THEORETICAL STUDY OF QUANTUM TRANSPORT IN REALISTIC SEMICONDUCTOR DEVICES

by

Pratik B. Vyas

APPROVED BY SUPERVISORY COMMITTEE:

Dr. Massimo V. Fischetti, Chair

Dr. Lawrence Overzet

Dr. Mark Lee

Dr. William Frensley

Dr. William G. VanDenberghe

Copyright © 2019

Pratik B. Vyas

All rights reserved

Dedicated to my grandmother.

THEORETICAL STUDY OF QUANTUM TRANSPORT IN REALISTIC SEMICONDUCTOR DEVICES

by

PRATIK B. VYAS, BS, MS

DISSERTATION

Presented to the Faculty of
The University of Texas at Dallas
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY IN
ELECTRICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS
May 2019

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Massimo (Max) Fischetti, for the continued support of my research and for his patience, motivation and immense knowledge. His vigorous passion and love for science is an inspiration for me and something I look up to. His guidance has helped me overcome the obstacles I have faced in my work and I cannot imagine having a better mentor for my PhD. Besides my advisor, I am deeply grateful to Dr. Maarten Van de Put for his time and invaluable support. His expertise in the subject and the insightful conversations we've had prove to be a constant source of motivation and encouragement. I am indebted and thankful to Dr. William Vandenberghe for his constructive guidance during the early years of my PhD and bearing with my amateur queries. I am very grateful for the time and support given by the other members of my PhD committee, Dr. Lawrence Overzet, Dr. William Frensley, Dr. Mark Lee and Dr. Vandenberghe. I value and cherish the friendship of my co-workers, Dr. Gautam Gaddemane and Shanmeng Chen, and my friends. Most importantly, I would like to thank my parents and my family for being there for me through thick and thin and without whom I would not be here. Last, but definitely not the least, I am grateful to that one person who has changed my life completely ever since she became a part of it. She has held my hand firmly through all the hurdles and, at the same time, kept me grounded through her constructive criticisms. Thank you Vrunda Desai for being my partner-in-crime.

April 2019

THEORETICAL STUDY OF QUANTUM TRANSPORT IN REALISTIC SEMICONDUCTOR DEVICES

Pratik B. Vyas, PhD
The University of Texas at Dallas, 2019

Supervising Professor: Dr. Massimo V. Fischetti, Chair

Semiconductor devices have transformed the world through tremendous technological advances in all aspects of life imaginable. An important aspect of the research into improving these devices is computer-aided simulation and modeling of their electrical behavior. The ability to study theoretically semiconductor devices allows us to predict their behavior as well as optimize their performance before having to physically fabricate the device, saving us money and time. To this end, we have developed a novel approach, based on the effective mass approximation, to study theoretically quantum transport, both ballistic and dissipative, in realistic semiconductor devices. Our model takes into account quantum confinement and other non-local quantum effects affecting electronic transport in the current and near-future generations of transistors. As an example of application, we have studied the electrical behavior of well-known silicon field-effect transistors (FETs) and the factors affecting their performance.

TABLE OF CONTENTS

ACKNOWLEDGMENTS v

ABSTRACT vi

LIST OF FIGURES x

CHAPTER 1 INTRODUCTION 1

- 1.1 Scaling of CMOS Technology 1
- 1.2 Simulation of Quantum Transport 1
- 1.3 Outline of Dissertation 2

CHAPTER 2 EQUILIBRIUM ANALYSIS OF A SEMICONDUCTOR DEVICE . . 4

- 2.1 2-D Schrödinger Equation with Dirichlet Boundary Conditions 5
- 2.2 Channel Oriented in the [110] Direction 8
- 2.3 Schrödinger Equation with Neumann Boundary Condition 10
  - 2.3.1 Electrons 11
  - 2.3.2 Holes 12
- 2.4 Solution of the Poisson Equation 12
- 2.5 Self-consistent Solution of the Schrödinger and Poisson Equations with *Closed* Boundary Conditions: State of the System at Equilibrium 13
  - 2.5.1 Conventional Newton's iteration scheme: Building a semiclassical Jacobian 14

CHAPTER 3 BALLISTIC SIMULATION OF DEVICE BEHAVIOR UNDER APPLIED BIAS: NON EQUILIBRIUM ANALYSIS 17

- 3.1 2-D Schrödinger Equation with *Open* Boundary Conditions: Quantum Transmitting Boundary Method 17
  - 3.1.1 QTBM boundary conditions 19
- 3.2 Calculation of Device Characteristics 24
  - 3.2.1 Transmission coefficient 24
  - 3.2.2 Local density of states 26
  - 3.2.3 Drain current 26
- 3.3 Self-consistent *Open* System 29
  - 3.3.1 *Open* system Schrödinger equation with Dirichlet leads 30

3.3.2 *Open* boundary system with Neumann leads 36

CHAPTER 4 THEORETICAL SIMULATION OF NEGATIVE DIFFERENTIAL TRANSCONDUCTANCE IN LATERAL QUANTUM WELL NMOS DEVICES 38

4.1 Preface 39

4.2 Introduction 40

4.3 Device Description and Experimental Observations 41

4.4 Theoretical Formulation 42

4.5 Simulation Results and Discussion 45

4.6 Silicon-on-Insulator QW CMOS 53

CHAPTER 5 SIMULATION OF QUANTUM CURRENT IN DOUBLE GATE MOSFETS: VORTICES IN ELECTRON TRANSPORT 55

5.1 Preface 56

5.2 Introduction 56

5.3 Device Description 58

5.4 Simulation Results 59

CHAPTER 6 QUANTUM MECHANICAL STUDY OF IMPACT OF SURFACE ROUGHNESS ON ELECTRON TRANSPORT IN ULTRA- THIN BODY SILICON FETS 65

6.1 Preface 66

6.2 Introduction 67

6.3 Device Description and Simulation of Interface Roughness 68

6.4 Simulation Results 70

CHAPTER 7 DISSIPATIVE QUANTUM TRANSPORT: MASTER EQUATION APPROACH 74

7.1 Introduction 74

7.1.1 Scattering between electronic states: Solution of Pauli master equation 75

7.2 Dissipative Transport in UTB FETs 85

7.2.1 Simulation of different scattering phenomena 86

7.2.2 Transport Characteristics 89

CHAPTER 8 CONCLUSION 95
REFERENCES 97
BIOGRAPHICAL SKETCH 105
CURRICULUM VITAE

LIST OF FIGURES

2.1 A simplified diagram showing how Dirichlet and Neumann boundary conditions behave at the device-contact interface. 4

3.1 Schematic illustration showing the implementation of the Quantum Transmitting Boundary Method. 17

4.1 Schematic cross-section of the lateral QW nMOSFET [1]. 42

4.2 Magnitude of the net doping profile of the device with a nominal gate length of 40 nm. The white region at the top represents the gate oxide [1]. 43

4.3 Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics at 46 K (black triangles) and 10 K (cyan circles) for the 10 nm device with $V_{\mathrm{DS}} = 10$ mV [1]. 45

4.4 Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics for the 20 nm device at 10 K. $V_{\mathrm{DS}} = 1$ mV [1]. 46

4.5 Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics at 10 K (cyan circle) and 300 K (black triangle) for the 40 nm device with $V_{\mathrm{DS}} = 10$ mV. Negligible current is seen at 10 K. Thermionic emission is therefore the major cause of the vastly larger current observed at 300 K [1]. 46

4.6 Average LDOS in the channel for 10 nm, 20 nm, 30 nm, and 40 nm devices (from left to right) at 10 K. The cyan colored lines represent the potential-energy profile at the semiconductor/gate-insulator interface in each device. The energies are measured with respect to the Fermi energy in the source contact. The bias conditions are $V_{\mathrm{GS}} = 1.18$ V, 1.93 V, 2.2 V, and 2.2 V, respectively (left to right) [1]. 47

4.7 Transmission coefficient *vs.* injection energy for a particular traveling mode in the 10 nm, 20 nm , 30 nm and 40 nm devices, respectively. The traveling mode energies E_m^r are, from left to right, -0.83 eV, -0.49 eV, -0.26 eV and -0.04 eV. These energies are chosen since they exhibit the best resonant behavior in the respective devices (wherever applicable). $V_{\mathrm{GS}} = 1.18$ V, 1.93 V, 2.2 V, and 2.2 V, respectively (left to right) [1]. 48

4.8 Magnitude of the modified net doping profile along the Si/gate-insulator interface of the 40 nm device compared to the original doping profile. The pSDE's have been broadened and their doping concentration has been reduced, to increase the conduction via resonant tunneling [1]. 48

4.9 Left: Average LDOS in the channel of the 40 nm device with the modified doping profile (shown in Fig. 4.8), at 10 K. The cyan colored line represents the potential-energy profile at the semiconductor/gate-insulator interface. Right: Transmission coefficient *vs.* total injection energy for the particular traveling mode that exhibits the highest resonant transmission. The energy E_m^r of the traveling mode is -0.16 eV and the gate bias $V_{\mathrm{GS}} = 2.0$ V [1]. 49

4.10 Magnitude of the net doping profile of the SOI QW nMOS device with a nominal gate length of 20 nm. The white region at the top represents the gate oxide. . . 53

4.11 Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics for the 20 nm SOI nMOS at 46 K. $V_{\mathrm{DS}} = 1$ mV 54

5.1 Net doping profile of the 10 nm UTB DG nMOS that we have studied. The white regions at the top and bottom represent the 1 nm thick gate oxide while the grey patches are used to highlight the position of the gate terminals. 58

5.2 I_{DS}-V_{GS} characteristics of a 10 nm UTB DG nMOS at 10 K and 300 K. V_{GS} is measured with respect to the flat band voltage of the device. V_{DS} =10 mV. . . . 59

5.3 Total charge distribution in a 10 nm UTB DG nMOS at 300 K. (a) The dark red regions show the creation of two separate inversion channels deep inside saturation, whereas in (b) volume inversion is seen in the linear region of operation. V_{GS}=0.6 V, V_{DS} =10 mV. 59

5.4 Transmission coefficient T *vs.* injection energy (a) for current injection from the source contact in a 10 nm UTD DG nMOS at 300 K. The different colored lines correspond to the different injected subbands. The energies are measured with respect to source Fermi level. The LDoS distribution averaged over a cross-sectional thickness of roughly 1.3 nm in the top channel (b) and middle of device (c) at 300 K for injection from the source contact. V_{GS} =1.6 V, V_{DS} =10 mV. . . 60

5.5 (a) Current density distribution in a 10 nm UTB DG nMOS at 300 K. (b) Current density distribution resolved for a single injection energy at 300 K. The energy is chosen to have the value at which a sharp dip in transmission occurs due to destructive interference at resonance. V_{GS} =0.6 V, V_{DS} =10 mV. (c) Current density distribution in the same device at 10 K. V_{GS} =0.6 V, V_{DS} =10 mV. The red arrows highlight the direction of the vortices. The plots are stretched to match the aspect ratio of the device. 61

5.6 Antiresonance features in the transmission spectra broaden in the presence of an asymmetric gate bias. The potential difference between the gates is 0.4 V. Temperature = 300 K, V_{DS} =10 mV. 62

