &= \frac{\varepsilon_s+\varepsilon_p}{2}+\left(V_{ss}+V_{pp}\right)\cos\left(K+G_n\right)a-\left[\begin{array}{l}\left(\dfrac{\varepsilon_s-\varepsilon_p}{2}+\left(V_{ss}-V_{pp}\right)\cos\left(K+G_n\right)a\right)^2 \\ +4{V_{sp}}^2\sin^2\left(K+G_n\right)a\end{array}\right]^{1/2} \\ E_{2n} &= \frac{\varepsilon_s+\varepsilon_p}{2}+\left(V_{ss}+V_{pp}\right)\cos\left(K+G_n\right)a+\left[\begin{array}{l}\left(\dfrac{\varepsilon_s-\varepsilon_p}{2}+\left(V_{ss}-V_{pp}\right)\cos\left(K+G_n\right)a\right)^2 \\ +4{V_{sp}}^2\sin^2\left(K+G_n\right)a\end{array}\right]^{1/2} \end{aligned} \tag{5.6}$$

where

$$G_n=\frac{2\pi}{3a}(n-2),\ n=1,2,3 \tag{5.7}$$

Examining the periodicity of the supercell the eigenvectors can be written as follows:

$$\left|\Psi_{2n-1,2n}\right\rangle=\frac{1}{\sqrt{3\left(1+\left(\gamma_{2n-1,2n}\right)^2\right)}}\left\{\begin{array}{l}\left\{\begin{array}{l}1\\ \gamma_{2n-1,2n}\end{array}\right\} \\ \left\{\begin{array}{l}1\\ \gamma_{2n-1,2n}\end{array}\right\}e^{i\left(K+G_n\right)a} \\ \left\{\begin{array}{l}1\\ \gamma_{2n-1,2n}\end{array}\right\}e^{2i\left(K+G_n\right)a}\end{array}\right\}\ \text{for}\ n=1,2,3 \tag{5.8}$$

where $\gamma_{2n-1,2n}=\dfrac{\varepsilon_s+2V_{ss}\cos\left(K+G_n\right)a-E_{2n-1,2n}}{2iV_{sp}\sin\left(K+G_n\right)a}$

The bandstructure (equation 5.6) of the supercell is plotted in Figure 5.3(c). The Brillouin zone of the supercell is one third of that of the small cell. However, supercell bandstructure contains thrice as many bands as the small cell bandstructure. Thus both representations contain the same physical information. The small cell bandstructure can be retrieved from the supercell bandstructure by shifting the bands by appropriate reciprocal lattice vectors, G_n's. Alternatively, supercell bandstructure can be obtained by shifting the small cell bands back into the supercell Brillouin zone. The zone-unfolding method to retrieve the small cell bandstructure from the supercell bandstructure is described in the next section.

5.3. The zone-unfolding method for perfect supercells

5.3.1. Method

The supercell and the small cell representations of a perfect elemental nanowire are equivalent because all small cells in a perfect supercell are identical. The eigenstate of a nanowire at energy E_p and wavevector K can be expressed as the following Bloch sum:

$$\left|\Psi_p(K)\right\rangle = \sum_{j=1}^{N_S}\sum_{\alpha,\mu}^{N_{O,cell}}\sum_{l=1}^{N_C}\frac{e^{iKR_j}}{\sqrt{N_S}}\beta_{l,p}^{\alpha,\mu}(K)\left|\alpha,\mu;R_j+\rho_l\right\rangle \tag{5.9}$$

Here, the summation index j runs over supercells and l runs over each small cell in a supercell, There are $N_{O,cell}$ orbitals per small cell, which are indexed by orbital type α (s, p_x, etc.) and atom within the small cell μ. There are N_C small cells in a supercell and N_S supercells in the nanowire. R_j and ρ_l denote the positions of the supercell along the nanowire and the position of small cell in this supercell respectively. β stands for the eigenvector coefficients.

Suppose that the same eigenstate (of energy E_p) in the small cell representation appears at a wavevector $k=K+G_n$. This eigenstate can be written as follows:

$$\left|\psi_p(K+G_n)\right\rangle = \sum_{j=1}^{N_S}\sum_{\alpha,\mu}^{N_{O,cell}}\sum_{l=1}^{N_C}\frac{e^{iKR_j}}{\sqrt{N_S N_C}}b_p^{\alpha,\mu}(K+G_n)e^{i(K+G_n)\rho_l}\left|\alpha,\mu;R_j+\rho_l\right\rangle \tag{5.10}$$

Here, G_n is a supercell reciprocal lattice vector such that $k=K+G_n$ lies in the first BZ of the small cell. G_n's are called allowed wavevectors. Note that the underlying periodicity of eigenvector coefficients $\beta_{l,p}^{\alpha,\mu}$ in equation (5.9) is a supercell periodicity while eigenvector coefficients $b_p^{\alpha,\mu}$ in the Bloch state of equation (5.10) have a small cell periodicity.

The determination of the allowed wavevectors G_n's depends on the supercell geometry, specifically on the number of small cells (N_C) contained in the supercell. Small cell BZ ranges from $-\pi/a$ to $+\pi/a$ while supercell BZ ranges from $-\pi/N_Ca$ to $+\pi/N_Ca$. Each band in the small cell eigenspectrum appears in N_C segments in the supercell eigenspectrum. To obtain the small cell eigenspectrum each of these segments are translated from the wavevector K in supercell BZ to the wavevector k in small cell BZ by the allowed wavevectors i.e. $k=K+G_{n(j)}$. There are N_C allowed wavevectors and they are given by the following prescription:

$$G_{n(j)} = \frac{2\pi}{N_C a} n(j)$$
$$n(j) = \begin{cases} -\frac{N_C-2}{2},\ldots,-1,0,1,\ldots,\frac{N_C}{2}, N_C even \\ -\frac{N_C-1}{2},\ldots,-1,0,1,\ldots,\frac{N_C-1}{2}, N_C odd \end{cases} \quad (5.11)$$

Note that all $G_{n(j)}$'s satisfy the Born-vonKarmann boundary condition $\Psi_{K+G}(x+N_Ca)=\Psi_{K+G}(x)$ and lie in the first BZ of the small cell. These criteria are used to calculate $G_{n(j)}$ [42, 70, 91]. Supercell and small cell eigenstates of equations (5.9) and (5.10) are eigenstates of the same periodic nanowire structure expressed as linear combinations of the same set of atomic-like orbitals $\left(\left|\alpha,\mu;R_j+\rho_l\right\rangle\right)$. Hence they must be linear combinations of each other, which can be expressed as

$$\left|\Psi_p(K)\right\rangle = \sum_{j=1}^{N_C} a_{p,n(j)} \left|\psi_{p,n}\left(K+G_{n(j)}\right)\right\rangle \quad (5.12)$$

In a pure elemental nanowire many of the coefficients $a_{p,n(j)}$ are zero. Non-zero $a_{p,n(j)}$ are interpreted as that the small cell state of energy E_p and wavevector $k=K+G_{n(j)}$ contributes

to the supercell state of the same energy E_p and the wavevector K. Equation (5.12) is solved for $a_{p,n(j)}$'s by employing the procedure of [42], which is briefly described here. Note that the supercell eigenstates, $\left|\Psi_p(K)\right\rangle$ are known from the supercell eigenspectrum calculation while expansion coefficients $a_{p,n(j)}$ and small cell eigenstates, $\left|\psi_{p,n}\left(K+G_{n(j)}\right)\right\rangle$ are unknown. However, as described below one can exploit the normalization property of the small cell eigenstates to compute $|a_{p,n(j)}|^2$, which is sufficient to obtain unfolded supercell bandstructure.

Substituting equations (5.9) and (5.10) into equation (5.12), projecting out the component for the ket $\left|\alpha,\mu;R_j+\rho_l\right\rangle$, and rearranging gives N_C equations for l=1,2,…,N_C

$$e^{-iK.\rho_l}\beta_{l,p}^{(\alpha,\mu)}(K)=\frac{1}{\sqrt{N_C}}\sum_{j=1}^{N_C}a_{p,n(j)}b_p^{(\alpha,\mu)}\left(K+G_{n(j)}\right)e^{iG_{n(j)}.\rho_l} \tag{5.13}$$

These equations expressed in the matrix form

$$B_p^{(\alpha,\mu)}(K)=\underline{U}\cdot C_p^{(\alpha,\mu)}(K) \tag{5.14}$$

where

$$B_p^{(\alpha,\mu)}(K)=\begin{bmatrix} e^{-iK.\rho_l}\beta_{1,p}^{(\alpha,\mu)}(K) \\ \vdots \\ \vdots \\ e^{-iK.\rho_{N_C}}\beta_{N_C,p}^{(\alpha,\mu)}(K)\end{bmatrix},$$

$$C_p^{(\alpha,\mu)}(K)=\begin{bmatrix} a_{p,n(1)}b_p^{(\alpha,\mu)}\left(K+G_{n(1)}\right) \\ \vdots \\ \vdots \\ a_{p,n(N_C)}b_p^{(\alpha,\mu)}\left(K+G_{n(N_C)}\right)\end{bmatrix} \tag{5.15}$$

and the $N_C\times N_C$ unitary matrix $\underline{U}$ is give by

$$\underline{U}=\frac{1}{\sqrt{N_C}}\begin{bmatrix} e^{i\rho_1.G_{n(1)}} & e^{i\rho_1.G_{n(2)}} & \cdots & e^{i\rho_1.G_{n(N_C)}} \\ e^{i\rho_2.G_{n(1)}} & \ddots & \ddots & e^{i\rho_2.G_{n(N_C)}} \\ \vdots & \ddots & \ddots & \vdots \\ e^{i\rho_{N_C}.G_{n(1)}} & \cdots & \cdots & e^{i\rho_{N_C}.G_{n(N_C)}}\end{bmatrix} \tag{5.16}$$

Exploiting the unitarity of $\underline{U}$, equation (5.14) is repeatedly solved for each supercell eigenvector of energy, E_p and (atom, orbital) combination (α,μ) to get $C_p^{(\alpha,\mu)}(K)$

$$C_p^{(\alpha,\mu)}(K) = \underline{U}^{+} \cdot B_p^{(\alpha,\mu)}(K) \tag{5.17}$$

Using the normalization of the bulk eigenstates, the coefficients of linear combinations $a_{p,n(j)}$ are obtained by summing over atoms and orbitals for a fixed energy, E_p, and supercell reciprocal lattice vector $G_{n(j)}$

$$\begin{aligned} \left|a_{p,n(j)}\right| &= \sqrt{\sum_{\alpha,\mu} \left|\left[C_p^{(\alpha,\mu)}\right]_j\right|^2}\,, \\ \left[C_p^{(\alpha,\mu)}\right]_j &= a_{p,n(j)} b_p^{(\alpha,\mu)}\left(K + G_{n(j)}\right) \end{aligned} \tag{5.18}$$

The coefficients of linear combinations $a_{p,n(j)}$ give the contribution of each small cell state of energy, E_p to a supercell state of the same energy, E_p. To obtain the small cell bandstructure, for each non-zero $a_{p,n(j)}$ energy (Ep) is translated from the wavevector K in supercell BZ to the wavevector k in small cell BZ by the corresponding allowed wavevector i.e. $k=K+G_{n(j)}$.

5.3.2. Analytical zone-unfolding calculation in 1D toy example

In the 1D toy example described in section 5.2, a supercell is composed of three small cells (i.e. N_C=3). For this simple case of the 1D toy model, the zone-unfolding calculation can be done analytically. The real space small cell lattice vectors are given by $\rho_1 = 0$, $\rho_2 = a$, and $\rho_3 = 2a$ while the allowed wavevectors for this supercell are respectively given by $G_{n(1)} = -2\pi/3a$, $G_{n(2)} = 0$, and $G_{n(3)} = 2\pi/3a$. The basis transformation between the supercell and small cell eigenstates of energy E_p can be expressed as

$$\left|\Psi_p(K)\right\rangle = \sum_{j=1}^{3} a_{p,n(j)} \left|\psi_{p,n(j)}\left(K + G_{n(j)}\right)\right\rangle \tag{5.19}$$

The expansion coefficients, $a_{p,n(j)}$'s can be obtained by following the procedure of section 5.3.1. Here, an analytical zone-unfolding calculation is shown for the supercell band, B_1 of Figure 5.3(c). The supercell eigenvector for this band is given by

$$\left|\Psi_1\right\rangle = \frac{1}{\sqrt{3\left(1+(\gamma_1)^2\right)}}\left\{\begin{Bmatrix}1\\ \gamma_1\end{Bmatrix} \begin{Bmatrix}1\\ \gamma_1\end{Bmatrix} e^{i\left(K+\frac{2\pi}{3a}\right)a} \begin{Bmatrix}1\\ \gamma_1\end{Bmatrix} e^{2i\left(K+\frac{2\pi}{3a}\right)a}\right\} \tag{5.20}$$

where γ_1 is given by equation (5.8). Following the procedure of section 5.3.1, equation (5.19) can be cast in matrix form of equation (5.17). Substituting $\underline{U}^{\dagger}$ and $\left|\Psi_1\right\rangle$ into equation (5.19) gives

Supercell band index (p)	$G_{n(1)} = -2\pi/3a$	$G_{n(2)} = 0$	$G_{n(3)} = 2\pi/3a$
	$\lvert a_{p,n(1)}\rvert^2$	$\lvert a_{p,n(2)}\rvert^2$	$\lvert a_{p,n(3)}\rvert^2$
1	0	0	1
2	1	0	0
3	0	0	1
4	0	1	0
5	1	0	0
6	0	1	0

Figure 5.4. A table of coefficients $a_{p,n(j)}$'s in equation 5.19.

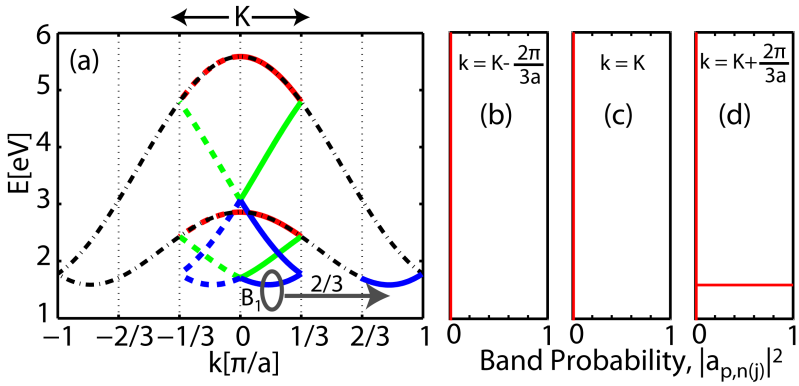

Figure 5.5. Unfolding of supercell band B_1. The contributions ($a_{p,n(j)}$) from small cell eigenvectors at $k=K+G_{n(j)}$ for $G_{n(j)}=-2\pi/3a,\ 0,\ 2\pi/3a$ are plotted in Figures (b), (c), and (d) respectively. Since the only non-zero contribution comes from $G=2\pi/3a$, the band B_1 is shifted by the reciprocal lattice vector $G=2\pi/3a$ from supercell wavevector K to a small cell wavevector $k=K+2\pi/3a$ in the small cell Brillouin zone.

$$C_1^{(\alpha,\mu)}(K)=\frac{1}{\sqrt{3}}\begin{bmatrix}1 & e^{i\frac{2\pi}{3}} & e^{i\frac{4\pi}{3}}\\ 1 & 1 & 1\\ e^{-i\rho_1.G_{n(N_C)}} & e^{-i\frac{2\pi}{3}} & e^{-i\frac{4\pi}{3}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{3\left(1+\gamma_1^{\,2}\right)}}\begin{pmatrix}1\\ \gamma_1\end{pmatrix}\\ e^{-iKa}\frac{1}{\sqrt{3\left(1+\gamma_1^{\,2}\right)}}\begin{pmatrix}e^{i\left(K+\frac{2\pi}{3a}\right)a}\\ \gamma_1 e^{i\left(K+\frac{2\pi}{3a}\right)a}\end{pmatrix}\\ e^{-2iKa}\frac{1}{\sqrt{3\left(1+\gamma_1^{\,2}\right)}}\begin{pmatrix}e^{2i\left(K+\frac{2\pi}{3a}\right)a}\\ \gamma_1 e^{2i\left(K+\frac{2\pi}{3a}\right)a}\end{pmatrix}\end{bmatrix} \tag{5.21}$$

which further simplifies to

$$C_1^{(\alpha,\mu)}=\frac{1}{3\sqrt{1+\gamma_1^{\,2}}}\begin{bmatrix}\begin{pmatrix}1\\ \gamma_1\end{pmatrix}+\begin{pmatrix}1\\ \gamma_1\end{pmatrix}e^{i\frac{4\pi}{3}}+\begin{pmatrix}1\\ \gamma_1\end{pmatrix}e^{i\frac{8\pi}{3}}\\ \begin{pmatrix}1\\ \gamma_1\end{pmatrix}+\begin{pmatrix}1\\ \gamma_1\end{pmatrix}e^{i\frac{2\pi}{3}}+\begin{pmatrix}1\\ \gamma_1\end{pmatrix}e^{i\frac{4\pi}{3}}\\ \begin{pmatrix}1\\ \gamma_1\end{pmatrix}+\begin{pmatrix}1\\ \gamma_1\end{pmatrix}+\begin{pmatrix}1\\ \gamma_1\end{pmatrix}\end{bmatrix}=\frac{1}{3\sqrt{1+\gamma_1^{\,2}}}\begin{bmatrix}0\\ 0\\ 3\begin{pmatrix}1\\ \gamma_1\end{pmatrix}\end{bmatrix} \tag{5.22}$$

Exploiting the normalization of small cell states, the expansion coefficients $|a_{1,n(j)}|$ are obtained as follows:

$$\left|a_{1,n(j)}\right|=\sqrt{\sum_{\alpha,\mu}^{N_{O,cell}}\left|\left[C_1^{(\alpha,\mu)}\right]_j\right|^2}=\begin{cases}\left|a_{1,n(1)}\right|=0\\ \left|a_{1,n(2)}\right|=0\\ \left|a_{1,n(3)}\right|=1\end{cases} \tag{5.23}$$

Since each supercell contains only one atom, the summation runs over *s* and *p* orbitals of an atom in each supercell. Similar calculations can be performed for all supercell bands, $B_1,\ldots,B_6$. The values of $|a_{p,n(j)}|$ for each supercell band are given in Figure 5.4. As discussed in [42], $|a_{p,n(j)}|^2$ is interpreted as the probability that the eigenvalue of energy E_p at the supercell wavevector K is shifted by the reciprocal lattice vector, $G_{n(j)}$ in the small cell eigenspectrum. This procedure is illustrated for supercell band B_1 in Figure 5.5, for B_2 in Figure 5.6, and for B_4 in Figure 5.7.

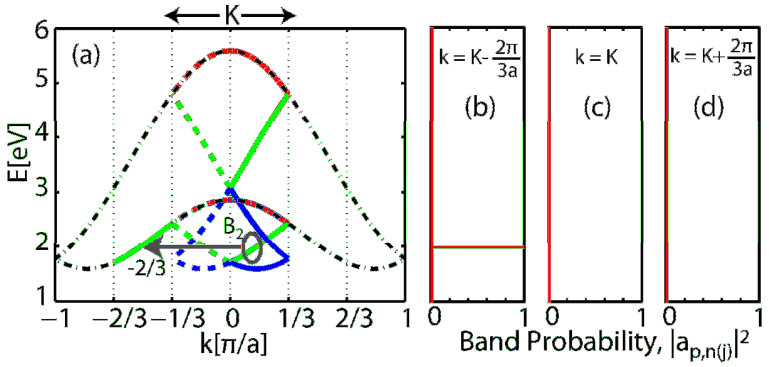

Figure 5.6. Unfolding of supercell band B_2. The contributions ($a_{p,n(j)}$) from small cell eigenvectors at $k=K+G_{n(j)}$ for $G_{n(j)}=-2\pi/3a,\ 0,\ 2\pi/3a$ are plotted in Figures (b), (c), and (d) respectively. Since the only non-zero contribution comes from $G=-2\pi/3a$, the band B_2 is shifted by the reciprocal lattice vector $G=-2\pi/3a$ from supercell wavevector K to a small cell wavevector $k=K-2\pi/3a$ in the small cell Brillouin zone.