5.7 Current density distribution in the device with asymmetric gate bias at 300 K. The red arrows highlight the direction of vortices in the current density. Potential difference between gates is 0.4 V for the latter case. 63

5.8 Comparison of the transfer characteristics of the device with equal (symmetric) and unequal (asymmetric) gate biases applied at 300 K. Potential difference between gates is 0.4 V for the latter case. 63

6.1 Doping profile of (a) DG and (b) SOI nMOS with surface roughness. The white region at the top and bottom represents the gate oxide while the grey patches are used to highlight position of the gate terminals. Green dashed lines are used to mark the oxide-semiconductor interface of the ideal devices. 69

6.2 Transfer characteristics of the simulated devices in (a) linear and (b) semi-log scale to highlight the impact of SR scattering in above-threshold and subthreshold regions, respectively. The dashed lines represent ideal device behavior while solid lines along with error-bars are used to plot statistical mean and standard deviation of the behavior of 'roughened' samples, respectively. 71

6.3 Current density distribution in (a) 4 nm and (b) 8 nm DG nMOS at 300 K. The plots show that SR scattering mostly affects the electron transport close to the oxide-semiconductor interface. A vortex in the current density is highlighted by a red curved arrow. V_{GS} =0.95 V, V_{DS} =100 mV. 72

7.1 Net doping profile of the simulated UTB DG nMOS. The white regions at the top and bottom represent the 1.2 nm-thick gate oxide, while the grey patches are used to highlight the position of the gate contacts. 85

7.2 Simulated potential-energy distribution in the 5 nm UTB DG nMOS plotted for reference. The potential energy is measured with respect to the source fermi level. 86

7.3 Scattering rates as a function of total electron-energy simulated for the different scattering mechanisms. V_{GS} =0.2 V, V_{DS} = 100 mV. 87

7.4 Diagonal elements of the density-matrix, as a function of total electron-energy, representing the final occupation of states in presence of (a) electron-phonon scattering, and (b) SR scattering. The solid and dashed red lines represent the ballistic occupation of states injected from the source and the drain, respectively. 88

7.5 LDOS distribution in (a) ballistic case, (b) in presence of phonon (optical and acoustic) scattering, and (c) in presence of scattering with surface roughness. The LDOS is averaged over a cross-sectional thickness of roughly 1 nm along the middle of the device. The red dashed line represents the potential distribution cut across the device mid-section. The energies are measured with respect to the source Fermi level. V_{GS} =0.2 V, V_{DS} = 100 mV. 90

7.6 (a) Spatial distribution of the electron kinetic energy in presence of SR scattering. Kinetic-energy distributions for the ballistic and electron-phonon are largely similar to the above, except in the channel. The latter is highlighted in (c) which shows the kinetic energy averaged over the cross-sectional thickness along the length of the device for the different scattering mechanisms. 91

7.7 Spatial distribution of the electron drift velocity in the presence of (a) electron-phonon scattering and (b) SR scattering. (c) Comparative plot of the drift velocity, averaged over the cross-sectional thickness along the length of the device, in presence of different scattering mechanisms. 92

7.8 Transfer characteristics of the simulated 5 nm UTB DG nMOS in (a) linear and (b) semi-log scale focussing on the above-threshold and subthreshold regions, respectively. 94

CHAPTER 1

INTRODUCTION

1.1 Scaling of CMOS Technology

The world has benefitted immeasurably from the development and use of better, faster, smarter, and more compact electronic devices ranging from the everyday-life gadgets, like the cell phone and laptops, to much more sophisticated tools used in space navigations and military operations. With all advancements come the challenges to maintain steady progress. Research into improving the structure and properties of the individual CMOS transistors is essential for the continued growth of the semiconductor industry. Aggressive scaling of the Complementary Metal Oxide Semiconductor (CMOS) technology into the nanoscale era is a stark evidence of the progress made in this field. As part of this trend, new device structures have been proposed and implemented, e.g., silicon-on-insulator (SOI) field effect transistors (FET) [2], ultra thin body (UTB) double gate (DG) FETs [3], tunnel FETs [4], graphene nano ribbon devices [5], to name a few.

1.2 Simulation of Quantum Transport

Theoretical analysis of semiconductor devices has traditionally been done through simulation of electron transport by semiclassical methods based on the Boltzmann Transport Equation (BTE) [6, 7, 8] or moments of the BTE, like the drift-diffusion model [9] or the hydrodynamics model [10]. However, the current and emerging generation of transistors have active regions of length comparable to the wavelength of an electron. At such small dimensions, the basic assumption of the BTE, that electrons are point-like objects with defined position and momentum, falls short. A full quantum study of electron transport is needed to simulate and understand the behavior of these devices. Explicit quantum mechanical effects like tunneling, quantization and resonance cannot be fully captured using semiclassical means. Moreover, the International Technology for Semiconductors (ITRS) [11] has set strong end-of-roadmap goals for RF and analog/mixed signal devices and circuit performance, well beyond the current state-of-the-art capabilities. Achieving these very aggressive targets through semiclassical devices seems to be a daunting task. However, it has been seen and documented that incorporating explicit quantum mechanical transport into a device's operation has led to transformative capabilities [12, 13, 14]. Thus, a possible industrially-scalable and cost-effective route to achieve the goals outlined by the ITRS would be to integrate explicit quantum transport into industrial Si CMOS technology. Modeling the behavior of such devices will also require a full quantum transport treatment.

The three well-known methods to study quantum transport are the Wigner function method [15, 16, 17], the Non-equilibrium Green's function method [18, 19, 20] and the self-consistent solution of the Schrödinger, Pauli master [21, 22] and Poisson equations. In our research we will be using the latter to model ballistic and dissipative electron transport. The model we employ uses the single-electron envelope approximation and the conduction bands are approximated with parabolic ellipsoidal or spherical valleys (six ellipsoidal valleys in the case of silicon). This model best suits our research interest in the present and near future generation of technologically significant semiconductor devices with channel lengths between 5 nm to 20 nm. Devices with longer gate or channel lengths can be modeled accurately by less expensive semiclassical models like the Monte Carlo method [6, 7] or drift-diffusion approach [9] while devices with smaller than 5 nm channel lengths will require a more complicated and computationally expensive full band quantum transport treatment [23, 24].

1.3 Outline of Dissertation

The dissertation is organized in a step-by-step fashion, each step representing an improvement over the previous, bringing our theoretical model closer to reality. We start with the simplest case — a purely ballistic description of transport. Our main objective is to solve a system that can communicate with its environment through its contacts, namely an *open* system. This requires discretizing the continuum of energies in the contacts. The optimal discretization is obtained by sampling the density-of-states using the solution of the system in equilibrium, namely a *closed* system. Chapter 2 details our theoretical scheme of obtaining these solutions of the *closed* system. We solve the Schrödinger equation with *closed* boundary conditions self-consistently with the Poisson equation in the two-dimensional (2-D) plane of the device to achieve this. The device is assumed wide enough so that a 3-D simulation is not required. In Chapter 3, our simulation methodology is extended to model an *open* system driven far from equilibrium under the influence of a drain-to-source bias. The Quantum Transmitting Boundary method (QTBM) [25, 26] is used to enforce *open* boundary conditions to model particle exchange, through current-carrying leads, between the device and environment. Furthermore, ways to derive various device characteristics of practical importance are highlighted here. Chapters 4 and 5 present two interesting applications of our ballistic simulation tool, where we use the same to model the behavior of quantum-well (QW). In addition, dissipation in the form of scattering of electrons with surface roughness is introduced in a statistical fashion in Chapter 6. Here the electron transport is still modeled in the ballistic regime, while dissipation due to surface-roughness (SR) scattering arises as a result of averaging over statistical ensembles of sample devices. Finally, in Chapter 7, dissipation is introduced explicitly as a perturbation away from the ballistic picture by solving the Pauli master equation self-consistently with the Schrödinger and Poisson equations to determine the final state of the device. In this work, we have considered only electron-phonon and SR scattering only, but the model can potentially be extended to include other scattering phenomena as well. As an application, the impact of the mentioned scattering mechanisms is presented for the case of UTB double-gate (DG) FETs. Reinstating the purpose and significance of the entire work, concluding remarks are summarized in Chapter 8.

CHAPTER 2

EQUILIBRIUM ANALYSIS OF A SEMICONDUCTOR DEVICE

The equilibrium level condition of the system is obtained by solving the two-dimensional (2-D) Schrödinger and Poisson equations self consistently [27] for different gate biases and at different temperatures, with no drain-to-source voltage applied. For very wide devices, we can assume translational invariance of the device profile along the out-of-plane y direction and therefore our 2-D approach is a good approximation. We define this *closed* state of the system as one in which the system does not exchange particles (and energy, thermodynamically) with the rest of the 'universe' through its contacts. Therefore, in its quantum description, either the wavefunctions must vanish outside the system or there should be zero net probability current entering or exiting the system. In order to obtain the correct *closed* system solution, it is better to solve the Schrödinger equation with both Dirichlet and Neumann boundary conditions, so that one gets the right density near the contacts and, so, the correct potential. The Dirichlet and Neumann solutions behave 'sine-like' and 'cosine-like' at the device-contact interface, respectively, as Fig. 2.1 shows. Thereby, a combined set of these two forms a complete set of physical solutions and any other solution to the *closed* system can be represented as a linear combination of the two. Moreover, another vital application

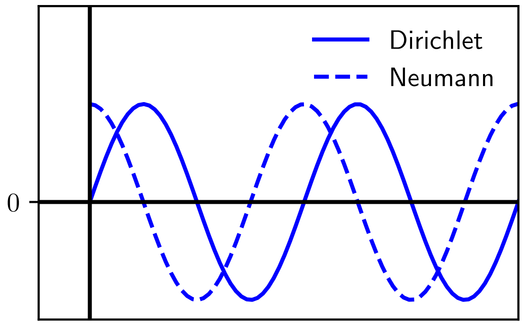

Figure 2.1: A simplified diagram showing how Dirichlet and Neumann boundary conditions behave at the device-contact interface.

of these two solutions of the *closed* system will be highlighted in Sec. 3.3 where we use them to sample the continuous energy spectrum of the *open* system. In all the cases, zero-value Dirichlet boundary conditions will be applied at the oxide-semiconductor interface as well as at the substrate of the device to account for the idealized yet physically consistent situation of having zero charges beyond the said interfaces.

Next, we describe how to obtain a self-consistent 2-D solution of the Schrödinger equation, considering both Dirichlet (Sec. 2.1) and Neumann boundary conditions (Sec. 2.3), and Poisson equations. When solving the Poisson equation, in order to account for electrostatic control through gates using externally applied bias, Dirichlet boundary conditions are applied at the gate-dielectric interfaces, while Neumann conditions are imposed elsewhere to ensure confinement of the electric field to the simulation domain.

We will assign the transport, confinement and out-of-plane directions to the x, z and y directions, respectively.

2.1 2-D Schrödinger Equation with Dirichlet Boundary Conditions

The finite-difference method is used to solve numerically the time-independent single electron 'effective mass' Luttinger-Kohn (Schrödinger) equation [28] using the parabolic-band approximation to approximate the anisotropic ('elliposidal') electron dispersion close to the six minima of the conduction band of Si. Only electrons in the first conduction band are considered. Now, the Luttinger-Kohn equation can be written as:

$$[E_m(-i\nabla) + V(x,y,z)]\psi(x,y,z) = E\psi(x,y,z) \ , \tag{2.1}$$

where, using the effective mass approximation, $E_m(k) = \frac{\hbar^2 k^2}{2m^*}$ and m^* is the effective mass of band m. For the simplest case in which the silicon (Si) channel is oriented along the [100] direction, the two-dimensional Schrödinger equation takes the form [27, 29]:

$$-\frac{\hbar^2}{2}\left[\frac{1}{m^v_{\text{x}}}\frac{\partial^2 \xi^v(x,z)}{\partial x^2} + \frac{1}{m^v_{\text{z}}}\frac{\partial^2 \xi^v(x,z)}{\partial z^2}\right] + \ V(x,z)\xi^v(x,z) = E^v_{xz}\xi^v(x,z) \ , \tag{2.2}$$

where the envelope wavefunction is $\psi(x, y, z) = e^{\mathrm{i}k_{\mathrm{y}}y}\xi(x, z)$ and m_{x}, m_{y} and m_{z} are the transport mass (on the plane (x, z)-plane of the Si/gate-insulator interface and along the channel), the out-of-plane mass (along the y-direction perpendicular to the interface), and the quantization mass (on the plane of the interface, but along the z direction perpendicular to the transport direction). V is the potential energy on the (x, z) plane, E_{xz} is the energy of the two-dimensional wavefunction $\xi(x, z)$ and k_{y} represents the wavevector in the y direction. These quantities take different values for each of the six ellipsoidal valleys close to the X symmetry-point in the Si Brillouin zone and the superscript v denotes the valley index. All calculations presented in the following have to be repeated three times, once for each pair of inequivalent valley-orientations.

We denote by $E^v_{xz} + \hbar^2 k_{\mathrm{y}}^2/2m^v_{\mathrm{y}}$ the total energy of an electronic state described by the wavefunction $\psi(x, y, z)$. In its discretized form, Eq. (2.2) can be recast as:

$$a_1\left[\frac{\xi^v(x_{i+1}, z_i)}{\Delta x_i^2} + \frac{\xi^v(x_{i-1}, z_i)}{\Delta x_i \Delta x_{i-1}}\right] + a_2\left[\frac{\xi^v(x_i, z_{i+1})}{\Delta z_i^2} + \frac{\xi^v(x_i, z_{i-1})}{\Delta z_i \Delta z_{i-1}}\right] - \left[\frac{a_1}{\Delta x_i^2} + \frac{a_1}{\Delta x_i \Delta x_{i-1}} \right.$$
$$\left. + \frac{a_2}{\Delta z_i^2} + \frac{a_2}{\Delta z_i \Delta z_{i-1}} + V(x_i, z_i)\right]\xi^v(x_i, z_i) = E^v_{xz}\xi^v(x_i, z_i) \ , \tag{2.3}$$

where $a_1 = -\hbar^2/2m^v_{\mathrm{x}}$, $a_2 = -\hbar^2/2m^v_{\mathrm{z}}$ and, i represents each of the N discretization points. We employ a two-dimensional mesh given by the tensor product of two one-dimensional meshes consisting of N_{x} and N_{z} points, so that $N = N_{\mathrm{x}} \times N_{\mathrm{z}}$. The $N \times N$ Hamiltonian matrix ($\mathbf{H}^{\mathrm{D}}$) is created using the terms on the left-hand-side of the discretized equation. Assuming Dirichlet boundary conditions, the wavefunctions vanish at the edges of the simulated region (the 'device'). Therefore, Eq. (2.3) takes the form of an eigenvalue problem of rank N:

$$\left[\mathbf{H}^{\mathrm{D}} + \mathbf{V}\right] \cdot \boldsymbol{\xi}_\mu \ = \ E_\mu \boldsymbol{\xi}_\mu \ , \tag{2.4}$$

where $\boldsymbol{\xi}_\mu$ and E_μ are the μ^{th} eigenfunction and eigenvalue, respectively, and, $\mu = 1, 2, ..., N$. Incorporating non uniformity of the mesh, the Hamiltonian matrix ($\mathbf{H}^{\mathrm{D}}$) is given by:

$$\mathbf{H}^{\mathrm{D}} = \mathbf{H}_{\mathrm{x}}^{\mathrm{D}} + \mathbf{H}_{\mathrm{z}}^{\mathrm{D}} \ , \tag{2.5}$$

where $\mathbf{H}_{\mathrm{x}}^{\mathrm{D}} =$

$$\begin{bmatrix}
-2\mathrm{A}_{1,1}^{1,1} & \mathrm{A}_{1,2}^{1,1} & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\
\mathrm{A}_{2,1}^{2,1} & \substack{-\mathrm{A}_{2,2}^{2,2} \\ -\mathrm{A}_{2,1}^{2,2}} & \mathrm{A}_{2,2}^{2,3} & 0 & & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & \cdots & & & 0 \\
\vdots & & \ddots & \ddots & \ddots & \cdots & & \vdots \\
0 & \cdots & \mathrm{A}_{\mathrm{i,i-1}}^{\mathrm{i,i-1}} & \substack{-\mathrm{A}_{\mathrm{i,i}}^{\mathrm{i,i}} \\ -\mathrm{A}_{\mathrm{i,i-1}}^{\mathrm{i,i}}} & \mathrm{A}_{\mathrm{i,i}}^{\mathrm{i,i+1}} & 0 & \cdots & 0 \\
0 & \cdots & \cdots & \ddots & \ddots & \ddots & \cdots & 0 \\
0 & \cdots & \cdots & & & & \mathrm{A}_{\mathrm{N,N-1}}^{\mathrm{N,N-1}} & -2\mathrm{A}_{\mathrm{N,N}}^{\mathrm{N,N}}
\end{bmatrix} ,$$

and, $\mathbf{H}_{\mathrm{z}}^{\mathrm{D}} =$

$$\begin{bmatrix}
-2\mathrm{B}_{1,1}^{1,1} & 0 & \cdots & \mathrm{B}_{1,1}^{1,1+\mathrm{N_x}} & 0 & \cdots & & & \\
0 & -2\mathrm{B}_{2,2}^{2,2} & 0 & \cdots & \mathrm{B}_{2,2}^{2,2+\mathrm{N_x}} & 0 & \cdots & & \\
0 & \cdots & \ddots & & & \cdots & & & \\
\vdots & & \cdots & \ddots & & \cdots & & & \\
 & & \mathrm{B}_{\mathrm{i,i-N_x}}^{\mathrm{i,i-N_x}} & \cdots & \substack{-\mathrm{B}_{\mathrm{i,i}}^{\mathrm{i,i}} \\ -\mathrm{B}_{\mathrm{i,i-N_x}}} & 0 & \cdots & \mathrm{B}_{\mathrm{i,i}}^{\mathrm{i,i+N_x}} & \cdots \\
\vdots & \cdots & & & & \ddots & \cdots & & \\
0 & & \cdots & & & & & \mathrm{B}_{\mathrm{N,N-N_x}}^{\mathrm{N,N-N_x}} & \mathrm{B}_{\mathrm{N,N}}^{\mathrm{N,N}}
\end{bmatrix} .$$

$\mathrm{A}_{i,j}^{k,l} = \frac{a_1}{\Delta x_i \Delta x_j}$ and $\mathrm{B}_{i,j}^{k,l} = \frac{a_2}{\Delta z_i \Delta z_j}$ are the matrix elements in row k and column l. $\mathbf{H}^{\mathrm{D}}$ actually gives wavefunctions that vanish just outside the device domain. All elements, except those on the diagonal, of the rows representing the left and right edges of the device, i.e., rows having indexes jN_{x} or jN_{x} (j being an integer within range $(0, N_{\mathrm{y}})$), are set to zero. This ensures application of zero-value Dirichlet boundary condition exactly at the left and right edge of the device. **V** is a diagonal matrix containing the potential distribution, as shown below,

$$\mathbf{V} = \begin{bmatrix} V_1 & & & & & \\ & V_2 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & V_{N-1} & \\ & & & & & V_N \end{bmatrix} , \tag{2.6}$$

where V_i is the potential energy at mesh point i.

The six equivalent conduction band valleys of Si are two-fold degenerate. Therefore, the eigenvalue problem of Eq. (2.4) is calculated separately for each of the three pairs of valleys.

2.2 Channel Oriented in the [110] Direction

We take a look at the effect that the channel orientation has on the Luttinger-Kohn equation Eq. (2.1). For channels oriented along the [100] direction, Eq. (2.1) can be solved directly, using the known effective masses of silicon, separately three times for the pair of 6 equivalent Si valleys. Here, we specifically take the more complex case of device channel oriented along the [110] direction, following the general trend of the VLSI technology. The effective mass tensor will be a full 3×3 matrix, instead of a diagonal matrix. This will result in the presence of mixed second-order derivatives in the Schrödinger equation, as we see below.

Let (x', y', z') and (x, y, z) be the co-ordinate systems based on the [110] and [100] orientations, respectively. Based on simple geometry, we can write

$$x' = \frac{1}{\sqrt{2}}(x + y) \quad y' = \frac{1}{\sqrt{2}}(x - y) \quad z' = z \ . \tag{2.7}$$

$$\text{Re-writing, } x = \frac{1}{\sqrt{2}}(x' + y') \quad y = \frac{1}{\sqrt{2}}(x' - y') \quad z = z' \ .$$

$$\text{Now, } \frac{\partial \xi}{\partial x} = \frac{\partial \xi}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial \xi}{\partial y'}\frac{\partial y'}{\partial x} = \frac{1}{\sqrt{2}}(\frac{\partial \xi}{\partial x'} + \frac{\partial \xi}{\partial y'}) \ . \tag{2.8}$$

Rotating the co-ordinate system from (x, y, z) to (x', y', z') by incorporating the above relations in Eq. (2.1) we obtain

$$-\frac{\hbar^2}{2}\left\{\frac{1}{2}\left[\frac{1}{m^v_{\mathrm{x}}} + \frac{1}{m^v_{\mathrm{y}}}\right]\left[\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2}\right] + \frac{1}{m^v_{\mathrm{z}}}\frac{\partial^2}{\partial z'^2} + \left[\frac{1}{m_{\mathrm{x}}} - \frac{1}{m_{\mathrm{y}}}\right]\frac{\partial^2}{\partial x' \partial y'} + V(x', y')\right\}\psi^v(x', y', z) = E^v \psi^v(x', y', z') \ . \tag{2.9}$$