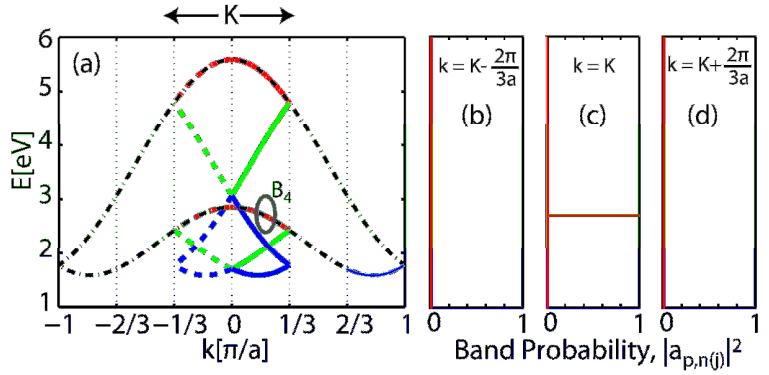

Figure 5.7. Unfolding of supercell band B_4. The contributions ($a_{p,n(j)}$) from small cell eigenvectors at $k=K+G_{n(j)}$ for $G_{n(j)}=-2\pi/3a,\ 0,\ 2\pi/3a$ are plotted in Figures (b), (c), and (d) respectively. Since the only non-zero contribution comes from $G=0$, the band B_4 is not shifted.

5.3.3. Perfect Si nanowire

The small cell and the supercell representations of a perfect elemental Si nanowire are shown in Figure 5.8. Figure 5.9(b) shows the bandstructure of a 2.2×2.2nm Si nanowire obtained by simulating a small cell similar to that shown by dotted rectangle in Figure 5.8 and by unfolding the supercell eigenspectrum of a 22×2.2×2.2nm supercell following the procedure of section 5.3.1. Both bandstructures match exactly as expected because no disorder is introduced in the supercell. Next section outlines the procedure to extract the small cell bandstructure from imperfect supercells.

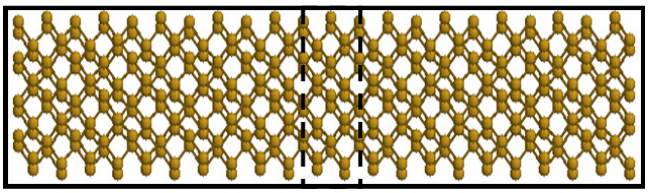

Figure 5.8. Two equivalent representations of a [100] Si nanowire. Small cell representation in which a wire is made up of a smallest repeating unit i.e. 4 atomic planes thick slab. Supercell representation in which a repeating unit consists of integer number of small cells. Here, nanowire supercell (solid lines) consists of 11 small cells (dotted lines).

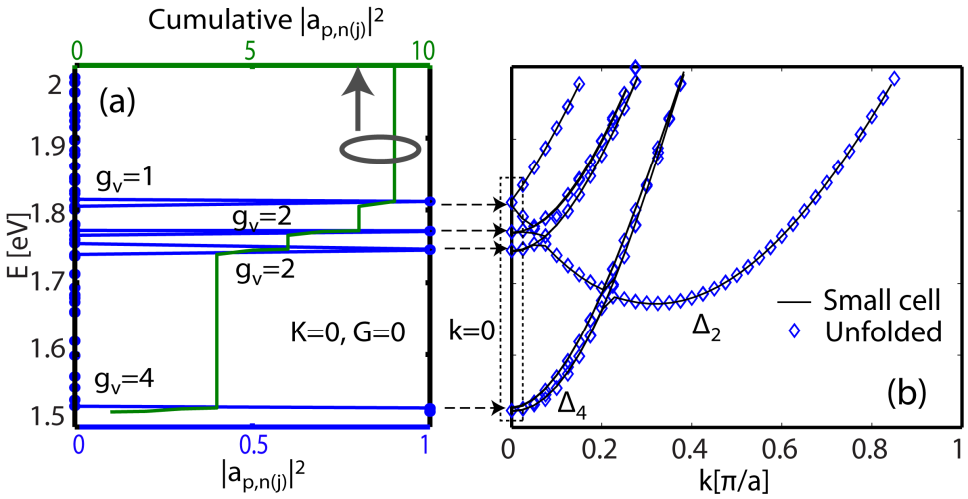

Figure 5.9. Small cell and unfolded supercell bandstructures of a perfect 2.2×2.2nm Si nanowire: (a) Coefficients $a_{p,n(j)}$ obtained by solving equation (5.12) for 22×2.2×2.2nm supercell wavevector $K=0$ and reciprocal lattice vector $G=0$. Energy values for which coefficients $a_{p,n(j)}$ are non-zero are shifted from $K=0$ in the supercell Brillouin zone to $k=K+G=0$ in the small cell Brillouin zone. This procedure is repeated for all $G_{n(j)}$ in equation (5.11) to obtain unfolded bandstructure in Figure (b).

5.4. Band-projection method for imperfect supercells

5.4.1. Method

Details of small cell band extraction in imperfect systems are given in [22]. Only the relevant procedure is summarized here. In an imperfect system the supercell state is generally a superposition of all small-cell Bloch states. Equation (5.12) gets modified to

$$\left|\Psi_p(K)\right\rangle = \sum_{\eta=1}^{N_{O,cell}} \sum_{j=1}^{N_C} a_{p;\eta,n(j)} \left|\psi_\eta\left(K+G_{n(j)}\right)\right\rangle \tag{5.24}$$

The summation index η runs over all $N_{O,cell}$ small cell bands. The number of bands, $N_{O,cell}$ in a small cell is given by number of atoms in a small cell multiplied by number of orbitals per atom. The supercell state of energy E_p is composed of small cell states of all energies E_η's. Generally the contribution of only few E_η's is significant and this contribution is quantified by expansion coefficients $a_{p;\eta,n(j)}$'s. One can project the small-cell states out of the supercell state by following similar procedure as described in section 5.2.2 to obtain the band probabilities

$$\begin{aligned} P_{p;n(j)} &= \sum_{\eta=1}^{N_{O,cell}} \left|a_{p;\eta,n(j)}\right|^2 = \sum_{\alpha,\mu}^{N_{O,cell}} \left|\left[C_p^{(\alpha,\mu)}\right]_j\right|^2, \\ \left[C_p^{(\alpha,\mu)}\right]_j &= \sum_{\eta=1}^{N_{O,cell}} a_{p;\eta,n(j)} b_\eta^{(\alpha,\mu)}\left(K+G_{n(j)}\right) \end{aligned} \tag{5.25}$$

Here $P_{p;n(j)}$ denotes the probability of there being a small cell state at energy E_p and small cell wavevector $k=K+G_{n(j)}$. The zone-unfolding procedure is summarized below:

1. Calculate the allowed supercell reciprocal lattice vectors $G_{n(j)}$.
2. Choose a small cell wavevector $k=K+G_{n(j)}$ at which to perform unfolding.
3. Compute and save the probability sums, $P_{p;n(j)}$, generated by repeatedly solving equation (5.25) for each supercell state of energy E_p.
4. Search for steps in the cumulative probability as a function of the energy, E_p.
5. Using the energies E_p and corresponding probability sums, $P_{p;n(j)}$, compute the mean energies $\tilde{E}$ and standard deviations ΔE between successive steps in the cumulative probability.

5.4.2. SiGe nanowire

Figure 5.10 shows the result of the band-projection method applied to unfold the bandstructure of 32.5×7.1×7.1nm $Si_{0.8}Ge_{0.2}$ alloy nanowire supercell to obtain bandstructure of 7.1×7.1nm $Si_{0.8}Ge_{0.2}$ alloy nanowire. Figure 5.10(a,b) show a typical

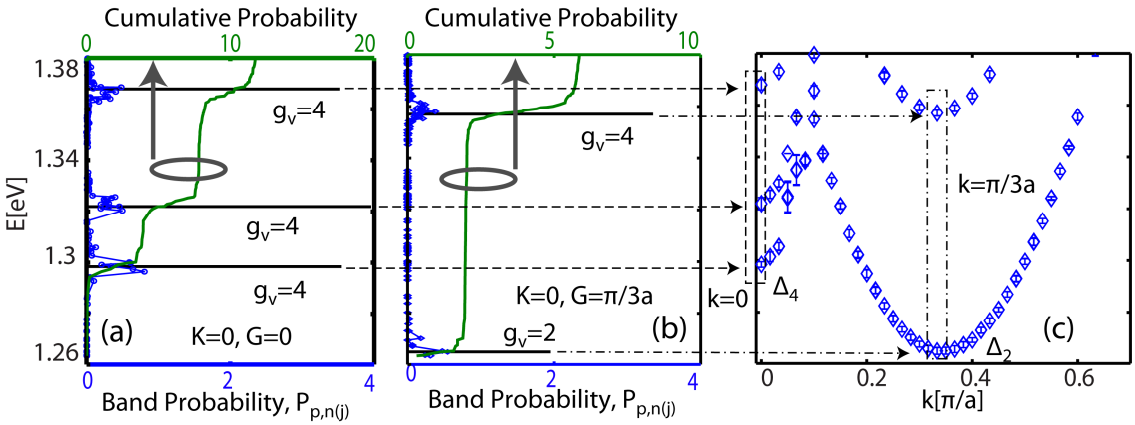

Figure 5.10. Zone unfolding of the bandstructure of 32.5×7.1×7.1nm $Si_{0.8}Ge_{0.2}$ alloy nanowire supercell to obtain bandstructure of 7.1×7.1nm $Si_{0.8}Ge_{0.2}$ alloy nanowire: Probability sums calculated using equation (5.25) for supercell wavevector *K=0* and reciprocal lattice vectors G=0 and *π/3a* are plotted as a function of energy in figures (a) and (b) respectively. The mean energies $\tilde{E}$ and standard deviations *ΔE* are calculated from the cumulative probability distribution. The degeneracies are denoted by g_v. The mean energy values in figures (a) and (b) are translated to *k=0* and *π/3a* respectively in small cell Brillouin zone in Figure (c). The procedure is repeated for all allowed reciprocal lattice vectors $G_{n(j)}$ to obtain small cell bandstructure shown in Figure(c).

result of the projection method for supercell wavevector K=0 and reciprocal lattice vectors *G=0* and *π/3a* respectively as a function of energy. For band detection, cumulative probability distribution (cdf) obtained from probability sums, $P_{p;n(j)}$, is used. It is easier to detect the steps in cdf than the detection of peaks in $P_{p;n(j)}$. If the band at E_p has degeneracy of g_v then the cdf jumps by g_v at E_p. Since the unfolded supercell bandstructure is approximate it has an error bar associated with each energy and k point in the dispersion. The error bars on the bands of Figure 5.10(c) represent the standard deviation error in the band detection algorithm. These energy uncertainties can be used to calculate the scattering time of the state according to the prescription of [92].

5.5. Conclusion

The zone unfolding method for 1D nanostructures is outlined in this chapter. The simple toy model of section 5.2 summarizes the essential physics of unfolding. The zone-unfolding method for realistically sized perfect nanowires is presented in section 5.3. In section 5.4, the generalized zone-unfolding method for imperfect supercells is presented.

In the next chapter the zone-unfolding method is employed to study of the electronic transport in alloy nanowires.

6. APPLICATIONS OF ZONE-UNFOLDING: ELECTRONIC STRUCTURE AND TRANSPORT IN ALLOY NANOWIRES

6.1. Introduction

Semiconductor nanowires are believed to be the potential candidates for devices beyond conventional planar Si-CMOS technology. Nanowires are also attractive for sensing applications due to their high surface-to-volume ratio. Nanowire transistors provide the possibility of 3-D integration to improve the device density on the chip. Nanowire transistors can be scaled down to much smaller sizes than their planar counterparts because of enhanced electrostatic control provided by the 3-D geometry. Recently several groups have demonstrated nanowire field-effect transistors (FETs) fabricated from pure elemental or compound semiconductors like Si [93], Ge [94], and GaAs [95] as well as semiconductor alloys like SiGe [95], and their III-V counterparts. Semiconductor alloy nanowires have local atomic and inhomogeneous strain disorder due to random distribution of atoms in the semiconductor lattice. As the nanowire dimension shrinks to below around 5 nm, the effective mass approximation breaks down even for pure elemental semiconductor nanowires [28] and an atomistic representation of the material is needed for transport calculations. Here, the nearest neighbor $sp^3d^5s^*$ semiempirical tight-binding technique is employed to model the electronic structure and the transport in semiconductor alloy nanowires. The tight-binding technique is readily suited for modeling atomic scale alloy disorders because of its atomistic nature of the Hamiltonian.

The tight-binding electronic structure calculations are typically performed within the virtual crystal approximation (VCA) that averages the tight-binding parameters according to the alloy composition. The VCA method can not reproduce the experimental band-gaps and effective masses even for the bulk alloys like AlGaAs [22]. This is because the

VCA method can not model the effect of random placement of atoms in the alloy. Such effects become more important in the nanoscale regime as they lead to the formation of localized states in the nanostructures. Therefore more rigorous models are required to study the electronic transport in alloy nanowires. This work employs the supercell approach, which allows us to relate features in the transport and bandstructure results in greater detail than previously possible. A unified picture of alloy nanowires emerges, in which the nanodevice (transmission) and nanomaterials (bandstructure) viewpoints complement each other and illuminate the physics of these interesting structures.

AlGaAs and SiGe systems are used to illustrate the influence of random-alloy disorder on the electronic structure and transport. Section 6.2 describes nanowire geometries and methods used to calculate bandstructure and transport characteristics. In section 6.3, the AlGaAs system is presented first because of its attraction from a theoretical point-of-view: nearly identical AlAs and GaAs bond lengths, so that disorder appears as the random cation distribution. Transport properties of SiGe nanowires that include random atom distribution and inhomogeneous strain due to different Si-Si and Ge-Ge bond lengths are described in section 6.4. Finally section 6.5 concludes the chapter.

6.2. Method

This section describes the nanowire geometries and methods used to model the electronic transport. Specifically, the virtual crystal approximation (VCA), local bandstructure, and zone-unfolding method are described. The electronic structure, and the transport calculations are performed on the whole nanowire in free standing configuration and substrate is not taken into account.

6.2.1. Nanowire geometries

The AlGaAs and SiGe alloy nanowires studied here have a square cross section, are oriented along [100] direction and have surfaces along *x*-*y* and *x*-*z* planes. The AlGaAs random alloy nanowire is shown in the atomistic representation in Figure 6.1. The nanowire geometry is specified in terms of conventional Zincblende (cubic) unit cells as

$n_x \times n_y \times n_z$ where n_α is the number of cubes in direction-α. Thus the nanowire dimension in direction α is given by $n_\alpha a$, where a is the lattice constant of the material. The nanowire supercell is periodic in the transport (x-direction) and the wire mantle surfaces are passivated to push the surface states out of the bandgap. The prescription of [96], where dangling bonds oriented along sp^3-hybridization directions are raised in energy (by 30eV) is used for surface passivation. The unit cell or small cell of such a nanowire is a $1 \times n_y \times n_z$ slab consisting of two anion-cation planes each.

The same nanowires geometries are used in both bandstructure and transmission calculations. The only difference in the two calculations is boundary conditions along x-direction. Periodic boundary conditions are applied in supercell bandstructure calculations while open boundary conditions [97] are applied in transport calculations.

6.2.2. The virtual crystal approximation (VCA)

The virtual crystal approximation (VCA) treats the alloy as a pseudo-material. For $Al_xGa_{1-x}As$ alloy, a pseudo-material in which all anions are the same (here, As), as are all cations (here Al_xGa_{1-x} pseudo-atoms). The $Al_xGa_{1-x}As$ VCA parameters are calculated in the usual manner,

$$P_{AlGaAs}(x) = xP_{AlAs} + (1-x)P_{GaAs} \tag{6.1}$$

where P is a tight-binding parameter (either onsite or nearest-neighbor). Thus, the VCA

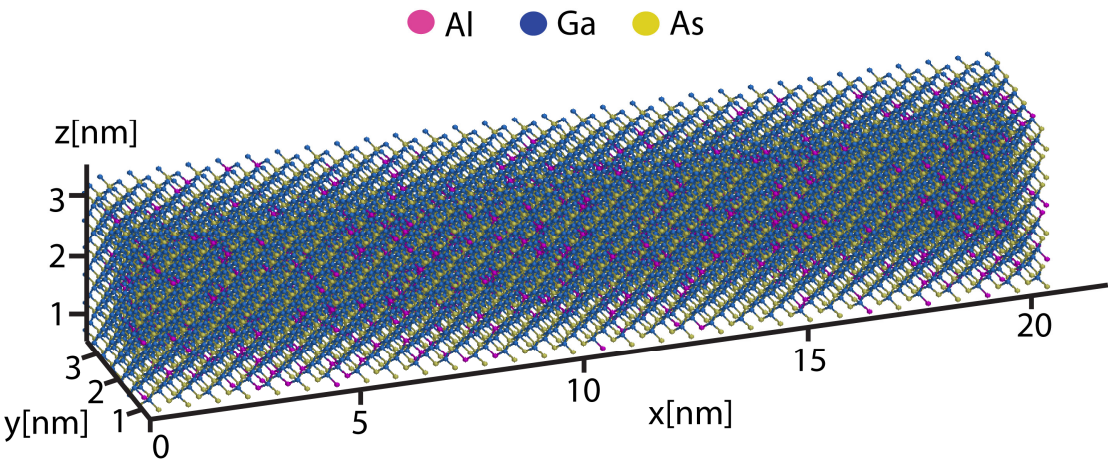

Figure 6.1. 22×3.3×3.3nm $Al_{0.15}Ga_{0.85}As$ random alloy nanowire. Nanowire is 40 fcc unit cells long in x-direction and 6 fcc unit cells long in both y- and z-directions.

does not incorporate randomness at an atomistic level, unlike the supercell calculation which is used here for approximate bandstructure calculation. The VCA parameters for Si_xGe_{1-x} alloy are calculated in the similar manner as

$$P_{SiGe}(x) = xP_{Si} + (1-x)P_{Ge} \tag{6.2}$$

6.2.3. Local bandstructure

Figure 6.2(a) depicts a sliver cut through the center of the 40×4×4 $Si_{0.8}Ge_{0.2}$ nanowire, indicating the atomistically resolved disorder of the wire. This nanowire is made of 40 1×4×4 slabs consisting of two anion-cation planes each. The so called local bandstructure of each slab is calculated assuming that each slab repeats infinitely along the nanowire. The bandstructure of the first slab is shown in Figure 6.2(b). Due to variations in atomic arrangements along the nanowire length one expects to see the different bandstructures for each slab. Figure 6.2(c) shows the conduction band-edge minima along the nanowire length. Fluctuations in the conduction band-edge minima cause reflections which lead to the formation of the localized states and peaks in transmission plots.

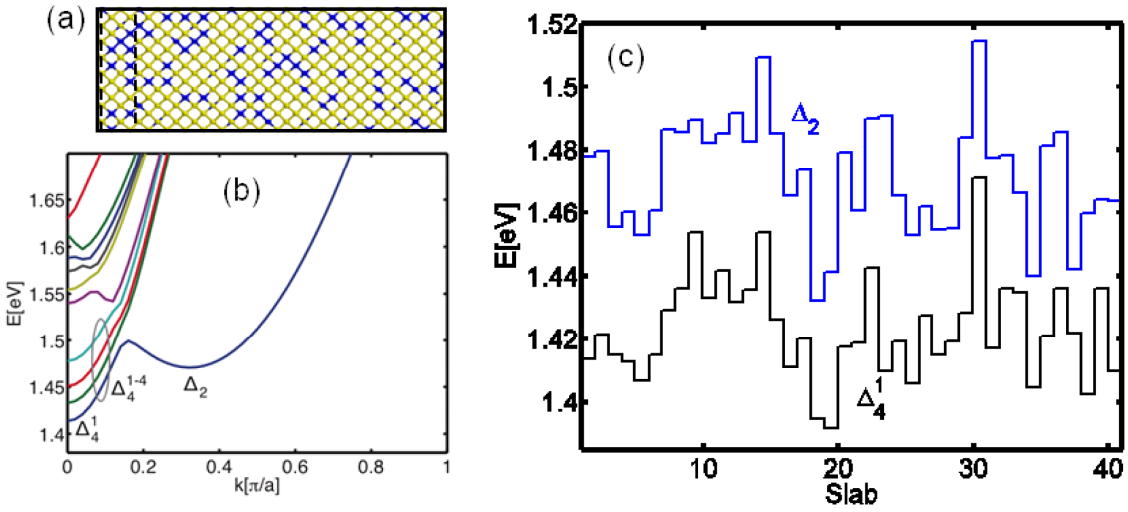

Figure 6.2. (a) Two equivalent representations of a [100] $Si_{0.8}Ge_{0.2}$ 40×4×4 nanowire. Small cell representation (dotted lines) in which a wire is made up of a smallest repeating unit i.e. 4 atomic planes thick slab. Supercell representation (solid lines) in which a repeating unit consists of integer number of small cells. (b) Conduction bandstructure of the first slab. Δ_4 valleys are split into four separate bands due to alloy disorder. Δ_2 valley bands are doubly degenerate. (c) Bandedge minima of lowest energy Δ_4 and Δ_2 valleys along length of the nanowire.

6.2.4. Bandstructure using zone-unfolding

The translational symmetry is broken in semiconductor alloy nanowires due to random placement of atoms and inhomogeneous strain. Figure 6.2(a) shows such disorder in a sliver cut through the center of the 40×4×4 $Si_{0.8}Ge_{0.2}$ nanowire. Thus one runs into the problem of choosing a unit cell for the bandstructure calculation. Due to the loss of the translational symmetry the concept of bandstructure can only be considered in an approximate sense. One way to take alloy disorder into account is to work with longer supercells (Figure 6.2(a)) as basic repeating units. The nanowire bandstructure obtained from the supercell calculation is folded due to which effective masses and energies of the off-zone center valleys can not be determined. The one dimensional version of the zone-unfolding method [42] described in chapter 5 is used to project out the approximate eigenspectrum of the nanowire supercell on the small cell Brillouin-Zone. The Projected probabilities are then used to extract the bandstructure of the alloy nanowire according the probability sum rule for band counting [22]. The small cell bandstructure thus obtained captures the effect of alloy disorder on the electronic structure. This approximate band structure is representative of the overall transport capabilities of the wire and correlates well to NEGF transport simulations. From this approximate dispersions, one can derive critical device parameters such as band gap and effective masses (along the transport direction), which can be used in an approximate top-of-the-barrier-model to predict device performance [98].

6.2.5. Transmission using NEGF formalism

The transmission coefficients through the nanowire are computed by using a hybrid method combining a recursive Non-Equilibrium Greens Function (NEGF) and a wavefunction method [97]. The scattering boundary method is used to calculate the open boundary conditions and the surface Green's function from the bandstructure of a reservoir. In these calculations the semi-infinite source (drain) region is assumed to be identical to the first (last) slab of the nanowire. That is the same atomic disorder as the first (last) slab is assumed to repeat throughout the semi-infinite source (drain) region so

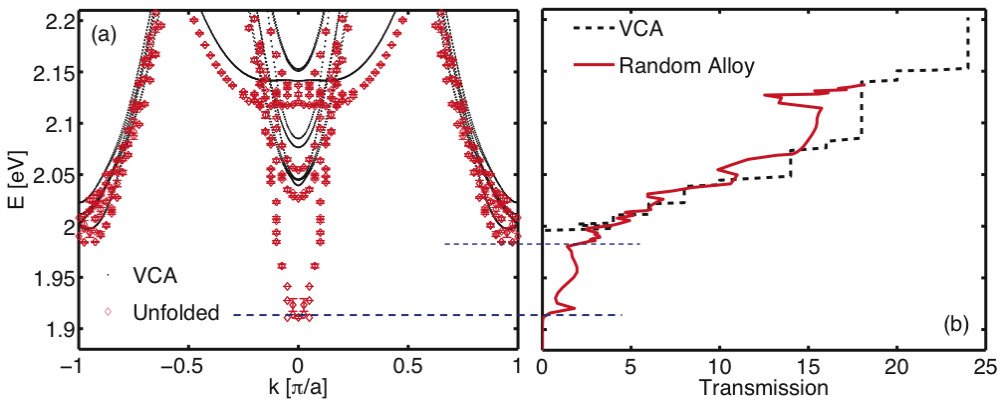

Figure 6.3. (a) Conduction bands of the 40×6×6 $Al_{0.15}Ga_{0.85}As$ nanowire as calculated with the VCA (small solid symbols) and as projected out of random-alloy supercell eigenstates (large, open symbols with error bars). Note in particular that the random alloy calculation gives a significantly lower minimum at $k = 0$. (b) Transmission characteristics. Dotted line: VCA nanowire; this nanowire is effectively a pure nanowire made of a pseudo-material, and shows step-like transmission. Solid line: random-alloy nanowire.

that the bandstructure of the source (drain) region is same as that of the first (last) slab. The tight-binding model in which atom and strain disorders are automatically incorporated by the atomistic nature of the Hamiltonian is used.

6.3. Transport characteristics of AlGaAs random alloy nanowires

AlGaAs nanowires have atom disorder due to random placement of Al atoms in GaAs lattice. Both GaAs and AlAs have almost same lattice constants hence AlGaAs nanowires do not have any inhomogeneous strain disorder. Atom disorder is automatically taken into account in supercell tight-binding calculations because of atomistic nature of the Hamiltonian. Figure 6.3(a) shows bandstructures of an $Al_{0.15}Ga_{0.85}As$ random alloy nanowire calculated using conventional VCA method and unfolding 40×6×6 supercell eigenspectrum. The random alloy calculation results in significantly lower conduction band minima than VCA calculation. Transmission coefficients through the same nanowire supercell are shown in Figure 6.3(b). The VCA transmission shows an ideal

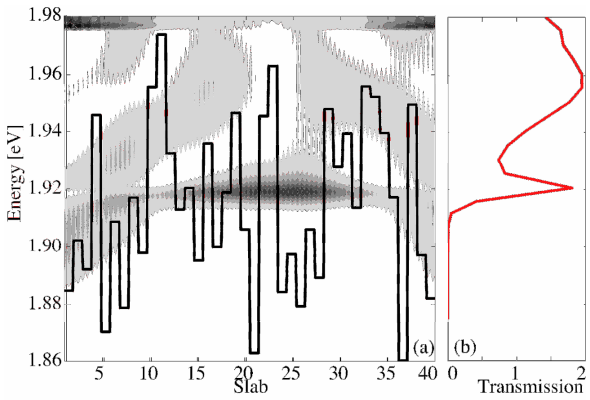

Figure 6.4. (a) Logarithm of the DOS (shaded region) neglecting spin, for the random-alloy nanowire superimposed over the wire conduction-band edge profile (thick line). (b) Transmission coefficient. The injected state is spin-up. The concentration of the DOS corresponds to the resonance peak in the transmission coefficient.

step like behavior, however, the transmission coefficient from random alloy calculations shows a noisy behavior as a consequence of random placement of Al atoms in AlGaAs nanowire. Features in transmission are related to the approximate bandstructure and vice versa. Each band corresponds to two transmission channels for up and down spins. The lowest approximate band at $k=0$ produces a transmission turn on near 1.92 eV. More channels turn on at about 1.97 eV due to approximate bands near the Brillouin zone boundary π/a_0.

Figure 6.4(a) shows the local conduction band minima for each slab (1×6×6 cell) and a density of states (DOS) of the same 40×6×6 AlGaAs nanowire. The transmission spike at about 1.92 eV in Figure 6.4(b) corresponds to a localized density of states seen in Figure 6.4(a). In summary, the approximate bandstructures from random alloy supercell calculations and atomistic NEGF transport calculations are complimentary and mutually supporting. Both methods provide better insight into the disordered nanowire device physics [23].

6.4. Transport characteristics of SiGe nanowires

The free standing square cross section $Si_{0.8}Ge_{0.2}$ alloy nanowires studied here have two types of disorders: random atom disorder due to alloying and inhomogeneous strain

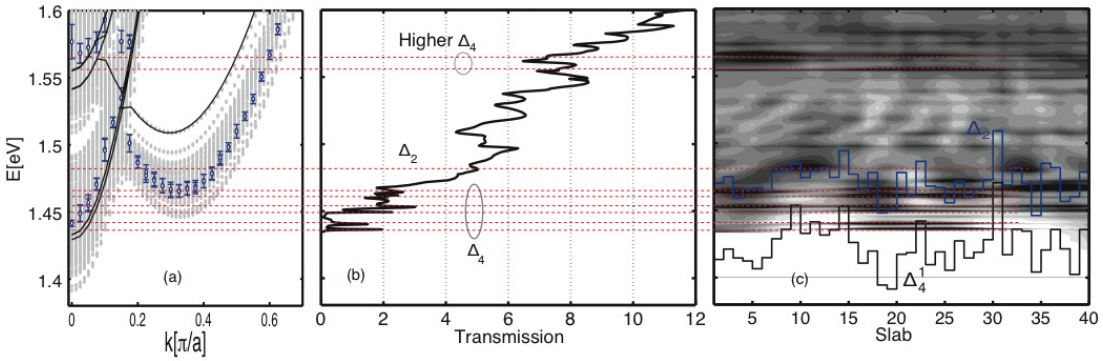

Figure 6.5. (a) Bandstructures of 40×6×6 $Si_{0.8}Ge_{0.2}$ alloy nanowire in local bandstructure (grey), VCA (black) and zone-unfolding (blue) formulations. (b) Transmission coefficient. Steps in transmission are identified as resulting from new bands appearing in projected bandstructure. (c) Local band-edge of lowest energy Δ_4 and Δ_2 valley minima along wire length and density of states (on a log scale). Peaks in the transmission arising from Δ_4 valley localized states have dominant DOS contributions coming from $|py|^2+|pz|^2$ while Δ_2 valley peaks have dominant $|px|^2$ contribution.

disorder due to different Si-Si, Ge-Ge, and Si-Ge bond lengths. Square nanowire cross-sections ranging from 2 to 7nm are simulated to study the effect of disorder. All electronic structure and transport calculations have been done in $sp^3d^5s^*$ tight-binding model. Relaxed wire geometries are calculated in NEMO-3D [39, 40, 77] from Valance Force Field approach [78]. The bulk and strain Si and Ge parameters are taken from [76] and [79].

The smaller wire (Figure 6.5) has the dimensions of 40×6×6 (22.3×3.3×3.3nm) i.e. it is constructed from 40 1×6×6 slabs along [100] crystallographic direction. Electronic structure of this nanowire is calculated with NEMO-3D in 3 different formulations in Figure 6.5(a). The first approach is the so called local bandstructure approach in which the bandstructure of each slab is calculated assuming that this slab repeats infinitely along the nanowire. Due to fluctuations in atomic arrangements along the nanowire length one expects to see the different bandstructures for each slab as shown in Figure 6.5(a). This local bandstructure approach does not deliver a meaningful overall band edge of the wire or a meaningful effective mass. The second approach is the conventional VCA approach which averages the atom potentials according to the atomic composition of the material. This results in a homogeneous wire without any disorder. The bandstructure might then as well be represented by a single slab. In the third approach eigenspectrum of the whole

nanowire supercell is unfolded to obtain the bandstructure of the best translationally symmetric nanowire small cell.

Figure 6.5(b) shows the transmission coefficient through this wire computed using coupled open boundary conditions and atomistic NEGF approach. Transmission coefficient shows the noisy behavior because of random SiGe alloy disorder and inhomogeneous strain disorder in the wire, which cause reflections along the nanowire direction. Steps in the transmission plot can be roughly related to the unfolded bandstructure (Figure 6.5(a)) from supercell calculations. Four separate Δ_4 valley bands appear as a single band with a finite energy spread in the projected bandstructure. These four bands turn on near 1.44 eV which corresponds to the conduction band transmission turn on. Two Δ_2 valley bands turn on near 1.47 eV which leads to a step in the transmission. Four more channels due to higher Δ_4 valley sub-bands turn on near 1.57eV. These transmission features can not be related to the conventional VCA bandstructure shown in Figure 6.5(a).

Peaks in the transmission plot can be related to the resonant transport through localized states in the wire. Local band-edge plots of the lowest Δ_4 and Δ_2 valley minima are shown in Figure 6.5(c). Variation of band-edges along the nanowire length cause reflections which lead to the formation of the localized states and peaks in transmission as seen in the density of states and transmission plots. Note that the presence of the peaks in the transmission can not be predicted from the unfolded bandstructure because the

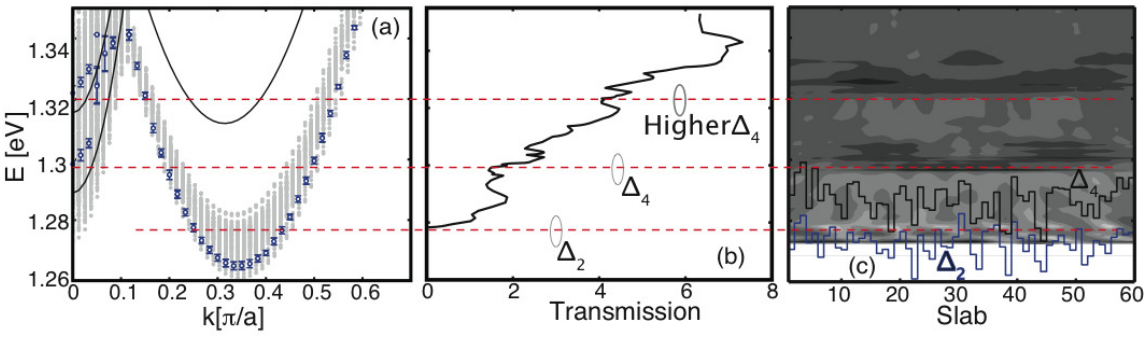

Figure 6.6. (a) Bandstructures of 60×13×13 $Si_{0.8}Ge_{0.2}$ alloy nanowire in local bandstructure (grey), VCA (black) and zone-unfolding (blue) formulations. (b) Transmission characteristics of the wire. (c) Local band-edge of lowest energy Δ_4 and Δ_2 valley minima along wire length and density of states (on a log scale). Transport characteristics show migration towards ideal 1D like transport.

unfolded bandstructure is an eigenspectrum of the best possible translationally symmetric Hamiltonian through the nanowire.

As the nanowire cross section increases random distribution of Ge atoms is expected to have less influence on the transport characteristic of the SiGe nanowires. The local bandstructure variations along the length of the nanowire decrease which reduce the number of localized states. The bandstructure and the transmission characteristics of a 60×13×13 (32.5×7.1×7.1nm) $Si_{0.8}Ge_{0.2}$ wire are shown in Figure 6.6. Error bars on the approximate bandstructure are smaller compared to smaller cross section nanowire of Figure 6.5. The transmission coefficient also shows smoother step like behavior compared to the smaller nanowire. The transmission behavior of the nanowire can not be predicted from the VCA bandstructure because the transmission turns on at 1.28eV before the first VCA band which only starts at 1.29eV.

It is important to point out here that the bandstructure calculation and the transport calculation use different boundary conditions. In the bandstructure calculation the wire is repeated infinitely with disorder, while the transport calculation assumes a homogeneous ordered alloy at the band edges as injectors. Therefore transmission turn on occurs at higher energies than predicted by the unfolded bandstructure because the local conduction band minima near the left reservoir (also the source of injected waves) are higher in energy which prevent propagation waves with energies lower than 1.28eV.

Extracted band parameters such as conduction band minima, effective masses and average band uncertainties are plotted as a function of the nanowire size in Figure 6.7. The approximate bandstructure predicts a smaller Δ_2 band minimum than the VCA

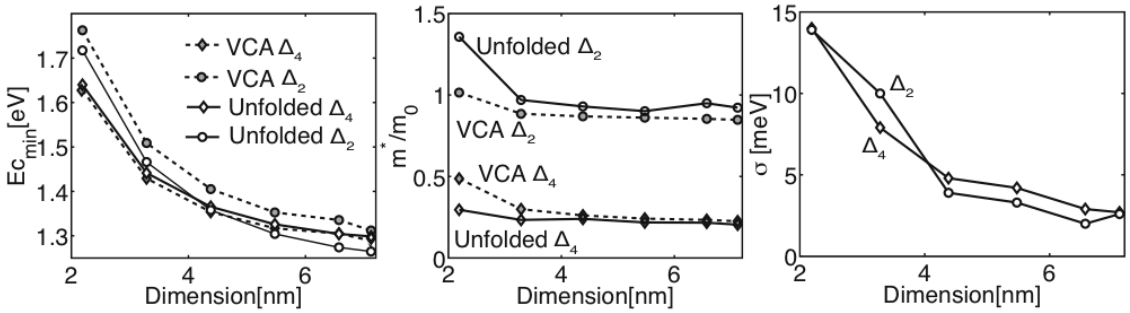

Figure 6.7. (a) Direct (Δ_4 valleys) and indirect (Δ_2 valleys) conduction band minima obtained form VCA and zone-unfolded bandstructures. (b) Effective masses of Δ_4 and Δ_2 valleys. (c) Energy uncertainties of Δ_4 and Δ_2 bands.

similar to AlGaAs nanowires [23] studied in section 3 and AlGaAs bulk [22]. The direct (Δ_4) and indirect (Δ_2) valley-band gaps show an interesting cross-over for wires with sizes larger than 4nm, which will significantly increase the density of states at the conduction band edge and influence device performance; the VCA assumption does not result in such a cross-over. The VCA and unfolded bandstructure effective masses are slightly different (Figure 6.7(b)). Unfolded bandstructure effective masses compare closely to the bulk $Si_{0.8}Ge_{0.2}$ Γ- and X-valley effective masses ($0.92 \cdot m_0$ and $0.19 \cdot m_0$ respectively) for nanowires with larger diameters.

Since the unfolded supercell bandstructure is approximate it has an error bar associated with each energy and wavevector in the dispersion. These energy uncertainties can be used to calculate the scattering time of the state according to the prescription of [92]. As the nanowire cross-section increases the error bars become smaller (Figure 6.7(c)) and the system becomes more bulk like [25].

6.5. Conclusion

The electronic structure and the transport characteristics of random AlGaAs and SiGe alloy nanowires indicate the critical importance of the treatment of atomistic disorder. Typical approaches of a smoothed out material (VCA) or considerations of bandstructure in just individual slices clearly fail to represent the disordered nanowire physics especially for very narrow cross section nanowires. Unfolded bandstructures from zone-unfolding of the supercell eigenspectrum and transmission characteristics combined are found to explain the relevant physics in these disordered systems. As the nanowire cross-section increases the transport behavior migrates from resonance dominated regime to smoother ideal 1D transport regime.