In order to remove the mixed second order derivatives, we express the solution as $\psi^v(x', y', z') = e^i k_{y'}(\alpha x' + y')\xi^v(x', z')$. Note here that we take into account the translational invariance of the potential in the y' direction by assuming a free electron solution in that direction. Substituting the guess solution in Eq. (2.9) we obtain,

$$-\frac{\hbar^2}{2}[\frac{1}{2}(\frac{1}{m^v_{\mathrm{x}}} + \frac{1}{m^v_{\mathrm{y}}})(-\alpha^2 k^2_{y'}\xi(x', z') + 2i\alpha k_{y'}\frac{\partial \xi^v(x', z')}{\partial x'} + \frac{\partial^2 \xi^v(x', z')}{\partial x'^2} - k^2_{y'}) + \frac{1}{m^v_{\mathrm{z}}}\frac{\partial^2 \xi^v(x', z')}{\partial z'^2} + (\frac{1}{m^v_{\mathrm{x}}} - \frac{1}{m^v_{\mathrm{y}}})(-\alpha k^2_{y'}\xi^v(x', z') + ik_{y'}\frac{\partial \xi^v(x', z')}{\partial x'}) + V(x', z')\xi^v(x', z')] = E^v \xi^v(x', z') \ . \tag{2.10}$$

Equating the coefficients of the first order derivatives w.r.t. x to 0,

$$i\alpha k_{y'}(\frac{1}{m^v_{\mathrm{x}}} + \frac{1}{m^v_{\mathrm{y}}}) + ik_{y'}(\frac{1}{m^v_{\mathrm{x}}} - \frac{1}{m^v_{\mathrm{y}}}) = 0 \ , \tag{2.11}$$

giving $\alpha = -\frac{m^v_{\mathrm{c}}}{2m^v_{\mathrm{xy}}}$, with $\frac{1}{m^v_{\mathrm{c}}} = \frac{1}{2}(\frac{1}{m^v_{\mathrm{x}}} + \frac{1}{m^v_{\mathrm{y}}})$ and $\frac{1}{m^v_{\mathrm{xy}}} = \frac{1}{m^v_{\mathrm{x}}} - \frac{1}{m^v_{\mathrm{y}}}$. Removing the 'prime' from the co-ordinate axes for convenience, the modified Schrödinger equation for the [110] direction then takes the form

$$-\frac{\hbar^2}{2}\left[\frac{1}{m^v_{\mathrm{c}}}\frac{\partial^2 \xi^v(x, z)}{\partial x^2} + \frac{1}{m^v_{\mathrm{z}}}\frac{\partial^2 \xi^v(x, z)}{\partial z^2}\right] + V(x, z)\xi^v(x, z) = E^v \xi^v(x, z) - \frac{\hbar^2 k^2_{\mathrm{y}}}{2}\left[\frac{1}{m^v_{\mathrm{c}}} - \frac{m^v_{\mathrm{c}}}{4(m^v_{\mathrm{xy}})^2}\right]\xi^v(x, z) \ , \tag{2.12}$$

where the full (envelope) wavefunction $\psi^v(x, y, z)$ is given as

$$\psi^v(x, y, z) = e^{\mathrm{i}k_\mathrm{y}y}\ e^{-\mathrm{i}\frac{m^v_\mathrm{c}}{2m^v_\mathrm{xy}}k_\mathrm{y}x}\xi^v(x, z)\ . \tag{2.13}$$

We calculate the electron wavefunctions $\xi_\mu(x, z)$ (labeled by the index μ) and the corresponding eigenvalues E^μ_{xz} by solving this eigenvalue problem. The second factor at the right-hand-side of Eq. (2.13) is incorporated into calculations at a later stage.

The more complicated form of the Schrödinger equation, Eq. (2.12), must be employed to treat only four of the six valleys, since the in-plane rotation from the [100] to the [110] direction does not affect the two ellipsoids with $m_\mathrm{z} = m_\mathrm{L}$. Here $m_\mathrm{L} = 0.91m_0$ (m_0 is the mass of an electron) is the longitudinal effective mass (we also assume $m_\mathrm{T} = 0.19m_0$ for the transverse effective mass). These valleys can be treated using the slightly simpler form given by Eq. (2.3). These calculations, shown here for Si, can be extended for any other material with similar orientations of their conduction band minima.

2.3 Schrödinger Equation with Neumann Boundary Condition

The same finite-differences form, given by Eq. (2.3), is used to solve Eq. (2.2) (and/or Eq. (2.12)) with Neumann boundary conditions at the left and right contacts (Dirichlet boundary condition for the top and bottom sides of the domain). The Hamiltonian matrix $\mathbf{H}^\mathrm{N}$ build to solve Eq. (2.4) for this case is very similar to $\mathbf{H}^\mathrm{D}$. The only difference is the rows representing the left and right contact surfaces, i.e., rows having indexes jN_x or jN_x (j being integer in the range $(0, N_\mathrm{y})$), will have $2\mathrm{A}^\mathrm{i,i+1}_\mathrm{i,i}$ and $2\mathrm{A}^\mathrm{i,i-1}_\mathrm{i,i-1}$ terms, respectively. This ensures application of Neumann boundary condition right at the drain and source edges.

$$\mathbf{H}^\mathrm{N} = \mathbf{H}^\mathrm{N}_\mathrm{x} + \mathbf{H}^\mathrm{N}_\mathrm{z}\ , \tag{2.14}$$

where $\mathbf{H}_\mathrm{z}^\mathrm{N} = \mathbf{H}_\mathrm{z}^\mathrm{D}$ and

$$\mathbf{H}_\mathrm{x}^\mathrm{N} = \begin{bmatrix} -\mathrm{A}_{1,1}^{1,1} & 2\mathrm{A}_{1,2}^{1,1} & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \mathrm{A}_{2,1}^{2,1} & \substack{-\mathrm{A}_{2,2}^{2,2} \\ -\mathrm{A}_{2,1}^{2,2}} & \mathrm{A}_{2,2}^{2,3} & 0 & & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \cdots & & & 0 \\ \vdots & & \ddots & \ddots & \ddots & \cdots & & \vdots \\ 0 & \cdots & \mathrm{A}_{\mathrm{i,i-1}}^{\mathrm{i,i-1}} & \substack{-\mathrm{A}_{\mathrm{i,i}}^{\mathrm{i,i}} \\ -\mathrm{A}_{\mathrm{i,i-1}}^{\mathrm{i,i}}} & \mathrm{A}_{\mathrm{i,i}}^{\mathrm{i,i+1}} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \ddots & \ddots & \ddots & \cdots & 0 \\ 0 & \cdots & \cdots & & & & 2\mathrm{A}_{\mathrm{N,N-1}}^{\mathrm{N,N-1}} & -\mathrm{A}_{\mathrm{N,N}}^{\mathrm{N,N}} \end{bmatrix} .$$

The terms A and B have their usual meaning, defined in Eq. (2.5).

2.3.1 Electrons

The electron wavefunctions obtained from the solution of the closed-system are used to calculate the electron charge distribution in the device. The information regarding the DoS along the x and z directions is already contained in these quantized 2-D wavefunctions. We have only to use the 1-D DoS representing the continuous energy spectrum along the 'homogenous' out-of-plane y direction given by the expression

$$D_{1D}^{v}(E_\mathrm{y}) = \frac{1}{\pi\hbar}\sqrt{\frac{m_\mathrm{y}^{v}}{2E_\mathrm{y}}} \; . \tag{2.15}$$

The expression for the electron charge distribution is

$$n(x,z) = \sum_{v}\sum_{\mu}\int_{E_{xz}^{\mu,v}}^{\infty} \mathrm{d}E_\mathrm{y} D_{1D}^{v}(E_\mathrm{y}) f(E_\mathrm{y}, E_\mathrm{F}) |\xi_\mu^{v}(x,z)|^2 \; , \tag{2.16}$$

where ξ_μ and E_μ are the wavefunction and energy of the quantum state μ, respectively. Using Eq. (2.15), the above expression becomes

$$n(x,z) = \sum_{v}\sum_{\mu}\frac{1}{\pi\hbar}\sqrt{\frac{m_\mathrm{y}^{v} k_\mathrm{B} T}{2}} F_{-\frac{1}{2}}\left(\frac{E_\mathrm{F} - E_{xz}^{\mu,v}}{k_\mathrm{B} T}\right) |\xi_\mu^{v}(x,z)|^2 \; , \tag{2.17}$$

Here $F_{-\frac{1}{2}}$ is the Fermi-Dirac integral of order -1/2, defined by Eq. (2.19).

2.3.2 Holes

The hole charge distribution, $p(x, y, z)$, is calculated semiclassically using the well known three-dimensional DoS expression,

$$p(x, y, z) = \frac{1}{2\sqrt{\pi}} \left(\frac{2m_{\mathrm{h}} k_{\mathrm{B}} T}{\pi \hbar^2}\right)^{\frac{3}{2}} F_{\frac{1}{2}} \left(\frac{V(x, z) - (E_{\mathrm{F}} + E_{\mathrm{g}})}{k_{\mathrm{B}} T}\right) , \tag{2.18}$$

where $m_{\mathrm{h}} = 0.8m_0$ is the hole effective mass, the E_{g} is the band gap energy of silicon, and, $F_{\frac{1}{2}}$ is the Fermi-Dirac integral of order 1/2, defined by Eq. (2.19).

We define the Fermi-Dirac integral of order σ as:

$$F_\sigma(\eta) = \int_0^\infty \frac{\varepsilon^\sigma \mathrm{d}\varepsilon}{1 + \exp(\varepsilon - \eta)} . \tag{2.19}$$

Eq. (2.19) is computed using the Gauss-Legendre quadrature method [30], as shown below for a function $g(y)$

$$\int_a^b \mathrm{d}y \frac{g(y)}{\sqrt{b - y}} = \sqrt{b - a} \sum_{i=1}^{n} w_i g(y_i) , \tag{2.20}$$

where $y_i = a + (b - a)x_i$, abscissas $x_i = 1 - \zeta_i^2$ and $w_i = 2w_i^{(2n)}$. ζ_i is the ith positive root of Legendre polynomial $P_{2n}(x)$ and $w_i^{(2n)}$ are the Gaussian weights of order $2n$. Note that the inclusion of temperature as an input parameter in the device simulation is achieved through the temperature dependence of Eqs. (2.17) and (2.18).

2.4 Solution of the Poisson Equation

Here we briefly describe the numerical solution of Poisson equation which constitutes an important part of most transport simulations. We solve the Poisson equation in the 2-D plane (x, z) of the device. To incorporate correctly the electrostatic effects caused by the change in permittivity along the dielectric-semiconductor interface, we solve the generalized Poisson equation:

$$\nabla . \left[\epsilon(x, z) \nabla V(x, z)\right] = \mathrm{e}^2 \left[p(x, z) - n(x, z) + N_{\mathrm{A}}(x, z) - N_{\mathrm{D}}(x, z)\right] , \tag{2.21}$$

where e is the electron charge and $\epsilon_{\mathrm{Si}}(x, z)$ is the permittivity of the material present at (x, z). The centered finite-differences method is used to solve this linear system problem :

$$\mathbf{P} \cdot \mathbf{V} \; = \; \mathbf{D} \; , \tag{2.22}$$

where $\mathbf{P}$ is an $N \times N$ matrix expressing the differential operators on the left hand side of Eq. (2.22), $\mathbf{V}$ and $\mathbf{D}$ are $N \times 1$ matrices expressing the potential energy distribution and the charge terms on the right hand side of Eq. (2.21), respectively, for each of the N mesh points. Neumann boundary conditions are employed along all the domain boundaries, excluding the portion of the domain edge representing the oxide-metal gate interface. Here, the gate potential V_{GS} is applied by setting Dirichlet boundary conditions. Mathematically this is done by assigning zeros to all the elements of the rows representing the said portion of the oxide-metal interface, except for the diagonal elements where unity is assigned. The gate potential in electron volts is added to the corresponding rows of the charge matrix $\mathbf{D}$. This assignment might lead to singularity issues while solving the linear system Eq. (2.22) computationally. The solution consists in multiplying the unity diagonal terms on $\mathbf{P}$ and the corresponding gate potential terms on $\mathbf{D}$ by a factor that has the same order of magnitude as the other non-zero terms on $\mathbf{P}$.

2.5 Self-consistent Solution of the Schrödinger and Poisson Equations with *Closed* Boundary Conditions: State of the System at Equilibrium

The electrostatic potential in the device is obtained by solving self-consistently the Schrödinger and Poisson equations, described in Sec. 2.1 and Sec. 2.4, respectively. The Fermi level of the device is first fixed at a value that results in charge neutrality deep in the substrate of the device. An initial 'guess' for the electrostatic potential is made that satisfies the condition of charge neutrality at each point in the device, using the three-dimensional density-of-states for electrons and holes. Proper selection of the initial state of the system is crucial as it is vital to attaining a good convergence and the previously mentioned 'guess' serves the purpose for our case. The Schrödinger equation is then solved numerically using this initial guess for the potential and the electron wavefunctions $\xi_\mu(x,z)$ and the corresponding energies, E_{xz}^{μ}, are calculated. The electron charge distribution is then calculated from these wavefunctions and the hole charge distribution using the classical expression Eq. (2.18). Thereafter, the Poisson equation is solved numerically, using the electron, hole, and doping charge densities, to generate a 'new' potential. The root-mean-square ('infinity-norm') error between the 'new' and 'old' potentials is then determined. If the error is greater than a predefined minimum value, typically of the order of 10^{-7} eV, then the procedure is repeated using the 'new' potential. Otherwise, the iterative procedure ends and the 'new' potential is the equilibrium electrostatic potential in the device.

This 'direct' self-consistent scheme seldom converges in practice. In order to accelerate the convergence of the iterative procedure, Newton's method [31] is used. In this method, the Poisson equation is not solved directly to generate the 'new' potential, instead a Jacobian matrix **J** is used to achieve convergence.

2.5.1 Conventional Newton's iteration scheme: Building a semiclassical Jacobian

In order to understand the role of Newton's method (also known as Newton-Raphson method) in achieving a faster rate of convergence, we need to first understand on which ideas the method is based. Basically, the method is a root-finding algorithm that is used iteratively to find better approximations to the roots of a function or a system of functions that are continuously differentiable. Let the vector $\mathbf{x}$ represent the set of variables $(x_1, x_2, ..., x_\mathrm{n})$ and the system of n functions $(f_1, f_2, ...f_n)$ is given by the vector $\mathbf{F}(\mathbf{x})$. Newton's method then states that the ith approximation $\mathbf{x}^i$ to the solution of the equation $\mathbf{F}(\mathbf{x}) = 0$ is,

$$\mathbf{x}^i = \mathbf{x}^{i-1} - \mathbf{J}^{-1}(\mathbf{x}^{i-1})\mathbf{F}(\mathbf{x}^{i-1}) \;, \tag{2.23}$$

where $\mathbf{J}(\mathbf{x}^i)$ is an $n \times n$ Jacobian matrix whose elements are the partial derivatives of the components of $\mathbf{F}(\mathbf{x})$ evaluated at $\mathbf{x}^i$, as shown

$$\mathbf{J}(\mathbf{x}^i) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x}^i)}{\partial x_1} & \frac{\partial f_2(\mathbf{x}^i)}{\partial x_1} & \cdots & \cdots & \frac{\partial f_\mathrm{n}(\mathbf{x}^i)}{\partial x_1} \\ \frac{\partial f_1(\mathbf{x}^i)}{\partial x_2} & \frac{\partial f_2(\mathbf{x}^i)}{\partial x_2} & \cdots & \cdots & \frac{\partial f_\mathrm{n}(\mathbf{x}^i)}{\partial x_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_1(\mathbf{x}^i)}{\partial x_\mathrm{n}} & \frac{\partial f_2(\mathbf{x}^i)}{\partial x_\mathrm{n}} & \cdots & \cdots & \frac{\partial f_\mathrm{n}(\mathbf{x}^i)}{\partial x_\mathrm{n}} \end{bmatrix} . \tag{2.24}$$

In our self-consistent scheme, instead of directly solving Eq. (2.21) to obtain the new potential (root), Newton's method is used to find the first approximation to the root of the Poisson equation, written as $\mathbf{P} \cdot \mathbf{V} - \mathbf{D} = 0$. The Jacobian $\mathbf{J}$ becomes a $N \times N$ diagonal matrix, with the diagonal terms being the first order derivatives w.r.t potential energy of the total charge, calculated semiclassically. The use of a semiclassical charge distribution is convenient since it can not only be determined using an analytical expression, but also its non-local nature is important in achieving fast convergence. The expression of the diagonal terms of $\mathbf{J}$ for a potential energy distribution $\mathbf{V}$ is,

$$\mathbf{J}^{i,i}(V^i) = -\frac{e^2}{\epsilon_{\mathrm{Si}}}\left(\frac{D_{\mathrm{sc}}(V^i + \Delta V) - D_{\mathrm{sc}}(V^i)}{\Delta V}\right) \;, \tag{2.25}$$

where $i = 1, 2, ..., N$ represents the N mesh points and V^i the potential energy at mesh point i. The term $D_{\mathrm{sc}}(V^i)$ is the total semiclassical electron and hole charge density at mesh point i given by

$$\begin{aligned} D_{\mathrm{sc}}(V^i) = \frac{1}{2\sqrt{\pi}}\left(\frac{2m_\mathrm{h}k_\mathrm{B}T}{\pi\hbar^2}\right)^{\frac{3}{2}} F_{\frac{1}{2}}\left(\frac{V(x,z) - (E_\mathrm{F} + E_\mathrm{g})}{k_\mathrm{B}T}\right) \\ + \sum_v \frac{1}{2\sqrt{\pi}}\left(\frac{2m_\mathrm{d}^v k_\mathrm{B}T}{\pi\hbar^2}\right)^{\frac{3}{2}} F_{\frac{1}{2}}\left(\frac{E_\mathrm{F} - V(x,z)}{k_\mathrm{B}T}\right) \;, \end{aligned} \tag{2.26}$$

where $m_{\mathrm{d}}^{v} = (m_x^v m_y^v m_z^v)^{1/3}$. The 'new' potential is obtained from:

$$\mathbf{V}_{\mathrm{new}} \;=\; \mathbf{V}_{\mathrm{old}} - \mathbf{J}^{-1}(\mathbf{V}_{\mathrm{old}})(\ \mathbf{P}\cdot\mathbf{V}_{\mathrm{old}} - \mathbf{D})\ . \tag{2.27}$$

Thus the 'new' potential contains an explicit dependence on the 'old' potential, instead of just depending on the charge distribution, which is the case when Eq. (2.21) is directly used to get the same. This procedure bears resemblance to the 'under-relaxation' approach. In this convergence scheme, again, the exact solution is not used in the subsequent self-consistent iterations. A combination of the current and previous solutions is employed instead,

$$\mathbf{V}^{i} = \lambda\mathbf{V}^{i} + (1-\lambda)\mathbf{V}^{i-1}\ , \tag{2.28}$$

where λ is the under-relaxation parameter lying in the interval $(0,1)$ and i is the iteration index. For most of our quantum simulations, Newton's method has been found to be much more successful in attaining convergence, compared to the relatively simpler under-relaxation approach. It is important to mention here that inverting the sparse matrix $\mathbf{J}$ is an expensive procedure and results in a dense matrix, which is not efficient to work with. Instead of doing this, we can see that the term $\mathbf{J}^{-1}(\mathbf{V})(\ \mathbf{P}\cdot\mathbf{V}-\mathbf{D})$ on the right hand side of Eq. (2.27), let us call it $\mathbf{L}$, is actually the solution of the linear system $\mathbf{J}(\mathbf{V})\cdot\mathbf{L} = (\ \mathbf{P}\cdot\mathbf{V}-\mathbf{D})$. Thus, the vector $\mathbf{L}$ can be determined using any linear system solver and the process is fast, inexpensive and does not burden the computer memory.

The self-consistent simulation typically achieves the desired degree of convergence in 25-30 iterations. On a quad core Linux machine with Intel i7 CPU, this process requires about 3 minutes for a 4 nm × 50 nm double gate MOSFET device with 110 × 100 mesh points.

CHAPTER 3

BALLISTIC SIMULATION OF DEVICE BEHAVIOR UNDER APPLIED BIAS: NON EQUILIBRIUM ANALYSIS

In the closed boundaries system, the wavefunction is assumed to vanish outside the device, so there is no current flow through the device, i.e., represents an isolated system. However, we are interested in the 'open' system in which electrons can flow into and out of the device. This describes the behavior of a device in the presence of an applied drain-source voltage V_{DS}. The method we follow is the QTBM proposed by Lent and Kirkner [25].

3.1 2-D Schrödinger Equation with *Open* Boundary Conditions: Quantum Transmitting Boundary Method

The source and drain contacts are imagined as infinite leads going into the device, as illustrated schematically in Fig. 3.1. A separate coordinate system (ω_s, κ_s) is defined for each lead s, as shown in the figure. This is just a matter of convenience as the same derivations/expressions become applicable for all the leads. The potential is assumed to be constant along the direction ω_s and, along κ_s, the potential profile is assumed to be same

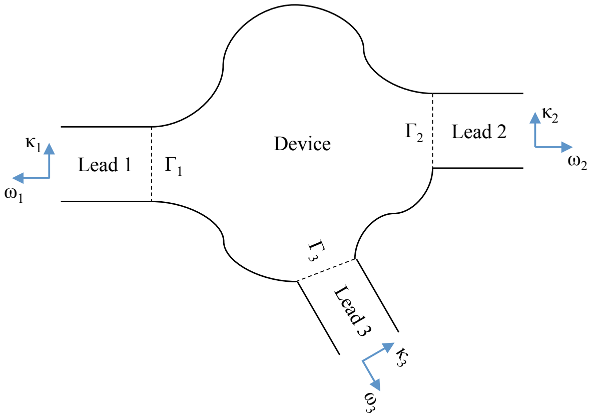

Figure 3.1: Schematic illustration showing the implementation of the Quantum Transmitting Boundary Method.