7. MULTISCALE MODELING AND PERFORMANCE ANALYSIS OF ULTRA-SCALED INAS QWFETS

7.1. Introduction

As the Si CMOS technology approaches the end of the ITRS roadmap, the semiconductor industry faces a formidable challenge to continue transistor scaling according to Moore's law [1]. Several industry and academic research groups have recently demonstrated high mobility III-V quantum well field effect transistors (QWFETs) that provide the possibility of achieving high speed operation at lower supply voltage for applications beyond Si CMOS technology [6-11, 99, 100]. In particular, InGaAs and InAs channel QWFETs show a great promise as their superior performance compared to Si MOSFETs and heterogenous integration on Si substrate have already been demonstrated [7, 8, 11].

Classical approximations such as the drift-diffusion model cannot capture the quantization of the energy levels resulting from the strong confinemnt of the electrons in a quantum well and the tunneling currents in nano-scale devices. To address these limitations quantum mechanical approaches based on the effective mass approximation [101] and on the tight-binding method [62] have already been proposed. While both approaches agree well with experimental data above threshold, they are not able to reproduce the OFF-current region where gate leakage dominates. The absence of a real dielectric layer between the channel and the gate contact, as in MOSFETs, makes the III-V QWFETs very sensitive to the gate leakage currents. Here, we use a two-dimensional (2-D), real-space Schrödinger-Poisson solver based on the effective mass approximation and accounting for gate leakage currents [102] to investigate the logic performances of InAs QWFETs. The transport and confinement effective masses are extracted from an atomistic $sp^3d^5s^*$ tight-binding calculation that includes strain and non-parabolic effects.

Strain is computed in the system with a Valence-Force-Field method [39, 40, 77]. The tight-binding parameters [80] have been validated against experimentally very different structures such as strained InAs quantum dots [90] and QWFETs [62].

This chapter is organized as follows: In section 7.2, we describe the 2-D real-space Schrödinger-Poisson solver based on the effective mass approximation, the tight-binding technique used to calculate the channel effective masses, and the Newton scheme used to calibrate the simulator to the experimental data. The approach is applied to calibrate the simulator to the experimental data on InAs QWFETs with gate lengths ranging from 30 to 50 nm, where a good quantitative match is obtained in both the OFF- and ON-state current regions. In section 7.3, the calibrated simulator is subsequently used to investigate the logic performance optimizations for the 20 nm InAs QWFET. It is shown that the best performance is achieved in thin InAs channel devices by reducing the insulator thickness and maintaining a flat gate geometry to improve the gate control and increasing the gate metal work function to suppress the gate leakage current. Finally, we present the conclusion and outlook of our work in section 7.4.

7.2. Approach

A three step multiscale modeling approach is adopted. First the strain is computed in an atomistic valence-force-field, secondly an atomistic tight-binding method is used to compute the electron dispersion in the quantum well channel and the corresponding electron effective masses. In a third step these effective masses are used in an effective mass based quantum transport simulator to obtain current-voltage characteristics of a QWFET which includes quantization effects and three-terminal effects such as gate tunneling.

7.2.1. Device structure

The InAs channel QWFET [8] is schematically shown in Figure 7.1. The channel region is composed of a 10 nm composite $In_{0.53}Ga_{0.47}As/InAs/In_{0.53}Ga_{0.47}As$ (2/3/5 nm, from top to bottom) multi-quantum-well grown on 500 nm thick $In_{0.52}Al_{0.48}As$ layer

lattice matched to InP substrate. The $In_{0.52}Al_{0.48}As$ layer between the multi-quantum-well channel and the gate contact acts as an insulator. The channel conduction electrons are supplied by the Si δ-doped layer placed just below the gate. The source/drain contacts are located on the top of the device almost 1 μm away from the gate contact. To reduce the computational burden we restrict the intrinsic device simulation domain to the gate contact and an extension L_{side} of 50 nm on each side. The extrinsic source/drain contacts are modeled via two series resistances R_S and R_D [103]. Due to this idealized contact assumption the phenomena such as "source starvation" that arise due to non-equilibrium contacts are not included in our simulations [104]. Two gate contact geometries, curved and flat which are a result of different gate-stack fabrication processes are considered. Edges of the curved gate contact are quarter circles with the radius of curvature equal to t_{InAlAs}-t_{ins}, where, t_{InAlAs} is the total thickness of InAlAs layer and and t_{ins} is the thickness of InAlAs layer between the gate contact and multi-quantum-well channel. Such curved gate contact geometries are a result of isotropic chemical etching techniques [8, 9].

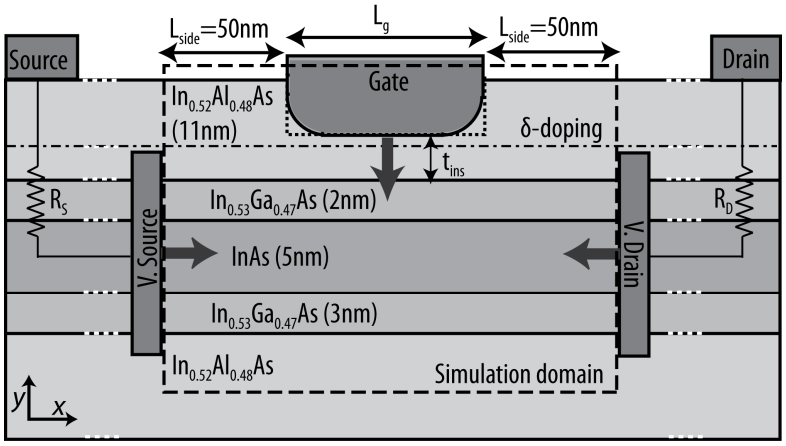

Figure 7.1. Schematic view of the InAs QWFET. The channel region is composed of a 10 nm InGaAs/InAs/InGaAs multi-quantum-well grown on a thick $In_{0.52}Al_{0.48}As$ layer lattice-matched to the InP substrate (not shown). The $In_{0.52}Al_{0.48}As$ layer between the gate contact and the multi-quantum-well-channel acts as an insulator. The black dash dotted line represents the Si δ-doped layer of concentration $N_D = 3\times10^{12}$ cm^{-2}. The dashed black rectangle encloses the quantum transport simulation domain, which is restricted to the gate contact region and the extension of L_{side} = 50 nm on source/drain sides. White arrows depict the direction of electron injection from contacts into the simulation domain. The source/drain extensions beyond virtual contacts are modeled by two series resistances R_S and R_D, respectively. Two gate contact geometries curved (black) and flat (dotted black) are investigated.

7.2.2. Channel effective masses

An accurate determination of the channel effective mass is crucial in devices subject to strain and strong material non-parabolicities. The effective mass determines channel properties such as injection velocity, source to drain tunneling, quantum capacitance, and density of states [98]. Effective masses are extracted from the $sp^3d^5s^*$ tight-binding bandstructure of the multi-quantum-well channel calculated using the general purpose electronic structure simulator NEMO-3D [39, 40, 77].

The compositions of InGaAs and InAlAs layers are chosen such that they are lattice matched to the InP substrate which has a lattice constant a_{InP} = 0.58688 nm. The InP lattice constant is smaller than the lattice constant of InAs (a_{InAs} = 0.60583 nm) which causes 3.2 % lattice mismatch and a biaxial compressive strain at the InAs-InGaAs

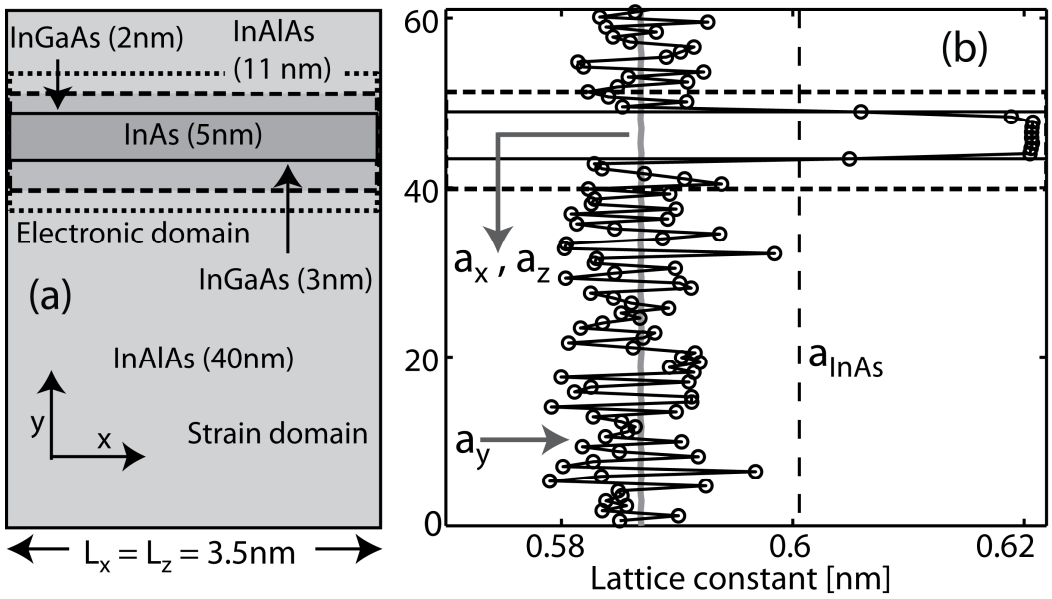

Figure 7.2. Strain modeling of the channel. InAs QWFET is fabricated in the heterostructure shown in (a). All layers except InAs are lattice matched to the InP substrate (not shown). The in-plane dimensions of the simulation domain are $L_x = L_z$ = 3.5 nm, which are sufficiently large to model the random placement of cations in the InGaAs and InAlAs layers. The relaxed atom positions are calculated from the valance-force-field method. As shown in Figure (b), the in-plane lattice constant (a_x, a_z) is smaller than the unstrained lattice constant of InAs. Due to the resulting biaxial compressive strain the lattice constant along growth direction (a_y) expands in the InAs quantum well region. The electronic structure calculation domain (enclosed in dotted black rectangle) has the same lateral dimensions as the strain relaxation domain and it includes 2 nm InAlAs around the InGaAs/InAs/InGaAs multi-quantum-well to capture the effect of wavefunction leakage into the InAlAs layers.

interface. The valance-force-field (VFF) method with Keating potential is used to compute the relaxed atom positions in the strained heterostructure [39, 77, 78]. The in-plane dimensions of the strain relaxation domain are L_x = L_z = 3.5 nm, which are sufficiently large to model the random placement of cations in InGaAs and InAlAs [40]. We remind the reader here that InGaAs shows a bimodal In-As and Ga-As bond distribution [105] which is quite important in quantum wells [90]. 11 nm thick InAlAs layer above and 40 nm thick InAlAs layer below the multi-quantum-well channel are included and InP lattice constant is imposed as the boundary condition at the bottom (Figure 7.2(a)).

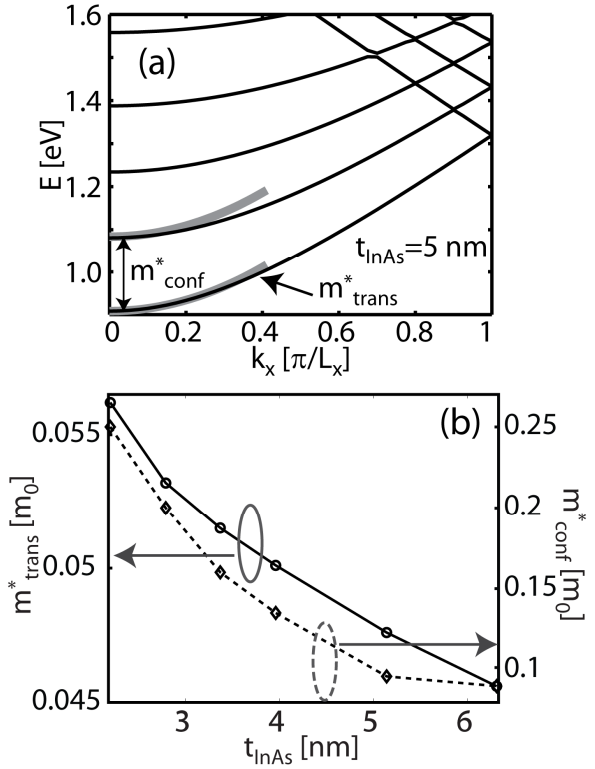

Figure 7.3. (a) Bandstructure (at k_z=0) obtained from the $sp^3d^5s^*$ tight-binding calculation on the electronic structure domain shown in Figure 7.2(a). The effect of random placement of cations in InGaAs and InAlAs layers, strain, and quantization due to band discontinuities at various heterostructure interfaces are included. The transport effective mass is extracted by fitting a parabola to the lowest conduction sub-band. The confinement effective mass is fitted to replicate the energy difference between first two sub-bands in the tight-binding bandstructure. (b) The transport and confinement effective masses extracted from the tight-binding bandstructure are significantly higher than the bulk InAs effective mass, $m^*_{InAs} = 0.023 \cdot m_0$. The effective mass is higher in thinner InAs quantum wells due to stronger confinement and a high degree of InAs non-parabolicity.

The lattice mismatch at the InAs/InGaAs interface causes biaxial compressive strain in InAs quantum well. The in-plane lattice constant along the center of InAs quantum well is labeled as (a_x,a_z) in Figure 7.2(b). Its value is same as the lattice constant of InP substrate. This causes in-plane biaxial compressive strain of $\varepsilon_{//} = a_x/a_{InAs} - 1 = -0.031$ in InAs quantum well. The lattice constant along the orthogonal direction (a_y) extends to 0.6207 nm (from bulk value of a_{InAs} = 0.60583 nm) in the InAs quantum well region which corresponds to orthogonal strain of $\varepsilon_{\perp} = 0.025$. Fluctuations in the lattice constant in InGaAs and InAlAs alloy regions arise from the local bond length variation due to random placement of cations.

The electronic structure calculation domain (dotted black rectangle in Figure 7.2(a)) is smaller than the strain relaxation domain. Only 2 nm thick InAlAs layers on the top and the bottom of the multi-quantum-well channel are included because the wavefunction leakage beyond this domain is negligible. The $sp^3d^5s^*$ tight-binding bandstructure of the multi-quantum-well with 5 nm thick InAs layer is shown in Figure 7.3(a). The effect of strain and quantization due to band discontinuities at InAs/InGaAs and InGaAs/InAlAs heterostructure interfaces are automatically included due to atomistic nature of the tight-binding Hamiltonian. The tight-binding parameters [80] were designed as transferable bulk parameters and have previously been benchmarked against complex experimental devices such as InAs/InGaAs/InAlAs quantum dots [90] and InAs QWFETs [62]. The transport effective mass is extracted by fitting a parabola to the lowest conduction sub-band. The confinement effective mass is fitted to replicate the energy difference between first two sub-bands in the tight-binding bandstructure. The transport and confinement effective masses in the InAs quantum well are raised from their bulk value (m^*_{InAs} = $0.023 \cdot m_0$) due to quantum confinement provided by InGaAs and InAlAs buffer layers. The effective masses become heavier as the quantum well thickness is reduced indicating a strong non-parabolic dispersion in InAs.

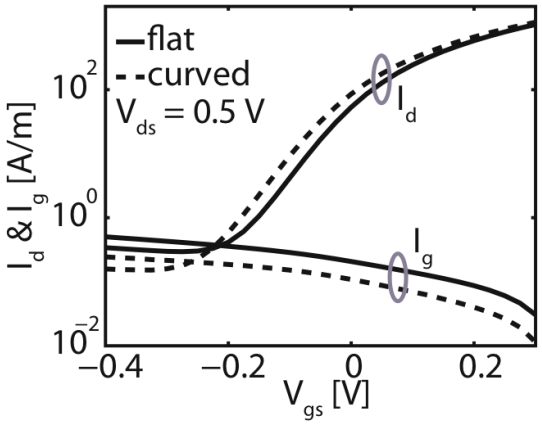

Figure 7.4. Intrinsic I_d-V_{gs} and I_g-V_{gs} characteristics of the L_g = 51nm InAs QWFETs with flat (dashed lines) and curved (solid lines) gate-insulator interface. Both devices perform similarly in the ON-state regime, however, the flat gate device exhibits higher gate-leakage current. Flat gate device exhibits superior gate control compared to the curved gate device. The flat gate device has subthreshold slope and *DIBL* of *SS* = 83.5 mV/dec, *DIBL* = 85.6 mV/dec while the curved gate device shows *SS* = 89.7 mV/dec, *DIBL* = 96.2 mV/dec.

7.2.3. 2-D Schrödinger-Poisson solver

Real-space Schrödinger and Poisson equations are solved self-consistently using the effective mass approximation and a 2-D finite-difference grid [102]. The grid is uniform and spacing along x and y-directions are Δx = 0.25 nm and Δy = 0.2 nm respectively. The device and the gate contact are treated as a single entity on a quantum mechanical level. Neumann boundary conditions are applied everywhere to the Poisson domain except at the gate contact where Dirichlet boundary conditions are applied. In the ballistic transport model used here, electrons are injected into the device at different wave-vector and energy values and the resulting contributions are summed up to give carrier and current densities.

7.2.4. Influence of the gate shape

The gate contact geometry plays an important role in determining the gate leakage current. The shape of the gate contact varies with the fabrication technique used to thin down the gate insulator before deposition of the metal gate stack. Anisotropic etching and

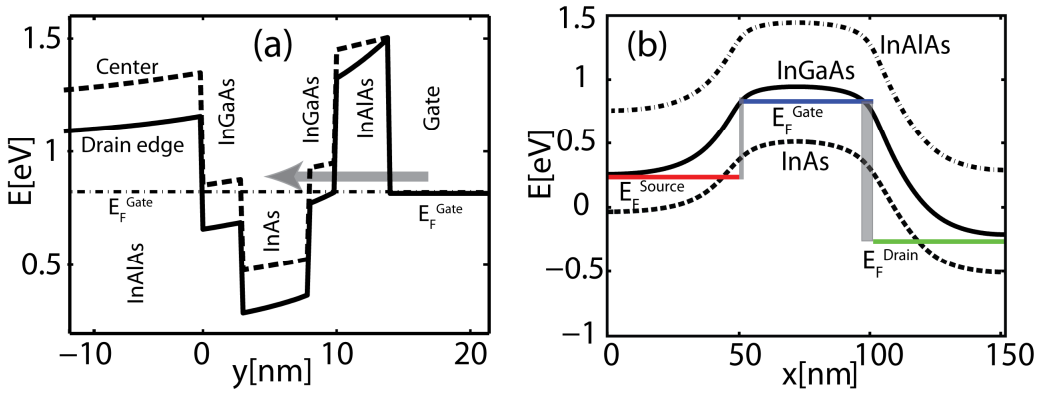

Figure 7.5. Gate leakage mechanism: (a) Band diagram along the vertical lines through the center and the drain side edge of the gate contact of a QWFET with L_g = 51 nm and a flat gate contact. As identified by the gray arrow, the gate leakage current arises due to the electron tunneling through the InAlAs insulator and InGaAs barrier layers between the gate contact and the InAs channel. Electrons at the edges of the gate contact encounter only the InAlAs tunnel barrier while the electrons at the center encounter an additional InGaAs tunnel barrier. Due to this gate-leakage current is concentrated at the edges of the gate contact. The high gate leakage paths are identified by the gray shaded rectangles in the band diagram along the channel direction in Figure (b). Bias conditions are V_{gs} = -0.4 V and V_{ds} = 0.5 V.

metal gate sinking ideally leads to a flat gate contact while isotropic etching leads to a curved gate contact (Figure 7.1). Flat or curved gate contact geometries act differently on the current magnitude as illustrated in Figure 7.4. Both devices perform similarly in the high V_{gs} regime, however, their drain currents differ significantly at low V_{gs}. The flat gate leads to lower subthreshold slope SS = 83.5 mV/dec compared to the curved gate device with SS = 89.7 mV/dec but has high gate leakage by a factor of 2.