as that along the lead-device interface. Another assumption made is that outside the device as well as outside the lead edges, the wavefunction vanishes. Using these conditions, the wavefunctions in the leads can be separated into two independent components - traveling waves along ω_s and wavefunctions with a discretized energy spectrum along κ_s due to the confinement in that direction. The latter part is determined by solving the one-dimensional Schrödinger equation along κ_s,

$$-\frac{\hbar^2}{2m_\mathrm{z}^v}\frac{\partial^2\varphi_m^{r,v}(\kappa_s)}{\partial\kappa_s^2}+V_s(\kappa_s)\varphi_m^{s,v}(\kappa_s)=E_m^{s,v}\varphi_m^{s,v}(\kappa_s)\ , \tag{3.1}$$

where $V_s(\kappa_s)$ is the electron potential energy along κ_s in lead s and m_z is the quantization mass. Eq. (3.1) becomes an eigenvalue problem as a consequence of the zero-value Dirichlet boundary conditions imposed by the lead edges. φ_m^s and E_m^s represent the m_th eigenstate and eigen-energy, respectively, in lead s of the one-dimensional (1-D) eigenvalue problem which can be written as,

$$\mathbf{H}_\mathrm{1D}\cdot\boldsymbol{\varphi}_m^{s,v}=E_m^{s,v}\boldsymbol{\varphi}_m^{s,v}\ , \tag{3.2}$$

where the 1-D Hamiltonian matrix $\mathbf{H}_\mathrm{1D}$ is a $N_\mathrm{x}\times N_\mathrm{x}$ tridiagonal matrix, built as follows:

$$\mathbf{H}_\mathrm{1D}=\begin{bmatrix}-2A_1 & B_1 & 0 & \cdots & \cdots & \cdots & 0\\ B_2 & -2A_2 & B_2 & 0 & \cdots & \cdots & 0\\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0\\ 0 & \cdots & \cdots & \cdots & B_{n-1} & -2A_{n-1} & B_{n-1}\\ 0 & \cdots & \cdots & \cdots & 0 & B_n & -2A_{n-1}\end{bmatrix}\ . \tag{3.3}$$

These eigenstates are normalized along the κ_s direction and the total wavefunction in lead s is then given by:

$$\Psi_{\beta,s}^v(\omega_s,\kappa_s)=\sum_{m=1}^{N_s^v}\left[a_m^{s,v}\varphi_m^{s,v}(\kappa_s)e^{\mathrm{i}k_{m,\beta}^{s,v}\omega_s}+b_{m,\beta}^{s,v}\varphi_m^{s,v}(\kappa_s)e^{-\mathrm{i}k_{m,\beta}^{s,v}\omega_s}\right]+\sum_{m=N_s^v+1}^{\infty}b_{m,\beta}^{s,v}\varphi_m^{s,v}(\kappa_s)e^{k_{m,\beta}^{s,v}\omega_s}\ . \tag{3.4}$$

Here, the terms inside the first summation on the right-hand-side represent the N_s^v traveling waves (traveling modes) with energy (along the x and z directions) $E_\beta^v > E_m^{s,v}$, going into and reflecting out of the device, respectively, through lead s. The index β denotes the different wavefunctions $\Psi_{\beta,s}^v$ and the corresponding energies E_β^v in the conduction band valley v. The third term on the right-hand-side represents the evanescent modes with energy $E_\beta^v < E_m^{s,v}$. The coefficients $a_m^{s,v}$'s are chosen as inputs for the different waves traveling into the device, while the coefficients $b_{m,\beta}^{s,v}$ need to be determined. For all other leads $j \neq s$, the injection amplitudes $a_m^{j,v} = 0$. The wavevectors $k_{m,\beta}^{s,v}$ for the traveling modes are given by: $[2m_{\mathrm{x}}^v(E_\beta^v - E_m^{s,v})]^{1/2}/\hbar$ and, for the evanescent modes, by: $[2m_{\mathrm{x}}^v(E_m^{s,v} - E_\beta^v)]^{1/2}/\hbar$. The energy E_β^v will be referred to as the 'injection energy'. Note that $\omega_s = -x$ or x depending on the left or right contact, respectively.

3.1.1 QTBM boundary conditions

The boundary conditions at the interface Γ_s between the device and lead s involve the continuity of both the wavefunction at the interface, $\phi_{\beta,s}^v\big|_{\Gamma_s} = \Psi_{\beta,s}^v(\omega_s = 0, \kappa_s)$, and, the normal derivative, $\nabla\phi_\beta^{s,v}.\hat{\kappa_s}\big|_{\Gamma_s} = \nabla\Psi_{\beta,s}^v(\omega_s = 0, \kappa_s).\hat{\kappa_s}$ for all $m \leq N_s^v$. Here $\phi_{\beta,s}^v(x, z)$ represents the wavefunction inside the device. Using Eq. (3.4) and combining the two boundary conditions together, we obtain:

$$\begin{aligned}\nabla\phi_\beta^{s,v}.\hat{\kappa_s}\big|_{\Gamma_s} = \sum_{m=1}^{N_s^v}&\left[a_m^{s,v}\varphi_m^{s,v}(\kappa_s)\frac{\partial\left(e^{\mathrm{i}k_{m,\beta}^{s,v}\omega_s}\right)}{\partial\omega_s}\big|_{\omega_s=0}\right.\\ &+\left. b_{m,\beta}^{s,v}\varphi_m^{s,v}(\kappa_s)\frac{\partial\left(e^{-\mathrm{i}k_{m,\beta}^{s,v}\omega_s}\right)}{\partial\omega_s}\big|_{\omega_s=0}\right] + \sum_{m=N_s^v+1}^{\infty} b_{m,\beta}^{s,v}\varphi_m^{s,v}(\kappa_s)\frac{\partial\left(e^{k_{m,\beta}^{s,v}}\right)}{\partial\omega_s}\big|_{\omega_s=0}\,. \qquad (3.5)\end{aligned}$$

Using finite-differences to discretize the right hand side of the above equation, we obtain

$$\begin{aligned}\frac{\partial \phi_{\beta}^{s,v}}{\partial x}\Big|_{\omega_s=0} = \sum_{m=1}^{N_s^v} \Bigg[& a_m^{s,v} \varphi_m^{s,v}(\kappa_s) \frac{e^{\mathrm{i}k_{m,\beta}^{s,v}(\omega_s=0)} \left(e^{\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1\right)}{\Delta\omega_s} \\ & + \; b_{m,\beta}^{s,v} \varphi_m^{s,v}(\kappa_s) \frac{e^{-\mathrm{i}k_{m,\beta}^{s,v}(\omega_s=0)} \left(e^{-\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1\right)}{\Delta\omega_s} \Bigg] \\ & + \sum_{m=N_s^v+1}^{\infty} b_{m,\beta}^{s,v} \varphi_m^{s,v}(\kappa_s) \frac{e^{k_{m,\beta}^{s,v}(\omega_s=0)} \left(e^{k_{m,\beta}^{s,v}\Delta\omega_s} - 1\right)}{\Delta\omega_s} \; . \qquad (3.6)\end{aligned}$$

Now, multiplying both sides of Eq. (3.4) by $\varphi_{m'}^{s,v*}(\kappa_s)$ and integrating over κ_s in the range $(0, d_s)$, where d_s is the vertical height of lead s, we obtain for $\omega_s = 0$,

$$\begin{aligned}\int_0^{d_s} \mathrm{d}\kappa_s \varphi_{m'}^{s,v*}(\kappa_s) \Psi_{\beta,s}^{v}(\omega_s = 0, \kappa_s) = \sum_{m=1}^{N_s^v} \left[a_m^{s,v} + b_{m,\beta}^{s,v}\right] & \int_0^{d_s} \mathrm{d}\kappa_{\mathrm{s}} \varphi_{\mathrm{m'}}^{\mathrm{s,v}*}(\kappa_{\mathrm{s}}) \varphi_{\mathrm{m}}^{\mathrm{s,v}}(\kappa_{\mathrm{s}}) \\ + \sum_{m=N_s^v+1}^{\infty} b_{m,\beta}^{s,v} & \int_0^{d_s} \mathrm{d}\kappa_{\mathrm{s}} \varphi_{\mathrm{m'}}^{\mathrm{s,v}*}(\kappa_{\mathrm{s}}) \varphi_{\mathrm{m}}^{\mathrm{s,v}}(\kappa_{\mathrm{s}}) \; . \qquad (3.7)\end{aligned}$$

The wavefunctions φ are eigenstates of a Hermitian operator ($\mathbf{H}_{1\mathrm{D}}$) and so are mutually orthogonal, $\int_0^{d_s} \mathrm{d}\kappa_{\mathrm{s}} \varphi_{\mathrm{m'}}^{\mathrm{s,v}*}(\kappa_{\mathrm{s}}) \varphi_{\mathrm{m}}^{\mathrm{s,v}}(\kappa_{\mathrm{s}}) = \delta_{\mathrm{mm'}}$. Using this relation and replacing Ψ_β with ϕ_β, observing continuity of the wavefunction at the interface, Eq. (3.7) gives us

$$b_{m,\beta}^{s,v} = \int_0^{d_s} \mathrm{d}\kappa_{\mathrm{s}} \varphi_{\mathrm{m}}^{\mathrm{s,v}*}(\kappa_{\mathrm{s}}) \phi_{\beta}^{\mathrm{s,v}}(\omega_{\mathrm{s}} = 0, \kappa_{\mathrm{s}}) - \mathrm{a}_{\mathrm{m}}^{\mathrm{s,v}} \; . \qquad (3.8)$$

Note that since evanescent waves are not injected, the above relation will not contain the $a_m^{s,v}$ term for $m > N_s^v$. Eq. (3.6) can be re-written as:

$$\begin{aligned}
\frac{\partial \phi_\beta^{s,v}}{\partial x}\Big|_{\omega_s=0} &= \sum_{m=1}^{N_s^v} \Bigg[a_m^{s,v} \varphi_m^{s,v}(\kappa_s) \left(\frac{e^{\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1}{\Delta\omega_s} - \frac{e^{-\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1}{\Delta\omega_s} \right) \\
&\quad + \left(\int_0^{d_s} \mathrm{d}\kappa_\mathrm{s} \varphi_\mathrm{m}^{\mathrm{s,v}*}(\kappa_\mathrm{s}) \phi_\beta^{\mathrm{s,v}}(\omega_\mathrm{s} = 0, \kappa_\mathrm{s}) \right) \varphi_m^{s,v}(\kappa_s) \frac{e^{-\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1}{\Delta\omega_s} \Bigg] \\
&\quad + \sum_{m=N_s^v+1}^{\infty} \left(\int_0^{d_s} \mathrm{d}\kappa_\mathrm{s} \varphi_\mathrm{m}^{\mathrm{s,v}*}(\kappa_\mathrm{s}) \phi_\beta^{\mathrm{s,v}}(\omega_\mathrm{s} = 0, \kappa_\mathrm{s}) \right) \varphi_m^{s,v}(\kappa_s) \frac{e^{k_{m,\beta}^{s,v}\Delta\omega_s} - 1}{\Delta\omega_s} \\
&= \sum_{m=1}^{N_s^v} \Bigg[\frac{2\mathrm{i} a_m^{s,v} \varphi_m^{s,v}(\kappa_s) \sin\!\big(k_{m,\beta}^{s,v}\Delta\omega_s\big)}{\Delta\omega_s} \\
&\quad + \left(\int_0^{d_s} \mathrm{d}\kappa_\mathrm{s} \varphi_\mathrm{m}^{\mathrm{s,v}*}(\kappa_\mathrm{s}) \phi_\beta^{\mathrm{s,v}}(\omega_\mathrm{s} = 0, \kappa_\mathrm{s}) \right) \varphi_m^{s,v}(\kappa_s) \frac{e^{-\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1}{\Delta\omega_s} \Bigg] \\
&\quad + \sum_{m=N_s^v+1}^{\infty} \left(\int_0^{d_s} \mathrm{d}\kappa_\mathrm{s} \varphi_\mathrm{m}^{\mathrm{s,v}*}(\kappa_\mathrm{s}) \phi_\beta^{\mathrm{s,v}}(\omega_\mathrm{s} = 0, \kappa_\mathrm{s}) \right) \varphi_m^{s,v}(\kappa_s) \frac{e^{k_{m,\beta}^{s,v}\Delta\omega_s} - 1}{\Delta\omega_s} \ . \qquad (3.9)
\end{aligned}$$