The difference between drain currents at low V_{gs} is a result of gate leakage (I_g) suppression in the curved gate contact device. The gate leakage current is concentrated at the edges due to lower tunneling barriers at the edges compared to the central region of the gate contact [63]. The edge-leakage mechanism of current crowding at the edges of the gate contact is explained in Figure 7.5. The current distribution in the gate leakage regime for both flat and curved gate contact devices is shown Figure 7.6. The curved gate device has a larger insulator thickness at the edges of the gate contact which leads to suppression of gate leakage current. Thus, an accurate description of the gate contact is crucial to reproduce the experimental I_d-V_{gs}, especially in the subthreshold regime. A

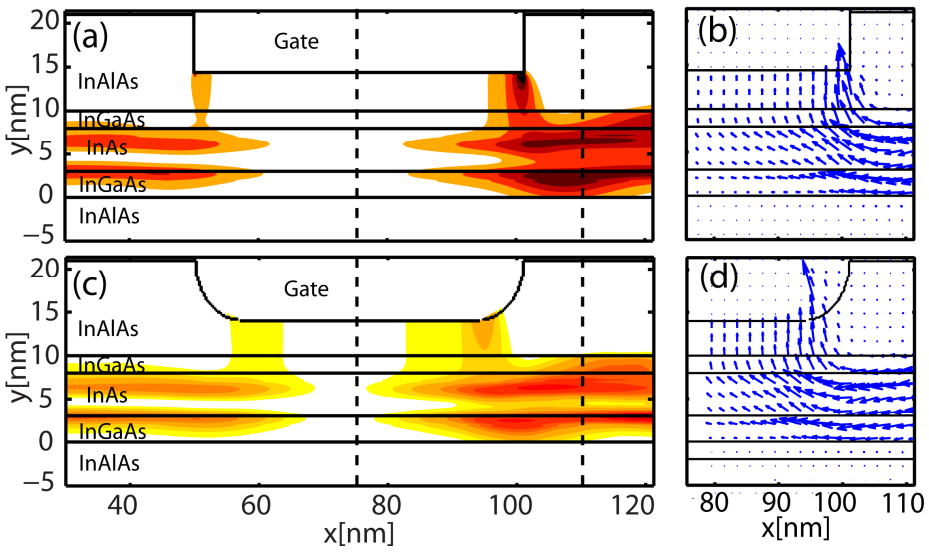

Figure 7.6. Influence of the gate geometry on the OFF-state current distribution - comparison of flat and curved gates. Same bias conditions (V_{gs} = -0.4 V and V_{ds} = 0.5 V) are used in both simulations. The same color scheme is used in both contour plots and the magnitude of the current decreases from dark to light colors. OFF-state current is dominated by the gate leakage. The electron tunneling from the gate contact into the channel gives rise to the leakage. The gate current is concentrated at the edges of the gate contact due to smaller tunneling barriers at the edges [63]. As seen in Figures (a) and (c), the leakage in the curved gate device is suppressed because of the thicker insulator layer at the edges compared to the flat gate device. Figures (b) and (d) show vector plots of current flow in the flat and the curved gate contact device respectively.

curved gate geometry which resembles to the experimental devices is used in the benchmarking procedure.

7.2.5. Benchmarking with experimental data

The device dimensions and material parameter values reported in the experiments [8] are only accurate within certain measurement errors. To calibrate the simulator to the experimental data [8] we use a Levenberg-Marquardt curve fitting scheme [106] to minimize the squared difference between simulated and experimentally measured drain current. The gate leakage and the subthreshold (low V_{gs}) bias regime is chosen for optimization because transport in this regime is very sensitive to the device dimensions and material parameters as compared to the high gate bias regime (Figure 7.7). The drain current is parameterized as $I_d(L_g, t_{ins}, m_{ins}, m_{buf}, \Phi_M)$. The notations used here are: L_g – gate

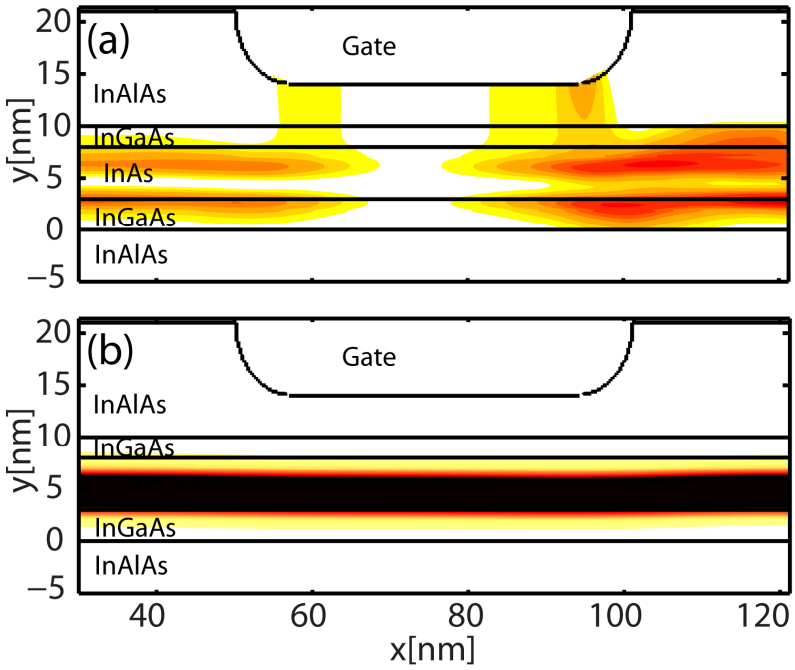

Figure 7.7. Current distribution in L_g = 51 nm QWFET at (a) low gate bias (V_{gs} = -0.4 V and V_{ds} = 0.5 V) and (b) high gate bias (V_{gs} = 0.3 V and V_{ds} = 0.5 V) regimes.

length, t_{ins} – thickness of InAlAs insulator layer between the multi-quantum-well channel and the gate contact, m_{ins},m_{buf} – effective masses of InAlAs and InGaAs respectively, and Φ_M – gate metal work function. The thickness of the InAs channel and the InGaAs buffer are not included in the fitting procedure because these are determined by the Molecular beam epitaxy (MBE) growth which is a precise atomic layer deposition technique. L_g and t_{ins}, however, are determined by lithography and wet chemical etching techniques which are prone to process variability. Likewise Φ_M may vary slightly due to alloy compositions and subsequent high temperature fabrication processes after gate stack deposition. Effective masses of InGaAs and InAlAs (m_{buf} and m_{ins} respectively) are included in the fitting procedure as they affect the electron tunneling probability from gate into the InAs channel which determines the gate leakage current. The electron affinity of InAs channel (χ = 4.9 eV [107]) and the conduction band offsets at the heterostructure interfaces ($\Delta E_{C,InGaAs/InAs}$ = 0.4 eV and $\Delta E_{C,InGaAs/InAlAs}$ = 0.5 eV [107, 108]) are not included in the variable parameter set because drain current is less sensitive to them and their values are well known from experimental measurements.

The fitting procedure is summarized in the flow chart in Figure 7.8. In each iteration, the subthreshold I_d-V_{gs} of the reference device is computed and compared with the experimental data. If the deviation from the experimental data is larger than the tolerance, then the next guess to the parameter vector is computed by using the Levenberg-Marquardt scheme [106]. A closed form expression for the drain current in terms of the fitting parameters does not exist, hence, partial derivatives in the Jacobian matrix which are required by the iteration scheme are computed numerically. To compute a partial

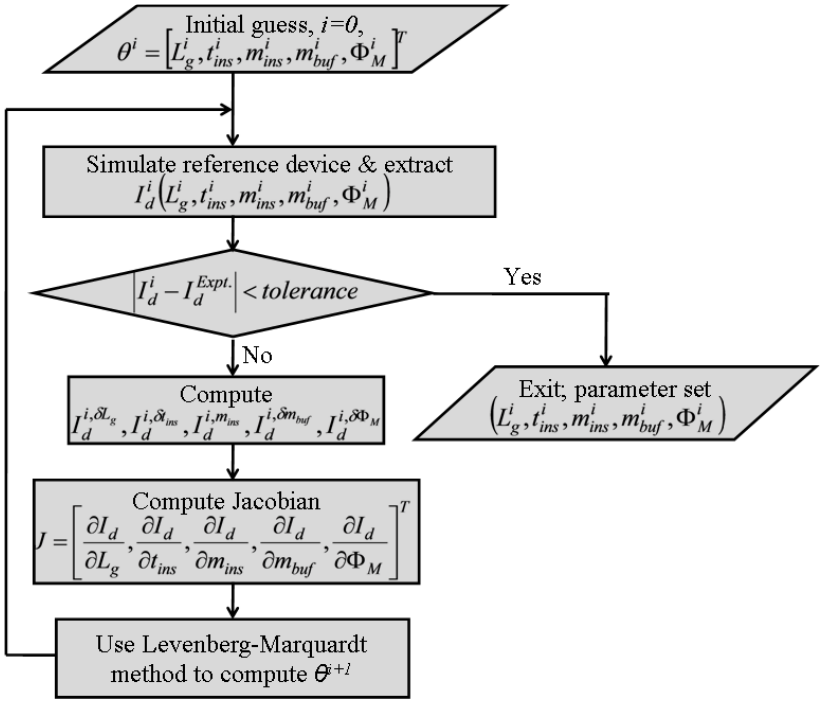

Figure 7.8. The flow chart of the parameter fitting procedure. To match the experimental I_d-V_{gs} data, the drain current is parameterized as $I_d\left(L_g, t_{ins}, m_{ins}, m_{buf}, \Phi_M\right)$. Here, L_g – gate length, t_{ins} – thickness of InAlAs between multi-quantum-well channel and the metal gate, m_{ins} – effective mass of InAlAs, m_{buf} – effective mass of InGaAs buffer, and Φ_M – metal work function. The simulated drain current in the leakage and the subthreshold region is fitted to the experimental data using Levenberg-Marquardt curve fitting scheme. $I_d^{i,\delta L_g}$ is the drain current of a device with the same parameters as the reference device except the gate length is changed to $L_g^i + \delta L_g$. The numerical derivative of the drain curren with respect to L_g is given by $\frac{\partial I_d}{\partial L_g} = \frac{I_d^{i,\delta L_g} - I_d^i}{\delta L_g}$. A similar procedure is used to compute all partial derivates in the Jacobian matrix.

derivative with respect to a parameter, I_d-V_{gs} characteristics of a new device are simulated by slightly changing this parameter and keeping all other parameters same as the reference device. For example, the partial derivative with respect to the gate length is $\frac{\partial I_d}{\partial L_g} = \frac{I_d^{i,\delta L_g} - I_d^i}{\delta L_g}$ where, i – iteration count, I_d^i – drain current of the reference device parameterized as $I_d^i\left(L_g^i, t_{ins}^i, m_{ins}^i, m_{buf}^i, \Phi_M^i\right)$, δL_g – change in the gate length of a new device from the reference device, and $I_d^{i,\delta L_g}$ – drain current of a new device parameterized as $I_d^{i,\delta L_g}\left(L_g^i + \delta L_g, t_{ins}^i, m_{ins}^i, m_{buf}^i, \Phi_M^i\right)$. The same procedure is repeated for calculating partial derivatives with respect to t_{ins}, m_{ins}, m_{buf}, and Φ_M. The parameter shifts used to compute the numerical derivatives are (δL_g, δt_{ins}, δm_{ins}, δm_{buf}, $\delta \Phi_M$) = (0.5 nm, 0.2 nm, 0.005·m_0, 0.005·m_0, 0.05 eV), where m_0 is the free electron mass. Each I_d-V_{gs} calculation requires typically 4 hours on 40 cores on a 2.5 GHz quad-core AMD 2380 processor.

The results of the parameter fitting for devices with gate lengths ranging from 30 to 50 nm are summarized in the table of Figure 7.9. Converged parameter values are close to the experimentally reported values which are used as an initial guess [8]. The effective masses are not allowed to vary among different devices because all devices are fabricated on similar heterostructure stacks. Converged effective mass values are within reported ranges in the literature which are 0.038·m_0 – 0.044·m_0 for $In_{0.53}Ga_{0.47}As$ and 0.070·m_0 – 0.083·m_0 for $In_{0.52}Al_{0.48}As$ [107, 108]. Converged values of gate length and insulator thickness are within the expected process variability of the wet chemical etching step used to thin down the InAlAs insulator prior to gate metal deposition [8]. InAlAs

Parameter	Initial	Final parameter set		
		30 nm	40 nm	50 nm
L_g [nm]	30, 40, 50	34.0	42	51.25
t_{ins} [nm]	4	3.6	3.8	4.0
m^*_{ins} [m_0]	0.075	0.0783	0.0783	0.0783
m^*_{buf} [m_0]	0.041	0.0430	0.0430	0.0430
Φ_M [eV]	4.7	4.6597	4.693	4.6779

Figure 7.9. Device dimensions and material parameters obtained from the fitting procedure. Typically convergence is obtained in less than 15 iterations. Here, m_0 is free electron effective mass.

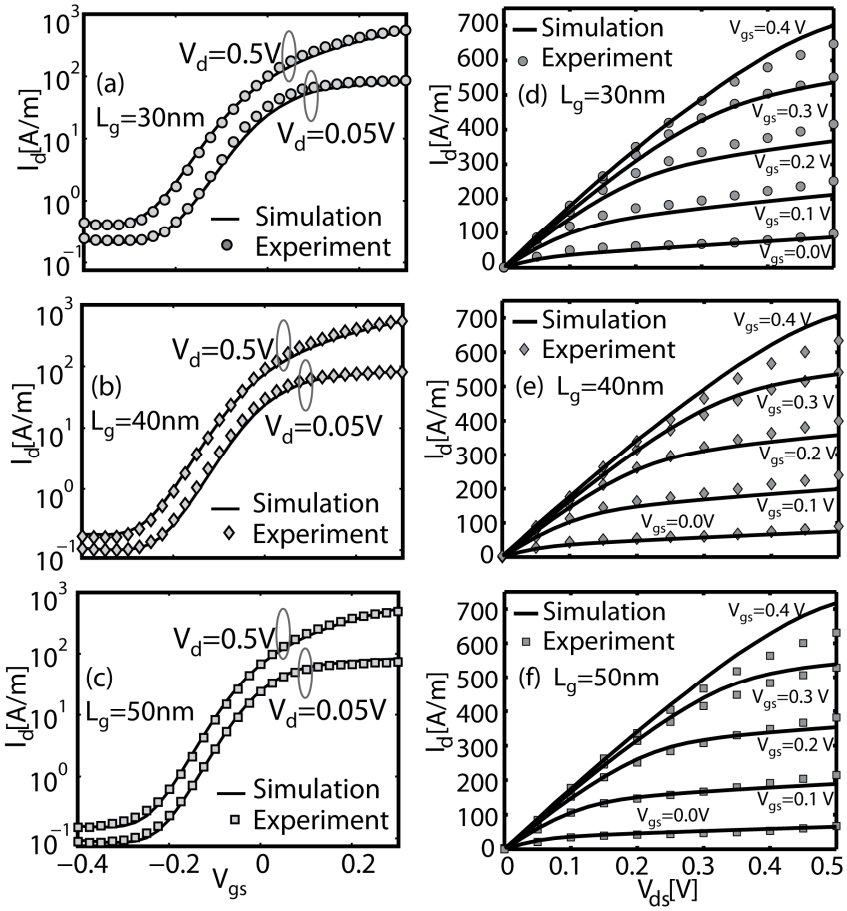

Figure 7.10. Comparison between the experimental [8] and simulated I_d-V_{gs} characteristics of (a) 30 nm, (b) 40 nm, and (c) 50 nm and I_d-V_{ds} characteristics (d) 30 nm, (e) 40 nm, and (f) 50 nm gate length InAs QWFETs. Parameters of Figure 7.9 are used in simulations. Simulated devices have curved gate contacts, which resemble closely to the gate contact geometries resulting from the wet chemical etching step in the gate stack fabrication process. The experimentally measured values of source/drain series resistance R_S = 0.21 Ω·m and R_D = 0.23 Ω·m are used in simulations. Deviation between simulated and experimental I_d-V_{gs} is larger for high bias voltages where scattering may play an important role.

insulator thickness and metal work function values automatically adjust to match different magnitude of gate leakage current in each device (Figure 7.9). Slightly different values of the gate metal work function are justified because the gate metal forms complex compounds at the semiconductor interface and the extent of this reaction is slightly different for each device.

L_g [nm]		S [mV/dec]	$DIBL$ [mV/V]	I_{ON}/I_{OFF}	v_{inj} [cm/s]
30	Expt.	106.9	168.9	0.47×10^3	
	Sim.	105.2	144.7	0.61×10^3	$3.0035\text{x}10^7$
40	Expt.	90.9	126.0	1.38×10^3	
	Sim.	89.4	99.3	1.86×10^3	$3.1127\text{x}10^7$
50	Expt.	85.1	97.2	1.80×10^3	
	Sim.	89.2	90.8	1.85×10^3	$3.1771\text{x}10^7$

Figure 7.11. Device performance parameters from simulated and experimental [8] devices. The threshold voltage (V_T) is defined as the V_{gs} that yields I_d = 1 μA/μm.

The experimental transfer characteristics I_d-V_{gs} and output characteristics I_d-V_{ds} are compared to simulation results in Figure 7.10. The parameters given in Figure 7.9 are used in the simulations. The source and drain series resistance values are R_S = 0.21 Ω·m and R_D = 0.23 Ω·m respectively as determined in the experiment [8]. The transport and confinement effective masses of InAs channel are $m^*_{conf} = 0.049 \cdot m_0$ and $m^*_{trans} = 0.096 \cdot m_0$ respectively (Figure 7.3). A very good quantitative agreement is enabled by consideration of a curved gate contact geometry, accurate estimation of channel effective masses, and parameter adjustments listed in Figure 7.9. Deviations in I_d-V_{ds} are larger at higher biases because scattering may play an important role in these biasing conditions. The performance parameters calculated from experimental and simulated I_d-V_{gs} in Figure 7.10 show a reasonable agreement as shown in the table of Figure 7.11.

7.3. Ultra-Scaled QWFET Design Exploration

After benchmarking our simulator to the "large" gate length experimental devices we explore the performance of a 20 nm gate length device. Since a flat gate contact provides superior gate control of the channel compared to a curved contact (Figure 7.4), we use this geometry for a 20 nm gate length device and propose that it should be used in ultra-scaled devices. A flat gate contact can be "easily" realised by replacing isotropic wet chemical etching step used to thin down the InAlAs insulator layer by anisotropic etching or by a metal gate sinking technique.