In practice the wavefunctions ϕ are calculated separately for each injected traveling wave and thus they depend on m as well (a superscript m will be used to denote this dependence henceforth). In our problem it has been seen that the extent of numerical errors is reduced with the introduction of more discretization into the system. For this reason, we also use a discretized version of the wave vectors $k_{m,\beta}^{s,v}$. The deduction of the same is given below in brief. We start by writing the Schrödinger equation for the forward traveling wave $e^{i\mathrm{i}k_{m,\beta}^{s,v}\omega_s}$,

$$-\frac{\hbar^2}{2m_\mathrm{x}^v} \frac{\partial^2 \left(e^{\mathrm{i}k_{m,\beta}^{s,v}\omega_r} \right)}{\partial x^2} = (E_\beta^v - E_m^{s,v}) e^{\mathrm{i}k_{m,\beta}^{s,v}\omega_s} \ .$$

Using centered differences method, the second order derivative can be written as:

$$-\frac{\hbar^2}{2m_\mathrm{x}^v} \frac{e^{\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} + e^{-\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 2}{\Delta\omega_s^2} = E_\beta^v - E_m^{s,v}$$

$$,\text{or, } \cos\!\big(k_{m,\beta}^{s,v}\Delta\omega_s\big) - 1 = -m_\mathrm{x}^v \Delta\omega_s^2 (E_\beta^v - E_m^{s,v})/\hbar^2 \ .$$

Using basic trigonometric identities, the discretized wave vector can be retrieved as,

$$k_{m,\beta}^{s,v} = \frac{2}{\Delta\omega_s} \arcsin\left(\sqrt{m_\mathrm{x}^v \Delta\omega_s^2 (E_\beta^v - E_m^{s,v})/(2\hbar^2)} \right) \ . \qquad (3.10)$$

Similarly for the evanescent waves ($m > N_s^v$), the wave vector can be written as:

$$k_{m,\beta}^{s,v} = \frac{2}{\Delta\omega_s} \operatorname{arcsinh}\left(\sqrt{m_{\mathrm{x}}^v \Delta\omega_s^2 (E_m^{s,v} - E_\beta^v)/(2\hbar^2)}\right) \ . \tag{3.11}$$

The QTBM Hamiltonian used to calculate the wavefunctions $\phi_{m,\beta}^{s,v}$ is built using the closed system Hamiltonian $\mathbf{H}^{\mathrm{D}}$, described in Sec. 2.1. The matrix elements that correspond to the drain and source edge of the device are modified to account for the QTBM boundary conditions. Let $l^1,\ l^2, ...,\ l^{N_{\mathrm{L}}}$ and $r^1,\ r^2, ...,\ r^{N_{\mathrm{R}}}$ be mesh indices of the N_{L} ($N_{\mathrm{L}} \leq N_{\mathrm{x}}$) and N_{R} ($N_{\mathrm{R}} \leq N_{\mathrm{x}}$) points in the range $(1, N_{\mathrm{x}} \times N_{\mathrm{y}})$, representing the left and right contacts, respectively. Two $N \times N$ matrices $\mathbf{\Sigma}_{\mathrm{L}}$ and $\mathbf{\Sigma}_{\mathrm{R}}$ are built which include the reflected and transmitted (both traveling and evanescent) waves traveling into and out of the device, respectively, shown below,

$$\mathbf{\Sigma}_{\mathrm{L}} =$$

$$\begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \mathrm{A}^{l^1,l^1} - \mathrm{L}^{l^1,l^1} & \cdots & -\mathrm{L}^{l^1,l^2} & \cdots & -\mathrm{L}^{l^1,l^j} & \cdots & 0 \\ \vdots & \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \cdots & -\mathrm{L}^{l^2,l^1} & \cdots & \mathrm{A}^{l^2,l^2} - \mathrm{L}^{l^2,l^2} & \cdots & -\mathrm{L}^{l^2,l^j} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots & 0 \\ 0 & \cdots & -\mathrm{L}^{l^j,l^1} & \cdots & -\mathrm{L}^{l^2,l^2} & \cdots & \mathrm{A}^{l^j,l^j} - \mathrm{L}^{l^j,l^j} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \end{bmatrix} \tag{3.12}$$

$\mathbf{\Sigma}_{\mathrm{R}} =$

$$\begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \mathrm{A}^{r^1,r^1}+\mathrm{L}^{r^1,r^1} & \cdots & \mathrm{L}^{r^1,r^2} & \cdots & \mathrm{L}^{r^1,r^i} & \cdots & 0 \\ \vdots & \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \cdots & \mathrm{L}^{r^2,r^1} & \cdots & \mathrm{A}^{r^2,r^2}+\mathrm{L}^{r^2,r^2} & \cdots & \mathrm{L}^{r^2,r^i} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots & 0 \\ 0 & \cdots & \mathrm{L}^{r^i,r^1} & \cdots & \mathrm{L}^{r^2,r^2} & \cdots & \mathrm{A}^{r^i,r^i}+\mathrm{L}^{r^i,r^i} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \end{bmatrix} \qquad (3.13)$$

where the A's are defined in the same way as for Eq. (2.4) and

$$\mathrm{L}^{p,q} = \frac{a_1}{\Delta x}\left[\sum_{m=1}^{N_s^v} \varphi_m^{s,v*}(\kappa_s = q)\phi_{m,\beta}^{s,v}(\omega_s = 0, \kappa_s = q)\Delta\omega_r \varphi_m^{s,v}(\kappa_s = p)\right.$$
$$\times\,(e^{-\mathrm{i}k_{m,\beta}^{s,v}\Delta\omega_s} - 1) \;+ \sum_{m=N_s^v+1}^{\infty} \varphi_m^{s,v*}(\kappa_s = q)\phi_{m,\beta}^{s,v}(\omega_s = 0, \kappa_s = q)\Delta\omega_s$$
$$\left. \times\, \varphi_m^{s,v}(\kappa_s = p)(e^{k_{m,\beta}^{s,v}\Delta\omega_s} - 1)\right]\,. \qquad (3.14)$$

Here, p and q represent the mesh indices l^i (or r^j), i (or j) being an integer in the range $(1, N_{\mathrm{L}})$ (or $(1, N_{\mathrm{R}})$, respectively). We also define a $N \times 1$ matrix $\mathbf{B}^m$, expressing the component of the wave (traveling) injected into the device:

$$\mathbf{B}^m = \begin{bmatrix} 0 \\ \vdots \\ \mathrm{G}^{l^1} \\ 0 \\ \vdots \\ \mathrm{G}^{l^j} \\ 0 \\ \vdots \end{bmatrix}\,, \qquad (3.15)$$

where $\mathrm{G}^{l^j} = -a_1 \left[2\mathrm{i}a_m^{s,v}\varphi_m^{s,v}(\kappa_s = l^j)\sin\left(k_{m,\beta}^{s,v}\Delta\omega_s\right)\right]/\Delta x^2$. The resulting linear system can be written as:

$$\left[\mathbf{H}^{\mathrm{D}} + \mathbf{\Sigma}_{\mathrm{L}} + \mathbf{\Sigma}_{\mathrm{R}} + \mathbf{V}\right]\boldsymbol{\phi}_{m,\beta}^{s,v} = \mathbf{B}^m \ . \tag{3.16}$$

Here $\boldsymbol{\phi}_{m,\beta}^{s,v}$ is a $N \times 1$ vector expressing the value of electron wavefunction $\phi_{m,\beta,s}$ at each of the N mesh points. Eq. (3.16) is solved to obtain $\phi_{m,\beta,s}$ separately for the leads s, the different injection energies E_β^v and the different traveling modes m (and valley v). These wavefunctions are then used to calculate the transmission coefficient, local density-of-states (LDOS) and, most importantly, the current.

3.2 Calculation of Device Characteristics

One of the most important applications of the open boundaries Schrödinger solver is the determination of current-voltage characteristics of semiconductor devices. The potential distribution obtained from the self-consistent solving of the open boundaries Schrödinger with the Poisson equation is usually taken as input for calculating these characteristics. This solution correctly corresponds to the state of the system far from equilibrium under an applied bias. Detailed discussion on this topic will be done at a later stage. On the other hand, for cases when the system is close to equilibrium, e.g., when the drain-source bias is very low, of the order of 1 meV, a self-consistent equilibrium solution as discussed in Sec. 2.5 can also provide a good approximation to the device characteristics. Nevertheless, whatever be the input potential, we show how some of the current transport parameters can be extracted from knowledge of the wavefunctions of the Schrödinger equation with open boundaries.

3.2.1 Transmission coefficient

The transmission coefficient, strictly from a device perspective, can be defined as the ratio of the reflected probability current to the probability current incident across any cross section of the device. The probability current is a quantum mechanical quantity describing probability flux, defined as:

$$\mathbf{S} = -\frac{\mathrm{i}\hbar}{2m^*}\left(\overline{\varphi}\nabla\varphi - \varphi\nabla\overline{\varphi}\right) \ . \tag{3.17}$$

Using this definition, the transmission coefficient $T^{rr'}(E^v_\beta, m_s)$ for a traveling mode m_s from lead s ('source') with injection energy E^v_β, going towards lead s' ('drain') can be written as:

$$T^{ss'}(E^v_\beta, m_s, v) = \sum_{m_{s'}} \frac{\iint\limits_A \mathrm{d}z\mathrm{d}y \ \mathbf{S}^{\mathrm{T}}_{m_{s'},\beta,s'}.\hat{\boldsymbol{x}}}{\iint\limits_A \mathrm{d}z\mathrm{d}y \ \mathbf{S}^{\mathrm{I}}_{m_s,\beta,s}.\hat{\boldsymbol{x}}} \ , \tag{3.18}$$

where A is cross section of the leads, $\mathbf{S}^{\mathrm{T}}$ and $\mathbf{S}^{\mathrm{I}}$ are the transmitted and incident probability current in the leads, respectively. Note here that for transmitted waves, all possible traveling modes $m_{s'}$ are taken into account for an incident wave of mode m_s. Assuming identical width of the leads and the uniformity of the system in the out-of-plane y direction, the above equation can be re written as:

$$T^{ss'}(E^v_\beta, m_s, v) = \sum_{m_{s'}} \frac{\int_{\mathrm{d}_{\mathrm{s}'}} \mathrm{d}z |b^{s',v}_{m_{s'},\beta}|^2 \varphi^{s',v*}_{m_{s'}} \varphi^{s',v}_{m_{s'}}}{\int_{\mathrm{d}_{\mathrm{s}}} \mathrm{d}z |a^{s,v}_{m_s,\beta}|^2 \varphi^{s,v*}_{m_s} \varphi^{s,v}_{m_s}} \times \frac{\left[e^{\mathrm{i}k^{s',v}_{m_{s'},\beta}x} \frac{\partial}{\partial x} e^{-\mathrm{i}k^{s',v}_{m_{s'},\beta}x} - e^{-\mathrm{i}k^{s',v}_{m_{s'},\beta}x} \frac{\partial}{\partial x} e^{\mathrm{i}k^{s',v}_{m_{s'},\beta}x}\right]}{\left[e^{\mathrm{i}k^{s,v}_{m_s,\beta}x} \frac{\partial}{\partial x} e^{-\mathrm{i}k^{s,v}_{m_s,\beta}x} - e^{-\mathrm{i}k^{s,v}_{m_s,\beta}x} \frac{\partial}{\partial x} e^{\mathrm{i}k^{s,v}_{m_s,\beta}x}\right]} \ .$$