Here, we explore the effects of InAlAs insulator thickness, gate metal work function, and InAs channel thickness on the transistor logic performance. The effective masses of

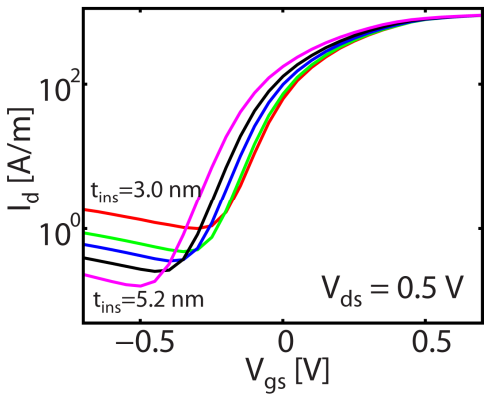

Figure 7.12. Transfer characteristics of a 20 nm InAs QWFET for InAlAs insulator thicknesses of 3.0, 3.4, 4.0, 4.4 and 5.0 nm. The gate leakage increases as the insulator thickness is reduced. Device dimensions are same as in Figure 7.1 with L_g = 20 nm. Metal work function is Φ_M = 4.7 eV.

the InAs channel, $In_{0.53}Ga_{0.47}As$ and $In_{0.52}Al_{0.48}As$ are taken from Figure 7.3 and Figure 7.9 respectively. The conduction band offsets used in the simulations are $\Delta E_{C,InGaAs/InAs}$ = 0.4 eV and $\Delta E_{C,InGaAs/InAlAs}$ = 0.5 eV and the electron affinity of InAs channel is χ = 4.9 eV [107, 108]. The source and the drain resistance values of R_S = 0.21 Ω·m and R_D = 0.23 Ω·m which are same as those for "large" gate length devices are used. In the device performance analysis, the threshold voltage (V_T) is determined from linear extrapolation of I_d-V_{gs} at the peak transconductance to zero I_d (maximum-g_m method) [109]. The ON-state is defined to be $V_g = V_T + 2V_{DD}/3$, $V_d = V_{DD}$, V_s = 0 while the OFF-state is defined to be $V_g = V_T - V_{DD}/3$, $V_d = V_{DD}$, V_s = 0, where V_{DD} = 0.5 V[3]. The capacitances are defined as (i) the gate capacitance: $C_g = dQ_s/dV_{gs}$, (ii) the insulator capacitance: $C_{ins} = \varepsilon_{ins}/t_{ins}$, and (iii) the inversion layer capacitance: $C_{inv} = dQ_s/d\psi_s$ [110]. Here, Q_s is the aerial charge density in the InGaAs/InAs/InGaAs composite channel, ε_{ins} is the dielectric constant of InAlAs insulator, and ψ_s is the potential at the interface between the InAlAs insulator and the InGaAs/InAs/InGaAs composite channel.

7.3.1. Insulator Thickness

The effect of InAlAs insulator thickness (t_{ins}) is depicted in Figure 7.12 and Figure 7.13. Performance metrics *SS*, *DIBL*, and I_{ON}/I_{OFF} ratio of a device with the gate metal

work function Φ_M = 4.7 eV improve as the insulator thickness is scaled until 3.4 nm because of stronger gate control of the channel in thinner insulator devices. However, due to excessive gate leakage, *SS* and I_{ON}/I_{OFF} ratio of this device degrade when the insulator thickness is scaled below 3.4 nm. This degradation can be controlled by increasing Φ_M to 5.1 eV, which increases the tunneling barrier heights between the gate and the InAs channel, reduces the gate leakage, and therefore improves *SS* and I_{ON}/I_{OFF} ratio, even with an InAlAs layer scaled down to 3 nm (Figure 7.13).

The gate capacitance (C_g) of InAs QWFETs is significantly smaller than the InAlAs insulator capacitance (C_{inv}) due to small inversion layer capacitance (C_{inv}) of InAs channel. The small C_{inv} is a result of the low effective mass and the low density of states in the channel. Compared to C_{ins}, C_g is significantly suppressed due to small C_{inv} of the InGaAs/InAs/InGaAs multi-quantum-well channel and the improvement in C_g by down scaling the InAlAs insulator thickness is almost negligible (Figure 7.13(c)).

7.3.2. Gate metal work function

The I_d-V_{gs} and I_g-V_{gs} characteristics of devices with Φ_M of 4.7 and 5.1 eV are shown in Figure 7.14. Both devices have very thin InAlAs insulator (t_{ins} = 3 nm). Threshold voltages of the Φ_M = 4.7 eV device and the Φ_M = 5.1 eV device are -0.11 V and 0.29 V respectively. The positive shift in the threshold voltage is equal to the work function difference $\Delta\Phi_M$ = 0.4 eV. desirable for CMOS logic applications. The positive threshold voltage value or the enhancement mode operation of the n-type FET is highly desirable for CMOS logic applications [7, 11]. In addition to a positive threshold voltage, the device with Φ_M = 5.1 eV shows 100× smaller gate leakage current (I_g) compared to a device with Φ_M = 4.7 eV resulting in much better logic performance. Suppression of the gate leakage results in the reduction of *SS* from 118 mV/dec for Φ_M = 4.7 eV device to 91 mV/dec for Φ_M = 5.1 eV device (Figure 7.13. 10(a)). The ON current for both devices is almost the same, however, as a result of the steeper *SS* the OFF current of a device with Φ_M = 5.1 eV is reduced by a factor of 7 compared to a device with Φ_M = 4.7 eV which results in 7 times higher I_{ON}/I_{OFF} ratio (Figure 7.13(b)) in a device with Φ_M = 5.1 eV.

The mechanism of the gate leakage suppression in high Φ_M devices is explained in Figure 7.15. In the Schottky gate FET considered here, the gate leakage current arises due

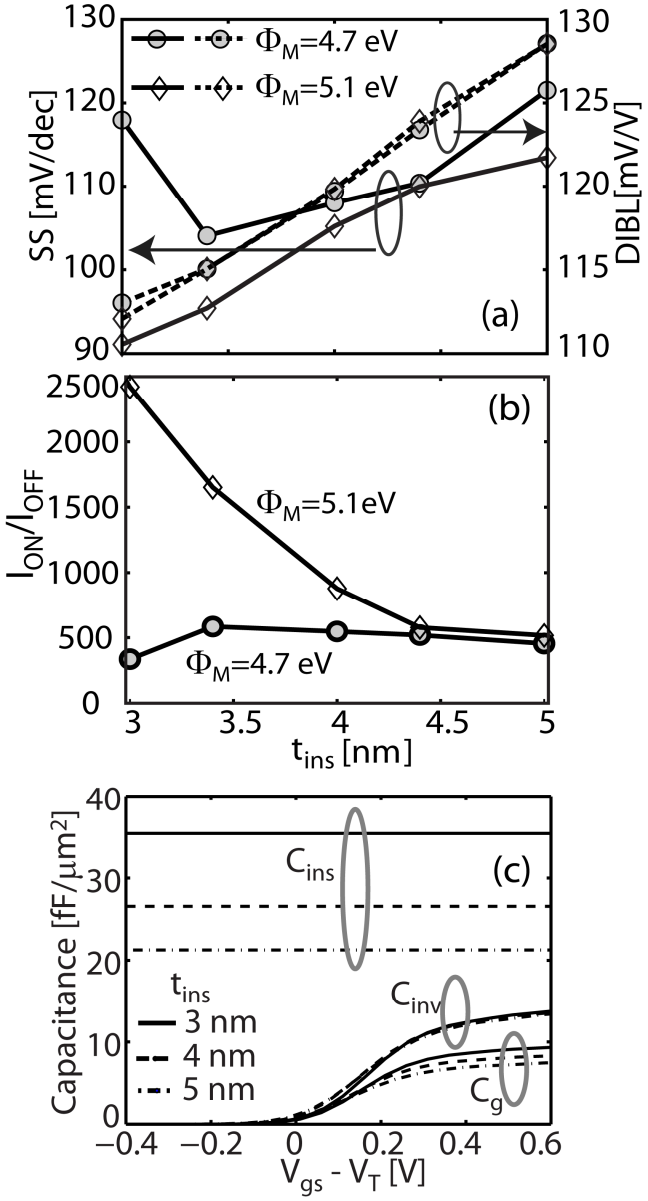

Figure 7.13. InAlAs insulator thickness (t_{ins}) scaling: (a,b) For metal work function Φ_M = 4.7 eV, the subthreshold slope (*SS*) and I_{ON}/I_{OFF} ratio improves as t_{ins} decreases till 3.4 nm. Devices with thinner insulator suffer from excessive gate leakage which leads to high *SS* and low I_{ON}/I_{OFF}. This degradation can be controlled by increasing the Φ_M to 5.1 eV. DIBL improves as t_{ins} is reduced. (c) Gate capacitance (C_g) and inversion layer capacitance (C_{inv}) as a function of gate overdrive (V_{gs}-V_T) for devices with fixed t_{InAs} = 5 nm and variable t_{ins}. C_{inv} is much lower than the insulator capacitance (C_{ins}), due to which overall gate capacitance C_g is significantly reduced. Device dimensions are same as in Figure 7.1 with L_g = 20 nm. Metal work function is Φ_M = 4.7 eV.

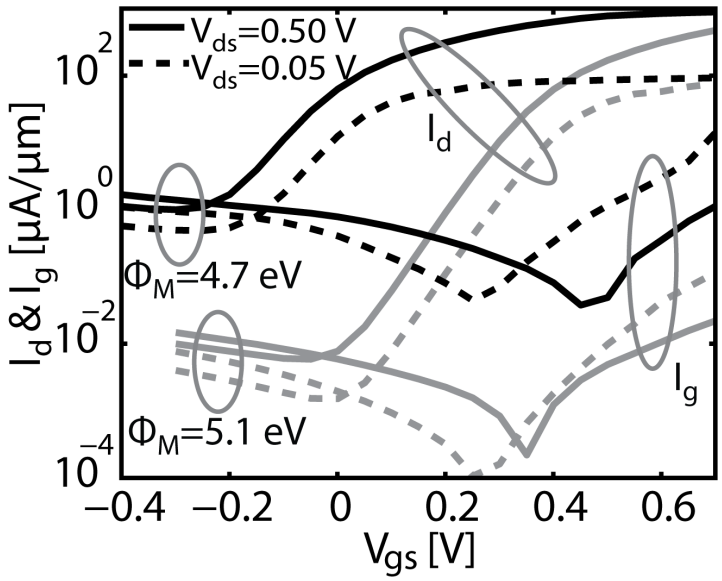

Figure 7.14. Gate leakage control by metal work function engineering: I_d-V_{gs} and I_g-V_{gs} characteristics of a L_g = 20 nm, t_{ins} = 3 nm InAs QWFET with metal work function Φ_M = 4.7 eV and 5.1 eV. Higher Φ_M significantly reduces the gate leakage current thereby improving sub-threshold slope and I_{ON}/I_{OFF} ratio. Device geometry is same as in Figure 7.1.

to the electron tunneling from the gate contact into the InAs channel through the InAlAs insulator and InGaAs barrier layers. To make a fair comparison we bias both devices at the same gate overdrive ($V_{gs} - V_T = -V_{DD}/3$) and drain bias ($V_{ds} = V_{DD}$ = 0.5 V) so that source to drain current is almost the same for both devices. The band diagram along the channel direction are shown in Figure 7.15(a,b). Same source/drain Fermi levels and similar band bending along channel direction ensure almost the same source to drain current in both devices. The gate Fermi levels in these devices are, however, separated by the gate metal work function difference of $\Delta\Phi_M$ = 0.4 eV. The electrons injected from the gate terminal experience higher InAlAs and InGaAs energy barriers in a device with Φ_M = 5.1 eV compared to a device with Φ_M = 4.7 eV. The energy barriers along the transverse direction through the center and the drain edge of the gate contact for both devices are shown in Figure 7.15(c,d). the direction of electron tunneling is shown by grey arrows. The thickness of the arrows schematically shows that the tunneling current in Φ_M = 5.1 eV device is suppressed compared to Φ_M = 4.7 eV device.

The transfer characteristics of a 20 nm InAs QWFET for metal work function values ranginf from 4.7 eV to 5.1 eV are shown in Figure 7.16. Gate leakage current is suppressed is higher metal work function devices. Figure 7.17 shows that a higher metal work function linearly pushes the threshold voltage to positive values which is desirable for CMOS logic applications. Increased Φ_M also improves the I_{ON}/I_{OFF} ratio in a

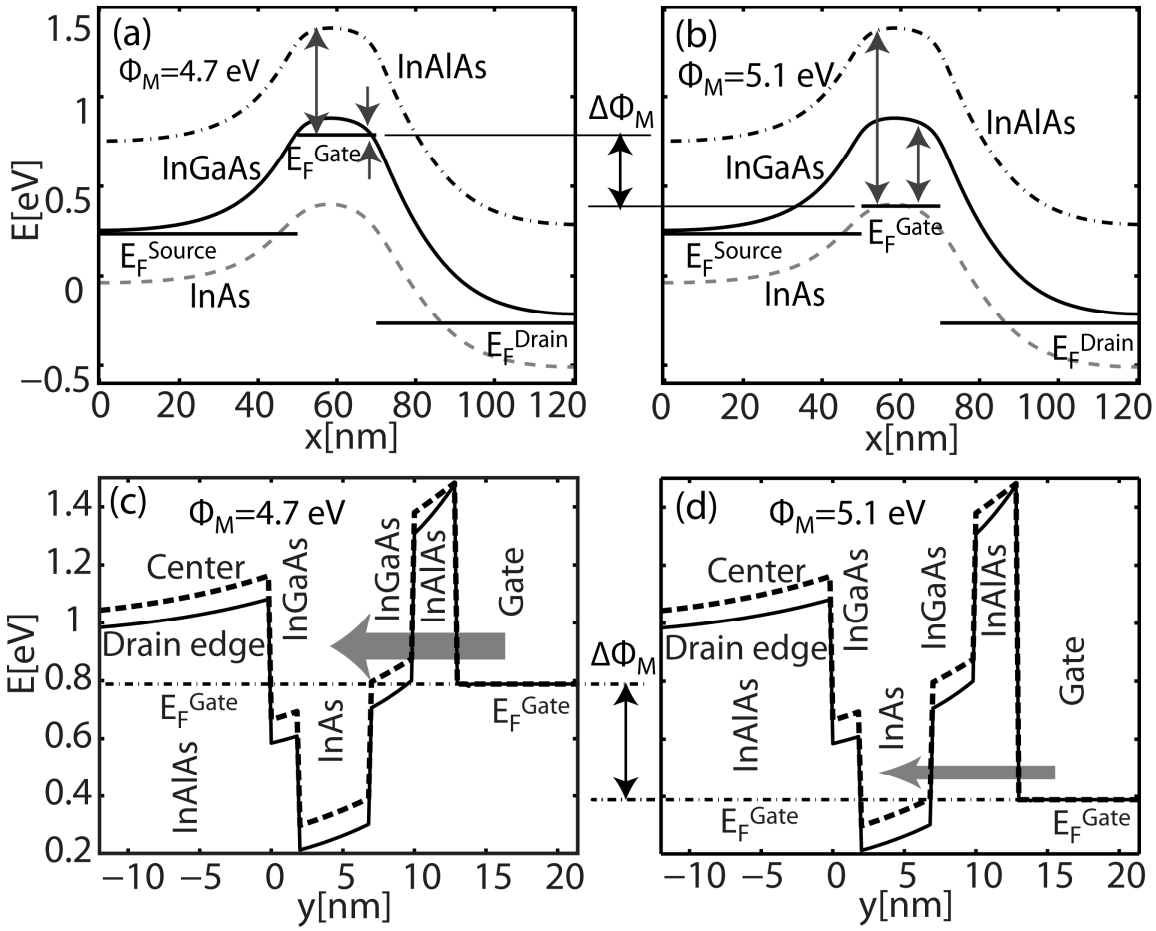

Figure 7.15. Band diagrams along the horizontal lines (a,b) through the center of InAlAs insulator, top InGaAs barrier, and InAs channel layer of same devices as in Figure 7.14 biased at the same gate overdrive $V_{gs} - V_T = -V_{DD}/3 = -0.1667$ V and $V_{ds} = 0.5$ V. Same source/drain Fermi levels and similar band bending ensures almost the same source to drain current in both devices. The electrons tunneling from the gate terminal into the channel experience taller energy barriers (grey arrows) in $\Phi_M = 5.1$ eV device (b) compared $\Phi_M = 4.7$ eV device (a) because of the metal gate Fermi level offset equal to the work function difference ($\Delta\Phi_M$). The energy barriers along the transverse directions through the center and the drain side edge of the gate contact are shown for $\Phi_M = 4.7$ eV device in (c) and for $\Phi_M = 5.1$ eV device in (d). The thickness of arrows in (c,d) schematically shows the direction of the electron tunneling and that the gate leakage is higher in $\Phi_M = 4.7$ eV device compared to $\Phi_M = 5.1$ eV device.

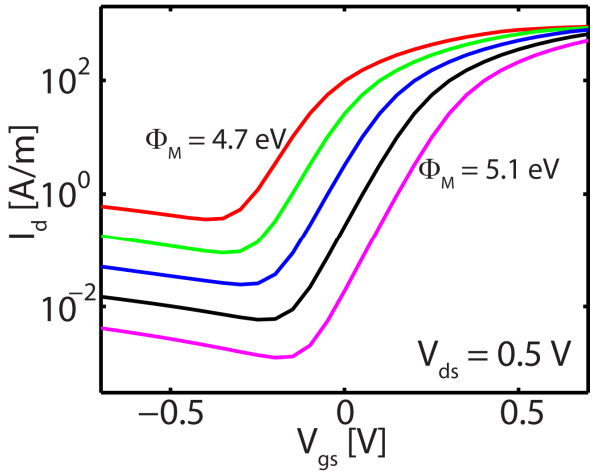

Figure 7.16. Transfer characteristics of a 20 nm InAs QWFET for metal work function values ranging from 4.7 eV to 5.1 eV in steps of 0.1 eV. The gate leakage decreases and the threshold voltage moves to positive values as the metal work function increases. Device dimensions are same as in Figure 7.1 with L_g = 20 nm and t_{ins} = 4 nm.

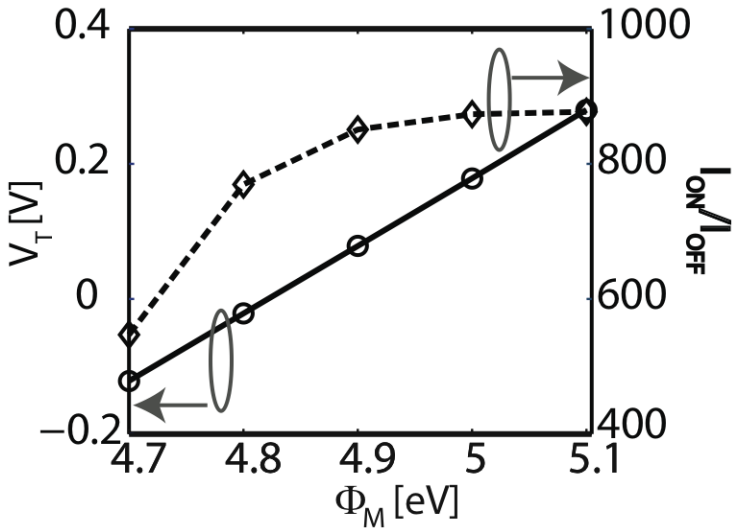

Figure 7.17. Gate metal work-function (Φ_M) engineering: Higher Φ_M results in higher threshold voltage and high I_{ON}/I_{OFF} ratio. Variation in I_{ON} with Φ_M is negligible. The high I_{ON}/I_{OFF} ratio is achieved mainly because I_{OFF} is significantly suppressed in the devices with higher Φ_M because of higher gate to channel tunneling barriers. Device dimensions are same as in Figure 7.1 with L_g = 20 nm and t_{ins} = 4 nm.

nonlinear fashion as the gate leakage is suppressed. The effects of t_{ins} and Φ_M on the device performance are strongly correlated and both these parameters should be optimized together to improve the device performance.

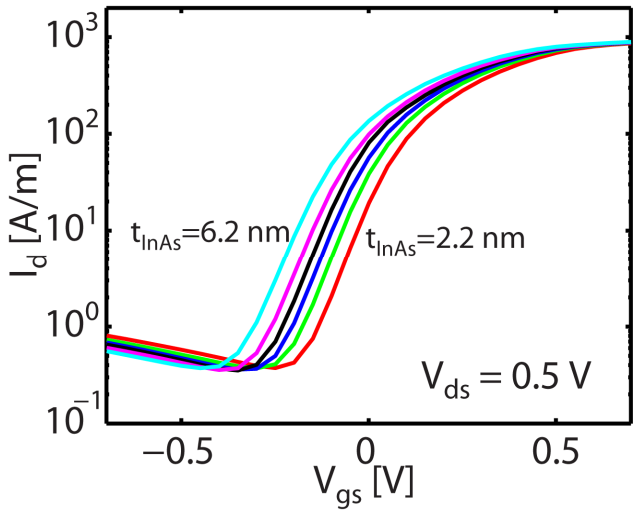

Figure 7.18. Transfer characteristics of a 20 nm InAs QWFET for InAs channel thicknesses of 2.2, 2.8, 3.4, 4.0, 5.0, and 6.2 nm. The threshold voltage moves to positive values in thin channel devices because stronger electron confinement raises energy levels in the channel. Device dimensions are same as in Figure 7.1 with L_g = 20 nm and t_{ins} = 4 nm. Metal work function is Φ_M = 4.7 eV

7.3.3. Channel thickness

The effect of scaling down the InAs channel thickness (t_{InAs}) on the logic performance can be analysed by noting that the QWFET is electrostatically very similar to the fully depleted silicon on insulator (FD-SOI) MOSFET[9]. In a QWFET, the InAs channel thickness plays a role similar to the role played by the Si body thickness to determine the scaling performance of the FD-SOI MOSFET. The effect of InAs channel thickness on transfer characteristic is depicted in Figure 7.18. Higher gate length to channel thickness ratio in thin InAs channel QWFETs result in stronger gate control which improves *SS* and *DIBL* (Figure 7.19(a)) similar to the ultra-thin Si body FD-SOI MOSFET [111]. The strong electrostatic confinement of electrons in thin InAs quantum well channel devices pushes channel conduction subbands higher in energy which subsequently results in higher threshold voltage in thinner InAs channel devices (Figure 7.19(b)). Similar threshold voltage enhancement due to scaling down the Si body thickness is observed in FD-SOI MOSFETs [112].

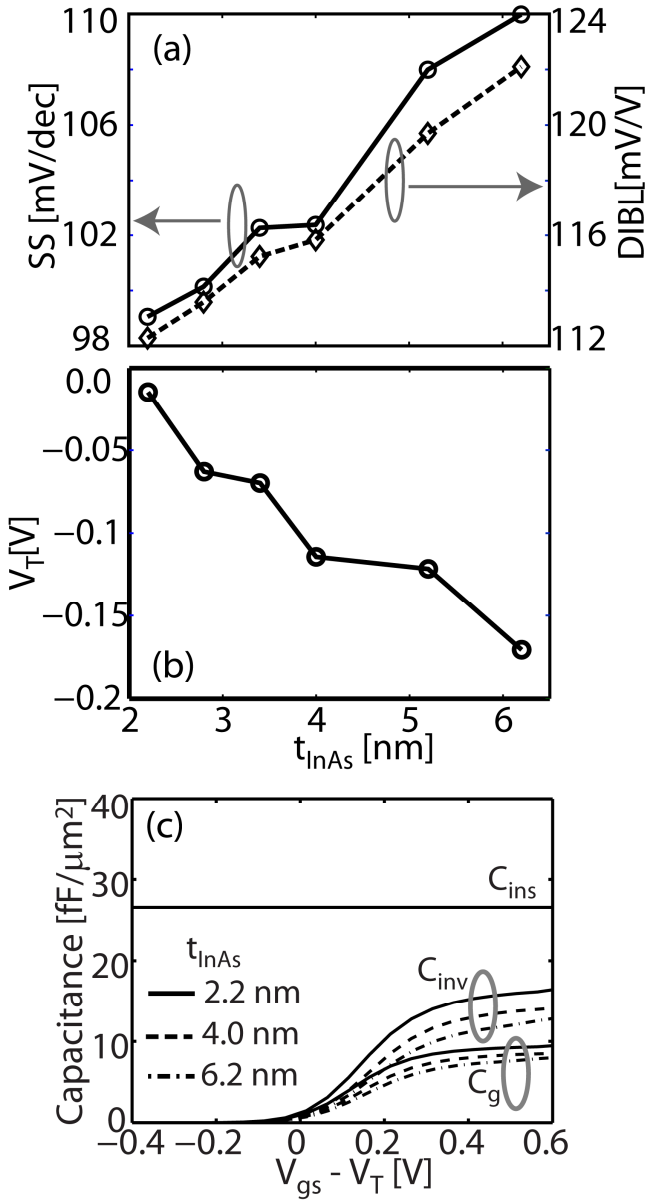

Figure 7.19. InAs quantum well thickness scaling: (a) *SS* and *DIBL* improve as the thickness of the InAs quantum well in the channel is reduced. The 2D electron gas in the channel is located closer to the gate in thinner quantum well devices which results in stronger gate control and improved short channel characteristics. (b) Devices with thinner quantum well have higher V_T as a result of the stronger quantum confinement in the channel. (c) Gate capacitance (C_g) and inversion layer capacitance (C_{inv}) as a function of gate overdrive (V_{gs}-V_T) for devices with fixed t_{ins} = 4 nm and variable t_{InAs}. C_{inv} is much lower than the insulator capacitance (C_{ins}), due to which overall gate capacitance C_g is significantly reduced. Device dimensions except InAs channel thickness are same as in Fig. 1 with L_g = 20 nm and t_{ins} = 4 nm. Metal work function is Φ_M = 4.7 eV.

The effect of InAs channel thickness on the gate capacitance is shown in Figure 7.19(c). The smaller gate capacitance (C_g) compared to the InAlAs insulator capacitance (C_{ins}) is a result of small inversion layer capacitance (C_{inv}) of InAs channel. C_{inv} can be increased by scaling down the InAs channel thickness (t_{InAs}) because stronger confinement in thinner InAs channels increases the effective mass (Figure 7.3(b)) and the density of states which results in higher C_{inv} and slightly higher C_g (Figure 7.19(b)) for the same insulator thickness.

7.4. Conclusion

Simulation tools and methodologies for the analysis and optimization of ultra-scaled InAs QWFETs are developed. The biaxial strain and quantum confinement provided by InGaAs and InAlAs barriers modify the channel effective masses. These effects are included in the atomistic supercell based strain relaxation and $sp^3d^5s^*$ tight-binding electronic structure calculations. Gate tunneling is found to be critical in the device analysis and it is included in the 2-D Schrödinger-Poisson solver by injecting carriers from the gate contact in addition to source and drain contacts. The accurate description of the shape of the gate contact is crucial to replicate the experimental results. The simulation approach is first verified for experimental devices with gate lengths ranging from 30 to 50 nm where a good quantitative match between experimental and simulated current-voltage characteristics is obtained. The scaling study on 20 nm InAs HEMT suggests that (i) the logic performance improves as InAlAs insulator thickness is scaled down to a limiting value beyond which it degrades due to excessive gate leakage, (ii) the gate leakage can be reduced by increasing the gate metal work function, (iii) the insulator thickness of a high metal work function device can be scaled down more aggressively than a low metal work function device, (iv) thin InAs channel devices exhibit superior logic performance, and (v) the gate capacitance is limited by the small inversion layer capacitance of the low effective mass InAs channel, (vi) a very good short channel performance and enhancement-mode operation can be achieved in a 20 nm InAs QWFET if a thin InAlAs insulator (3 nm) is combined with a thin InAs channel (3.4 nm), and a high metal work function (5.1 eV). Performance parameters of this optimized device are

V_T = 0.34 V, SS = 87 mV/dec, $DIBL$ = 110 mV/V, $I_{ON}/I_{OFF} = 2.9\times10^3$, C_g = 9.14 fF/μm^2, and $g_{m,max}$ = 1.7 mS/μm. A simulation tool *OMEN_FET* that generates results presented in this paper is available on nanoHUB.org [113].

8. SUMMARY

The objective of this thesis was to investigate the influence of disorder on the electronic structure and transport in nanoscale devices. The tight-binding method, which is ideally suited for modeling sharp variations on an atomic scale, was employed. The tight-binding calculations were performed in the supercell framework. Sufficiently large supercells were simulated to include random disorders such as inhomogeneous strain, surface roughness, and alloy disorder. The simulation domain contained several thousand to several million atoms depending on the size of the supercell. The electronic structure calculations were performed using NEMO-3D, while the transport calculations were performed using OMEN.

The valley degeneracy of (111) Si quantum wells was investigated. Previous effective mass based theories predicted 6-fold valley degeneracy for (111) Si quantum wells. The experimental measurements, however, show lower 2-fold and higher 4-fold valley degeneracies. It was shown that, this discrepancy between theory and experiments is a consequence of a peculiar reconstruction on (111) Si surface, which results in the formation of staircase like geometries. The supercell electronic structure calculations that included staircase like features resolved the long lasting puzzle of degeneracy breaking in (111) Si quantum wells.

Valley splitting is a critical design parameter in quantum computing devices because it controls decoherence time of a qubit. (001) Si quantum wells are of particular interest as several quantum computing architectures based on (001) SGe/Si/SiGe heterostructures have recently been proposed. Existing theories could not explain the linear magnetic field dependence and suppression of valley splitting in (001) SiGe/Si/SiGe heterostructures. It was shown that the suppression and the linear magnetic field dependence of valley splitting is associated with the fact that SiGe/Si/SiGe heterostructures were grown on

miscut (001) substrate as opposed to flat (001) substrate. It was found that the models that assume perfect miscut geometry under-estimate valley splitting. The supercell calculations that include step roughness at the miscut interface and SiGe alloy disorder were found to predict the experimentally observed valley splitting within variations induced by the presence of disorders.

The supercell technique was used to calculate the electronic structure of alloy nanowires. Although the supercell calculation includes effect of disorders, it can only deliver the absolute band extrema. Important transport parameters such as effective mass and relative positions of valley minima can not be determined from supercell calculations alone. To solve this problem, the zone-unfolding method, which can unfold the supercell bandstructure to obtain an "approximate" bandstructure of the disordered system, was adopted. Electronic transport in AlGaAs and SiGe alloy nanowires was studied using the supercell technique and the zone-unfolding method. The "approximate" bandstructures were compared with the transmission coefficients through the same nanowires. It was found that, as opposed to smooth step-like transmission coefficients in ideal nanowires, the transmission coefficients in disordered nanowires are noisy and smaller in magnitude. The supercell approach provides a unified picture of alloy nanowires, in which the nanodevice (transmission) and nanomaterials (bandstructure) viewpoints complement each other and illuminate the interesting physics of these disordered nanostructures that otherwise can not be explained using the traditional averaging methods such as the virtual crystal approximation.

The supercell approach to the bandstructure calculation was coupled with the effective mass based quantum transport model to develop a multiscale device modeling approach. The approach was used to investigate the scaling performance of composite InGaAs/InAs/InGaAs multi-quantum-well channel quantum well field effect transistors (QWFETs). The 2D real space effective mass based quantum transport simulator used here is capable of simulating the gate contact at the quantum mechanical level. The simulation methodology developed here was able to reproduce the experimental I_d-V_{gs} characteristics in both the OFF- and ON-state as opposed to earlier approaches which could not match the OFF-state regime of I_d-V_{gs} characteristics that is dominated by the

gate leakage. The simulation methodology was employed to explore the device design optimizations for 20 nm InAs QWFET. It was found that InAs QWFETs can be successfully scaled to 20 nm gate lengths and that the best performance can be achieved in thin InAs channel devices by reducing the insulator thickness to improve the gate control, increasing the gate work function to suppress the gate leakage, and maintaining a flat gate geometry.

In general, the supercell tight-binding approach is a very useful tool to model atomic scale disorders such as surface roughness, inhomogeneous strain, and random alloy disorder. The supercell approach coupled with the zone-unfolding method provides an efficient way to compute the “approximate” bandstructure of disordered nanostructures. The transport parameters such as effective mass and relative locations of valley-minima, extracted from the “approximate” bandstructure include the effect of atomic scale disorders. The electronic structure parameters extracted from the “approximate” bandstructure can be used in simple models such as the top-of-the-barrier-model or more advanced effective mass based NEGF simulators to analyse device performance.

LIST OF REFERENCES

LIST OF REFERENCES

[1] ITRS. "International Technology Roadmap for Semiconductors," HTTP://WWW.ITRS.NET/LINKS/2007ITRS/HOME2007.HTM.

[2] M. Bohr, “Nanotechnology goals and challenges for electronic applications,” *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 1, no. 1, pp. 56-62, 2002.

[3] R. Chau, S. Datta, M. Doczy *et al.*, “Benchmarking nanotechnology for high-performance and low-power logic transistor applications,” *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 4, no. 2, pp. 153-158, 2005.

[4] B. Doyle, S. Datta, M. Doczy *et al.*, “High performance fully-depleted tri-gate CMOS transistors,” *IEEE ELECTRON DEVICE LETTERS,* vol. 24, no. 4, pp. 263-265, 2003.

[5] B. Yu, L. Chang, S. Ahmed *et al.*, “FinFET scaling to 10nm gate length,” *INTERNATIONAL ELECTRON DEVICES 2002 MEETING, TECHNICAL DIGEST*, pp. 251-254, 2002.

[6] D. Kim, and J. del Alamo, “Scaling behavior of In0.7Ga0.3AsHEMTs for logic,” *2006 INTERNATIONAL ELECTRON DEVICES MEETING, VOLS 1 AND 2*, pp. 587-590, 2006.

[7] D. Kim, and J. del Alamo, “30 nm E-mode InAs PHEMTs for THz and Future Logic Applications,” *IEEE INTERNATIONAL ELECTRON DEVICES MEETING 2008, TECHNICAL DIGEST*, pp. 719-722, 2008.

[8] D. Kim, and J. del Alamo, “30-nm InAs pseudomorphic HEMTs on an InP substrate with a current-gain cutoff frequency of 628 GHz,” *IEEE ELECTRON DEVICE LETTERS,* vol. 29, no. 8, pp. 830-833, 2008.

[9] D. Kim, and J. del Alamo, "Lateral and Vertical Scaling of In0.7Ga0.3As HEMTs for Post-Si-CMOS Logic Applications," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 55, no. 10, pp. 2546-2553, 2008.

[10] G. Dewey, M. Hudait, K. Lee *et al.*, "Carrier transport in high-mobility III-V quantum-well transistors and performance impact for high-speed low-power logic applications," *IEEE ELECTRON DEVICE LETTERS,* vol. 29, no. 10, pp. 1094-1097, 2008.

[11] M. Hudait, G. Dewey, S. Datta *et al.*, "Heterogeneous integration of enhancement mode In0.7Ga0.3As quantum well transistor on silicon substrate using thin (<= 2 gm) composite buffer architecture for high-speed and low-voltage (0.5V) logic applications," *2007 IEEE INTERNATIONAL ELECTRON DEVICES MEETING, VOLS 1 AND 2*, pp. 625-628, 2007.

[12] M. Radosavljevic, T. Ashley, A. Andreev *et al.*, "High-Performance 40nm Gate Length InSb P-Channel Compressively Strained Quantum Well Field Effect Transistors for Low-Power (V-CC=0.5V) Logic Applications," *IEEE INTERNATIONAL ELECTRON DEVICES MEETING 2008, TECHNICAL DIGEST*, pp. 727-730, 2008.

[13] Y. Xuan, Y. Wu, H. Lin *et al.*, "Submicrometer inversion-type enhancement-mode InGaAs MOSFET with atomic-layer-deposited Al2O3 as gate dielectric," *IEEE ELECTRON DEVICE LETTERS,* vol. 28, no. 11, pp. 935-938, 2007.

[14] P. McEuen, M. Fuhrer, and H. Park, "Single-walled carbon nanotube electronics," *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 1, no. 1, pp. 78-85, 2002.

[15] J. Guo, A. Javey, H. Dai *et al.*, "Performance analysis and design optimization of near ballistic carbon nanotube field-effect transistors," *IEEE INTERNATIONAL ELECTRON DEVICES MEETING 2004, TECHNICAL DIGEST*, pp. 703-706, 2004.

[16] A. Javey, R. Tu, D. Farmer *et al.*, "High performance n-type carbon nanotube field-effect transistors with chemically doped contacts," *NANO LETTERS,* vol. 5, no. 2, pp. 345-348, 2005.

[17] P. Burke, "An RF circuit model for carbon nanotubes," *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 2, no. 1, pp. 55-58, 2003.

[18] M. A. Nielsen, and I. L. Chuang, *Quantum computation and quantum information*, Cambridge ; New York: Cambridge University Press, 2000.

[19] P. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM JOURNAL ON COMPUTING,* vol. 26, no. 5, pp. 1484-1509, 1997.

[20] L. Grover, "Quantum mechanics helps in searching for a needle in a haystack," *PHYSICAL REVIEW LETTERS,* vol. 79, no. 2, pp. 325-328, 1997.

[21] J. Wang, E. Polizzi, A. Ghosh *et al.*, "Theoretical investigation of surface roughness scattering in silicon nanowire transistors," *APPLIED PHYSICS LETTERS,* vol. 87, no. 4, pp. 043101 1-3, 2005.

[22] T. Boykin, N. Kharche, G. Klimeck *et al.*, "Approximate bandstructures of semiconductor alloys from tight-binding supercell calculations," *JOURNAL OF PHYSICS-CONDENSED MATTER,* vol. 19, no. 3, pp. 036203 1-14, 2007.

[23] T. Boykin, M. Luisier, A. Schenk *et al.*, "The electronic structure and transmission characteristics of disordered AlGaAs nanowires," *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 6, no. 1, pp. 43-47, 2007.

[24] N. Kharche, M. Prada, T. Boykin *et al.*, "Valley splitting in strained silicon quantum wells modeled with 2 degrees miscuts, step disorder, and alloy disorder," *APPLIED PHYSICS LETTERS,* vol. 90, no. 9, pp. 092109 1-3, 2007.

[25] N. Kharche, M. Luisier, T. Boykin *et al.*, "Electronic Structure and Transmission Characteristics of SiGe Nanowires," JOURNAL OF COMPUTATIONAL ELECTRONICS*,* vol. 7, pp. 350-354, 2008.

[26] M. Luisier, A. Schenk, and W. Fichtner, "Atomistic treatment of interface roughness in Si nanowire transistors with different channel orientations," *APPLIED PHYSICS LETTERS,* vol. 90, no. 10, pp. 102103 1-3, 2007.

[27] N. Azizi, M. Khellah, V. De *et al.*, "Variations-aware low-power design and block clustering with voltage scaling," *IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS,* vol. 15, no. 7, pp. 746-757, 2007.

[28] J. Wang, A. Rahman, A. Ghosh *et al.*, "On the validity of the parabolic effective-mass approximation for the I-V calculation of silicon nanowire transistors," *IEEE*

TRANSACTIONS ON ELECTRON DEVICES, vol. 52, no. 7, pp. 1589-1595, 2005.

[29] G. Klimeck, R. Lake, R. Bowen *et al.*, "Quantum device simulation with a generalized tunneling formula," APPLIED PHYSICS LETTERS*,* vol. 67, no. 17, pp. 2539-2541, 1995.

[30] R. Lake, G. Klimeck, R. Bowen *et al.*, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *JOURNAL OF APPLIED PHYSICS,* vol. 81, no. 12, pp. 7845-7869, 1997.

[31] G. Klimeck, F. Oyafuso, R. Bowen *et al.*, "3-D atomistic nanoelectronic modeling on high performance clusters: multimillion atom simulations," *SUPERLATTICES AND MICROSTRUCTURES,* vol. 31, no. 2-4, pp. 171-179, 2002.

[32] A. Canning, L. Wang, A. Williamson *et al.*, "Parallel empirical pseudopotential electronic structure calculations for million atom systems," *JOURNAL OF COMPUTATIONAL PHYSICS,* vol. 160, no. 1, pp. 29-41, 2000.

[33] T. Mattila, L. Wang, and A. Zunger, "Electronic consequences of lateral composition modulation in semiconductor alloys," *PHYSICAL REVIEW B,* vol. 59, no. 23, pp. 15270-15284, 1999.

[34] T. Mattila, S. Wei, and A. Zunger, "Localization and anticrossing of electron levels in GaAs1-xNx alloys," *PHYSICAL REVIEW B,* vol. 60, no. 16, pp. 11245-11248, 1999.

[35] P. Kent, and A. Zunger, "Theory of electronic structure evolution in GaAsN and GaPN alloys," *PHYSICAL REVIEW B,* vol. 64, no. 11, pp. 115208 1-23, 2001.

[36] L. Wang, L. Bellaiche, S. Wei *et al.*, ""Majority representation" of alloy electronic states," *PHYSICAL REVIEW LETTERS,* vol. 80, no. 21, pp. 4725-4728, 1998.

[37] J. Li, and L. Wang, "First principles calculations of ZnS : Te energy levels," *PHYSICAL REVIEW B,* vol. 67, no. 20, pp. 205319 1-11, 2003.

[38] K. Kim, P. Kent, A. Zunger *et al.*, "Atomistic description of the electronic structure of InxGa1-xAs alloys and InAs/GaAs superlattices," *PHYSICAL REVIEW B,* vol. 66, no. 4, pp. 045208 1-15, 2002.

[39] G. Klimeck, S. Ahmed, H. Bae *et al.*, "Atomistic simulation of realistically sized nanodevices using NEMO 3-D - Part I: Models and benchmarks," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 54, no. 9, pp. 2079-2089, 2007.

[40] G. Klimeck, S. Ahmed, N. Kharche *et al.*, "Atomistic simulation of realistically sized nanodevices using NEMO 3-D - Part II: Applications," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 54, no. 9, pp. 2090-2099, 2007.

[41] T. Whall, and E. Parker, "SiGe heterostructures for FET applications," *JOURNAL OF PHYSICS D-APPLIED PHYSICS,* vol. 31, no. 12, pp. 1397-1416, 1998.

[42] T. Boykin, and G. Klimeck, "Practical application of zone-folding concepts in tight-binding calculations," *PHYSICAL REVIEW B,* vol. 71, no. 11, pp. 115215 1-6, 2005.

[43] N. Kharche, S. Kim, T. Boykin *et al.*, "Valley degeneracies in (111) silicon quantum wells," *APPLIED PHYSICS LETTERS,* vol. 94, no. 4, pp. 042101 1-3, 2009.

[44] N. Kharche, M. Prada, T. Boykin *et al.*, "Valley splitting in strained silicon quantum wells modeled with 2 degrees miscuts, step disorder, and alloy disorder," *APPLIED PHYSICS LETTERS,* vol. 90, no. 9, pp. 092109 1-3, 2007.

[45] K. Eng, R. McFarland, and B. Kane, "Integer quantum Hall effect on a six-valley hydrogen-passivated silicon (111) surface," *PHYSICAL REVIEW LETTERS,* vol. 99, no. 1, pp. 016801 1-4, 2007.

[46] H. Watanabe, K. Fujita, and M. Ichikawa, "Atomic-step observation at buried SiO2/Si(111) interfaces by scanning reflection electron microscopy," *SURFACE SCIENCE,* vol. 385, no. 2-3, pp. L952-L957, 1997.

[47] H. Momose, T. Ohguro, S. Nakamura *et al.*, "Ultrathin gate oxide CMOS on (111) surface-oriented Si substrate," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 49, no. 9, pp. 1597-1605, 2002.

[48] H. Zandvliet, and H. Elswijk, "Morphology of monatomic step edges on vicinal Si(001)," *PHYSICAL REVIEW B,* vol. 48, no. 19, pp. 14269-14275, 1993.

[49] F. Stern, and W. Howard, "Properties of Semiconductor Surface Inversion Layers in the Electric Quantum Limit," *PHYSICAL REVIEW,* vol. 163, no. 3, pp. 816-835, 1967.

[50] T. Ando, "Valley splitting in the silicon inversion layer: Misorientation effects," *PHYSICAL REVIEW B,* vol. 19, no. 6, pp. 3089-3095, 1979.

[51] T. Ando, A. Fowler, and F. Stern, "Electronic-properties of two-dimensional systems," *REVIEWS OF MODERN PHYSICS,* vol. 54, no. 2, pp. 437-672, 1982.

[52] T. Boykin, G. Klimeck, and F. Oyafuso, "Valence band effective-mass expressions in the sp3d5s* empirical tight-binding model applied to a Si and Ge parametrization," *PHYSICAL REVIEW B,* vol. 69, no. 11, pp. 115201 1-10, 2004.

[53] D. Loss, and D. DiVincenzo, "Quantum computation with quantum dots," *PHYSICAL REVIEW A,* vol. 57, no. 1, pp. 120-126, 1998.

[54] R. Vrijen, E. Yablonovitch, K. Wang *et al.*, "Electron-spin-resonance transistors for quantum computing in silicon-germanium heterostructures," *PHYSICAL REVIEW A,* vol. 62, no. 1, pp. 012306 1-10, 2000.

[55] M. Friesen, P. Rugheimer, D. Savage *et al.*, "Practical design and simulation of silicon-based quantum-dot qubits," *PHYSICAL REVIEW B,* vol. 67, no. 12, pp. 121301(R) 1-4, 2003.

[56] M. A. Eriksson, M. Friesen, S. N. Coppersmith *et al.*, "Spin Based Quantum Dot Quantum Computing in Silicon," QUANTUM INFORMATION PROCESSING, vol. 3, no. 1-5, pp. 14, 2004.

[57] C. Tahan, M. Friesen, and R. Joynt, "Decoherence of electron spin qubits in Si-based quantum computers," *PHYSICAL REVIEW B,* vol. 66, no. 3, pp. 035314 1-11, 2002.

[58] T. Boykin, G. Klimeck, M. Friesen *et al.*, "Valley splitting in low-density quantum-confined heterostructures studied using tight-binding models," *PHYSICAL REVIEW B,* vol. 70, no. 16, pp. 165325 1-12, 2004.

[59] T. Boykin, G. Klimeck, M. Eriksson *et al.*, "Valley splitting in strained silicon quantum wells," *APPLIED PHYSICS LETTERS,* vol. 84, no. 1, pp. 115-117, 2004.

[60] S. Goswami, K. Slinker, M. Friesen *et al.*, "Controllable valley splitting in silicon quantum devices," *NATURE PHYSICS,* vol. 3, no. 1, pp. 41-45, 2007.

[61] S. Lee, and P. von Allmen, "Magnetic-field dependence of valley splitting in Si quantum wells grown on tilted SiGe substrates," *PHYSICAL REVIEW B,* vol. 74, no. 24, pp. 245302 1-5, 2006.

[62] M. Luisier, N. Neophytou, N. Kharche *et al.*, "Full-Band and Atomistic Simulation of Realistic 40 nm InAs HEMT," *IEEE INTERNATIONAL ELECTRON DEVICES MEETING 2008, TECHNICAL DIGEST*, pp. 887-890, 2008.

[63] N. Kharche, G. Klimeck, D.-H. Kim *et al.*, "Performance Analysis of Ultra-Scaled InAs HEMTs," *IEEE INTERNATIONAL ELECTRON DEVICES MEETING 2009, TECHNICAL DIGEST*, pp. 491-494, 2009.

[64] B. Kane, "A silicon-based nuclear spin quantum computer," *NATURE,* vol. 393, no. 6681, pp. 133-137, 1998.

[65] I. Appelbaum, B. Huang, and D. Monsma, "Electronic measurement and control of spin transport in silicon," *NATURE,* vol. 447, no. 7142, pp. 295-298, 2007.

[66] D. Tsui, and G. Kaminsky, "Valley degeneracy of electrons in accumulation and inversion layers on Si (111) surface," *SOLID STATE COMMUNICATIONS,* vol. 20, no. 1, pp. 93-95, 1976.

[67] T. Cole, and B. Mccombe, "Intersubband spectroscopy and valley degeneracy of Si(110) and Si(111) n-type inversion layers," *PHYSICAL REVIEW B,* vol. 29, no. 6, pp. 3180-3192, 1984.

[68] M. Kelly, and L. Falicov, "Electronic Structure of Inversion Layers in Many-Valley Semiconductors," *PHYSICAL REVIEW LETTERS,* vol. 37, no. 15, pp. 1021-1024, 1976.

[69] E. Hwang, and S. Das Sarma, "Transport properties of two-dimensional electron systems on silicon (111) surfaces," *PHYSICAL REVIEW B,* vol. 75, no. 7, pp. 073301 1-4, 2007.

[70] T. Boykin, N. Kharche, and G. Klimeck, "Non-primitive rectangular cells for tight-binding electronic structure calculations," *PHYSICA E-LOW-*

DIMENSIONAL SYSTEMS & NANOSTRUCTURES, vol. 41, no. 3, pp. 490-494, 2009.

[71] M. Friesen, M. Eriksson, and S. Coppersmith, "Magnetic field dependence of valley splitting in realistic Si/SiGe quantum wells," *APPLIED PHYSICS LETTERS,* vol. 89, no. 20, pp. 202106 1-3, 2006.

[72] M. Friesen, S. Chutia, C. Tahan *et al.*, "Valley splitting theory of SiGe/Si/SiGe quantum wells," *PHYSICAL REVIEW B,* vol. 75, no. 11, pp. 115318 1-12, 2007.

[73] L. Sham, S. Allen, A. Kamgar *et al.*, "Valley-Valley Splitting in Inversion Layers on a High-Index Surface of Silicon," *PHYSICAL REVIEW LETTERS,* vol. 40, no. 7, pp. 472-475, 1978.

[74] J. Salles, H. Closs, J. Senna *et al.*, "Minigaps and high-density activated transport in n-channel Si inversion layers," *PHYSICAL REVIEW B,* vol. 37, no. 15, pp. 8912-8914, 1988.

[75] T. Thornton, F. Ge, A. Andresen *et al.*, "Minigaps in strained silicon quantum wells on tilted substrates," *JOURNAL OF VACUUM SCIENCE & TECHNOLOGY B,* vol. 17, no. 4, pp. 1757-1760, 1999.

[76] T. Boykin, G. Klimeck, and F. Oyafuso, "Valence band effective-mass expressions in the sp(3)d(5)s(*) empirical tight-binding model applied to a Si and Ge parametrization," *PHYSICAL REVIEW B,* vol. 69, no. 11, pp. 115201 1-10, 2004.

[77] G. Klimeck, F. Oyafuso, T. Boykin *et al.*, "Development of a nanoelectronic 3-D (NEMO 3-D) simulator for multimillion atom simulations and its application to alloyed quantum dots," *CMES-COMPUTER MODELING IN ENGINEERING & SCIENCES,* vol. 3, no. 5, pp. 601-642, 2002.

[78] 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, pp. 637-645, 1966.

[79] T. Boykin, N. Kharche, and G. Klimeck, "Brillouin-zone unfolding of perfect supercells having nonequivalent primitive cells illustrated with a Si/Ge tight-binding parameterization," *PHYSICAL REVIEW B,* vol. 76, no. 3, pp. 035310 1-7, 2007.

[80] T. Boykin, G. Klimeck, R. Bowen *et al.*, “Diagonal parameter shifts due to nearest-neighbor displacements in empirical tight-binding theory,” *PHYSICAL REVIEW B,* vol. 66, no. 12, pp. 125207 1-6, 2002.

[81] T. Boykin, R. Bowen, and G. Klimeck, “Electromagnetic coupling and gauge invariance in the empirical tight-binding method,” *PHYSICAL REVIEW B,* vol. 6324, no. 24, pp. 245314 1-17, 2001.

[82] M. Graf, and P. Vogl, “Electromagnetic fields and dielectric response in empirical tight-binding theory,” *PHYSICAL REVIEW B,* vol. 51, no. 8, pp. 4940-4949, 1995.

[83] T. Boykin, N. Kharche, and G. Klimeck, “Valley splitting in finite barrier quantum wells,” *PHYSICAL REVIEW B,* vol. 77, no. 24, pp. 245320 1-8, 2008.

[84] B. Swartzentruber, Y. Mo, R. Kariotis *et al.*, “Direct determination of step and kink energies on vicinal Si(001),” *PHYSICAL REVIEW LETTERS,* vol. 65, no. 15, pp. 1913-1916, 1990.

[85] M. Korkusinski, G. Klimeck, H. Xu *et al.*, "Atomistic simulations in nanostructures composed of tens of millions of atoms: importance of long-range strain effects in quantum dots."

[86] R. Rahman, C. Wellard, F. Bradbury *et al.*, “High precision quantum control of single donor spins in silicon,” *PHYSICAL REVIEW LETTERS,* vol. 99, no. 3, pp. 036403 1-4, 2007.

[87] S. Park, R. Rahman, G. Klimeck *et al.*, “Mapping Donor Electron Wave Function Deformations at a Sub-Bohr Orbit Resolution,” *PHYSICAL REVIEW LETTERS,* vol. 103, no. 10, pp. 106802 1-4, 2009.

[88] R. Rahman, S. Park, J. Cole *et al.*, “Atomistic simulations of adiabatic coherent electron transport in triple donor systems,” *PHYSICAL REVIEW B,* vol. 80, no. 3, pp. 035302 1-7, 2009.

[89] G. Lansbergen, R. Rahman, C. Wellard *et al.*, “Gate-induced quantum-confinement transition of a single dopant atom in a silicon FinFET,” *NATURE PHYSICS,* vol. 4, no. 8, pp. 656-661, 2008.

[90] M. Usman, H. Ryu, I. Woo *et al.*, "Moving Toward Nano-TCAD Through Multimillion-Atom Quantum-Dot Simulations Matching Experimental Data," *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 8, no. 3, pp. 330-344, 2009.

[91] T. Boykin, N. Kharche, and G. Klimeck, "Allowed wavevectors under the application of incommensurate periodic boundary conditions," *EUROPEAN JOURNAL OF PHYSICS,* vol. 27, no. 1, pp. 5-10, 2006.

[92] T. Boykin, N. Kharche, and G. Klimeck, "Evolution time and energy uncertainty," *EUROPEAN JOURNAL OF PHYSICS,* vol. 28, no. 4, pp. 673-678, 2007.

[93] Y. Cui, L. Lauhon, M. Gudiksen *et al.*, "Diameter-controlled synthesis of single-crystal silicon nanowires," *APPLIED PHYSICS LETTERS,* vol. 78, no. 15, pp. 2214-2216, 2001.

[94] D. Wang, Q. Wang, A. Javey *et al.*, "Germanium nanowire field-effect transistors with SiO2 and high-kappa HfO2 gate dielectrics," *APPLIED PHYSICS LETTERS,* vol. 83, no. 12, pp. 2432-2434, 2003.

[95] C. Kim, J. Yang, H. Lee *et al.*, "Fabrication of Si1-xGex alloy nanowire field-effect transistors," *APPLIED PHYSICS LETTERS,* vol. 91, no. 3, pp. 033104 1-3, 2007.

[96] S. Lee, F. Oyafuso, P. von Allmen *et al.*, "Boundary conditions for the electronic structure of finite-extent embedded semiconductor nanostructures," *PHYSICAL REVIEW B,* vol. 69, no. 4, pp. 045316 1-8, 2004.

[97] M. Luisier, A. Schenk, W. Fichtner *et al.*, "Atomistic simulation of nanowires in the sp3d5s* tight-binding formalism: From boundary conditions to strain calculations," *PHYSICAL REVIEW B,* vol. 74, no. 20, pp. 205323 1-12, 2006.

[98] A. Rahman, J. Guo, S. Datta *et al.*, "Theory of ballistic nanotransistors," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 50, no. 9, pp. 1853-1864, 2003.

[99] S. Datta, T. Ashley, J. Brask *et al.*, "85nm gate length enhancement and depletion mode InSb quantum well transistors for ultra high speed and very low power digital logic applications," *IEEE INTERNATIONAL ELECTRON DEVICES MEETING 2005, TECHNICAL DIGEST*, pp. 783-786, 2005.

[100] S. Datta, and R. Chau, "Silicon and III-V nanoelectronics," *2005 INTERNATIONAL CONFERENCE ON INDIUM PHOSPHIDE AND RELATED MATERIALS*, pp. 7-8, 2005.

[101] N. Neophytou, T. Rakshit, and M. Lundstrom, "Performance Analysis of 60-nm Gate-Length III-V InGaAs HEMTs: Simulations Versus Experiments," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 56, no. 7, pp. 1377-1387, 2009.

[102] M. Luisier, and A. Schenk, "Two-dimensional tunneling effects on the leakage current of MOSFETs with single dielectric and high-kappa gate stacks," *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 55, no. 6, pp. 1494-1501, 2008.

[103] R. Venugopal, S. Goasguen, S. Datta *et al.*, "Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors," *JOURNAL OF APPLIED PHYSICS,* vol. 95, no. 1, pp. 292-305, 2004.

[104] M. Fischetti, L. Wang, B. Yu *et al.*, "Simulation of electron transport in high-mobility MOSFETs: Density of states bottleneck and source starvation," *2007 IEEE INTERNATIONAL ELECTRON DEVICES MEETING, VOLS 1 AND 2*, pp. 109-112, 2007.

[105] J. Mikkelsen, and J. Boyce, "Atomic-Scale Structure of Random Solid Solutions: Extended X-Ray-Absorption Fine-Structure Study of Ga1-xInxAs," *PHYSICAL REVIEW LETTERS,* vol. 49, no. 19, pp. 1412-1415, 1982.

[106] P. William, F. Brian, T. Saul *et al.*, *Numerical Recipes in C: The Art of Scientific Computing*: Cambridge University Press, 2002.

[107] NSM. "NSM Archive - Physical properties of Semiconductors," HTTP://WWW.IOFFE.RSSI.RU/SVA/NSM/SEMICOND/INDEX.HTML.

[108] P. Bhattacharya, *Properties of Lattice-Matched and Strained Indium Gallium Arsenide*: Institution of Engineering and Technology, 1993.

[109] D. K. Schroder, *Semiconductor material and device characterization, Chapter 4*, 3rd ed., [Piscataway, NJ] Hoboken, N.J.: IEEE Press ; Wiley, 2006.

[110] S. Takagi, and A. Toriumi, “Quantitative understanding of inversion-layer capacitance in Si MOSFET's,” *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 42, no. 12, pp. 2125-2130, 1995.

[111] R. Yan, A. Ourmazd, and K. Lee, “Scaling the Si MOSFET: from bulk to SOI to bulk,” *IEEE TRANSACTIONS ON ELECTRON DEVICES,* vol. 39, no. 7, pp. 1704-1710, 1992.

[112] S. Ramey, and D. Ferry, “Threshold voltage calculation in ultrathin-film SOI MOSFETs using the effective potential,” *IEEE TRANSACTIONS ON NANOTECHNOLOGY,* vol. 2, no. 3, pp. 121-125, 2003.

[113] N. Kharche, M. Luisier, G. Howlett *et al.* "OMEN_FET: Effective mass based 2-D real-space quantum transport simulator," https://nanohub.org/tools/omenhfet/.

VITA

VITA

Neerav Kharche was born in Ahmedabad, India, in 1981. Neerav received the B.Tech. and the M.Tech. degrees in metallurgical engineering and materials science from the Indian Institute of Technology, Bombay, India, in 2003. From 2003-2009, Neerav has been pursuing his PhD degree at Purdue University in the area of Microelectronics and Nanotechnology under the supervision of Professor Gerhard Klimeck and Professor Timothy Boykin, who is Professor in The University of Alabama in Huntsville. Neerav's research interests include theoretical and computational modeling of the electronic structure and transport in nanoscale devices.