Making use of orthogonality of the wavefunctions $\varphi^{s,v}_m$ and discretizing the partial derivatives using finite-differences (forward), the above equation can be simplified to

$$T^{ss'}(E^v_\beta, m_s, v) = \sum_{m_{s'}} \frac{|b^{s',v}_{m_{s'},\beta}|^2}{|a^{s,v}_{m_s,\beta}|^2} \frac{\sin\left(k^{s',v}_{m_{s'},\beta}\Delta x\right)}{\sin\left(k^{s,v}_{m_s,\beta}\Delta x\right)} \ . \tag{3.19}$$

In order to compensate for discretization errors and maintain unitariety, we again utilize here the discretized version of $k^{r,v}_{m,\beta}$ given by Eq. (3.10),

$$T^{ss'}(E^v_\beta, m_s, v) = \sum_{m_{s'}} \frac{|b^{s',v}_{m_{s'},\beta}|^2}{|a^{s,v}_{m_s,\beta}|^2} \frac{\sin\left[2\arcsin\left(\sqrt{m^v_{\mathrm{x}}\Delta x^2(E^v_\beta - E^{s',v}_{m_{s'}})/(2\hbar^2)}\right)\right]}{\sin\left[2\arcsin\left(\sqrt{m^v_{\mathrm{x}}\Delta x^2(E^v_\beta - E^{s,v}_{m_s})/(2\hbar^2)}\right)\right]} \ .$$

Using basic trigonometric identities the above expression reduces to:

$$T^{ss'}(E^v_\beta, m_s, v) = \sum_{m_{s'}} \frac{|b^{s',v}_{m_{s'},\beta}|^2}{|a^{s,v}_{m_s,\beta}|^2} \sqrt{\frac{E^v_\beta - E^{D,v}_{m_{s'}}}{E^v_\beta - E^{S,v}_{m_s}}} \sqrt{\frac{\hbar^2 - m^v_{\mathrm{x}} \Delta x^2 (E^v_\beta - E^{s',v}_{m_{s'}})}{\hbar^2 - m^v_{\mathrm{x}} \Delta x^2 (E^v_\beta - E^{s,v}_{m_s})}} \ . \tag{3.20}$$

The transmission coefficient T^{DS} of a wave incident on the 'drain' and traveling towards the 'source' can be similarly calculated.

3.2.2 Local density of states

The local density of states (LDoS) is an important tool that can provide good insight into how charge carriers behave inside a device driven far from equilibrium. It is basically the spatial variation of density of states inside the system domain. In our case, the 2-D LDoS is calculated by assigning the corresponding 2-D electron probability distribution to the 1-D density of the states (Eq. (2.15)) of the system along the transport direction x. The LDoS $\mathcal{D}^{s,v}_{\mathrm{loc}}(E, x, z)$ for injection from lead r can be calculated for each injection energy, E^v_β, from the expression:

$$\mathcal{D}^{s,v}_{\mathrm{loc}}(E^v_\beta, x, z) \ = \sum_{m=1}^{N_r} 2\sqrt{\frac{2m^v_{\mathrm{x}}}{\hbar^2}} \frac{1}{\sqrt{E^v_\beta - E^{s,v}_m}} |\phi^{s,v}_{m,\beta}(x, z)|^2 \ . \tag{3.21}$$

Note that wavefunction $\phi_{m,\beta,s}$ has been normalized assuming infinite-volume normalization along the x direction but using a finite-volume normalization along the z direction. Therefore $|\phi_{m,\beta,s}|^2$ has dimensions of inverse length and the local density of states given in Eq. (3.21) expresses states per unit energy and area.

3.2.3 Drain current

The total current flowing through the device is obtained by solving the open boundary-conditions Schrödinger problem for both the left and right sides (drain and source) using different Fermi levels for the two cases, separated by the applied bias V_{DS}, and then subtracting the calculated current from both sides. To obtain the expression for the current, we first start with the well known current equation:

$$I_s = \sum_v \frac{2e}{(2\pi)^3} \iiint\limits_{-\infty}^{\infty} \mathrm{d}k_{\mathrm{x}}\mathrm{d}k_{\mathrm{y}}\mathrm{d}k_{\mathrm{z}}\ f(k_{\mathrm{x}}, k_{\mathrm{y}}, k_{\mathrm{z}}) \left(\iint\limits_{A} \mathrm{d}x\mathrm{d}y\ \mathbf{S}^v_{k_{\mathrm{x}},k_{\mathrm{y}},k_{\mathrm{z}}}(x, y, z).\hat{\boldsymbol{x}} \right) , \tag{3.22}$$

where I_s is the current flowing from lead s and f_s is the Fermi Dirac distribution for the corresponding lead. The term inside the parentheses represents the total probability current (flux) S_{tot} flowing through any cross section of the device. The integral over k_{z} can be replaced by a summation over the quantized k_{z} states caused by the confinement along z.

$$I_s = \sum_v \sum_{m_s} \frac{e}{2\pi^2} \iint\limits_{-\infty}^{\infty} \mathrm{d}k_{\mathrm{x}}\mathrm{d}k_{\mathrm{y}}\ f(k_{\mathrm{x}}, k_{\mathrm{y}}, m_r) \left(\iint\limits_{A} \mathrm{d}x\mathrm{d}y\ \mathbf{S}^v_{k_{\mathrm{x}},k_{\mathrm{y}},m_s}(x, y).\hat{\boldsymbol{x}} \right) .$$

Switching the variables of integration from wave vector to energy using the relations $E_{\mathrm{y}} = \hbar^2 k_{\mathrm{y}}^2 / 2m^v_{\mathrm{y}}$ and $E_{\mathrm{x}} = E^v_\beta - E^{s,v}_{m_s} = \hbar^2 k_{\mathrm{x}}^2 / 2m^v_{\mathrm{x}}$, the above equation becomes:

$$I_s = \sum_v \sum_{m_s} \frac{e}{2\pi^2\hbar^2} (m^v_{\mathrm{y}} m^v_{\mathrm{x}})^{1/2} \int_0^\infty \int_{E^{s,v}_{m_s}}^\infty \frac{\mathrm{d}E_{\mathrm{y}}}{E_{\mathrm{y}}^{1/2}} \frac{\mathrm{d}E^v_\beta}{(E^v_\beta - E^{s,v}_{m_s})^{1/2}}\ f_s(E^v_\beta, E_{\mathrm{y}}) \times \left(\iint\limits_{A} \mathrm{d}x\mathrm{d}y\ \mathbf{S}^v_{m_s,E^v_\beta}(x, y).\hat{\boldsymbol{x}} \right) . \tag{3.23}$$

Note that we have utilized the fact that f_s is an even function of k_{y} and changed the limits over the corresponding integral to $(0, \infty)$. Current continuity requires that S_{tot} be equal to the total transmitted flux which, in turn, is the product of the total incident probability current and transmission coefficient. Using Eq. (3.20) and calculating over unit width of device, the total flux $S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v)$ for traveling mode m_s and injection energy E^v_β going from lead r to lead r' can be written as:

$$S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v) = 2|a^{s,v}_{m_s,\beta}|^2 \frac{\sqrt{m^v_{\mathrm{x}}}}{\hbar^2} \sqrt{E^v_\beta - E^{r,v}_{m_s}} \times \sqrt{\hbar^2 - m^v_{\mathrm{x}} \Delta x^2 (E^v_\beta - E^{s,v}_{m_s})}\ T^{ss'}(E^v_\beta, m_s, v) . \tag{3.24}$$

Note that $S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v)$ does not depend on E_{y}. Using the above relation in Eq. (3.23), we obtain:

$$I_s = \sum_v \sum_{m_s} \frac{e}{2\pi^2\hbar^2}(m^v_{\mathrm{y}} m^v_{\mathrm{x}})^{1/2} \int_{E^{r,v}_{m_s}}^{\infty} \frac{\mathrm{d}E^v_\beta}{(E^v_\beta - E^{s,v}_{m_s})^{1/2}}\ S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v) \times \int_0^\infty \frac{\mathrm{d}E_{\mathrm{y}}}{E_{\mathrm{y}}^{1/2}} f_s(E^v_\beta, E_{\mathrm{y}}) \ . \quad (3.25)$$

In our case, instead of using a continuous spectrum for E^v_β, we use the eigen-energies obtained from the closed system. Replacing the integral over E^v_β by summation, we can write the equation for current as:

$$I_s = \sum_v \sum_\beta \sum_{m_s} \frac{e}{2\pi^2\hbar^2}(m^v_{\mathrm{y}} m^v_{\mathrm{x}})^{1/2} \frac{\Delta E^v_\beta}{(E^v_\beta - E^{s,v}_{m_s})^{1/2}} \times S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v) F_{-\frac{1}{2}}\left(E^s_{\mathrm{F}} - E^v_\beta\right) \ , \quad (3.26)$$

where E^s_{F} is the Fermi level associated with lead s. The total drain current I_{D}, as mentioned before, is calculated as the difference between the currents from the drain and source:

$$I_{\mathrm{D}} = \sum_s \sum_v \sum_\beta \sum_{m_s} \frac{\eta_s e}{2\pi^2\hbar^2}(m^v_{\mathrm{y}} m^v_{\mathrm{x}})^{1/2} \frac{\Delta E^v_\beta}{(E^v_\beta - E^{r,v}_{m_s})^{1/2}} \times S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v) F_{-\frac{1}{2}}\left(E^s_{\mathrm{F}} - E^v_\beta\right) \ , \quad (3.27)$$

where s and s' represent source and drain, respectively, $\eta_{r=\mathrm{D}} = -1$ for the drain-to-source term and $\eta_{r=\mathrm{S}} = 1$ for the source-to-drain term. The distribution (both magnitude and direction) of current density Γ at different mesh points inside the system provides important insight into the physics of electron transport and can be calculated similar to the current, using the following equation:

$$\mathbf{\Gamma}(x, z) = \sum_r \sum_v \sum_\beta \sum_{m_r} \frac{\eta_r e}{2\pi^2\hbar^2}(m^v_{\mathrm{y}} m^v_{\mathrm{x}})^{1/2} \frac{\Delta E^v_\beta}{(E^v_\beta - E^{r,v}_{m_r})^{1/2}} \times \mathbf{S}^{r,v}_{\beta,m}(x, z) F_{-\frac{1}{2}}\left(E^s_{\mathrm{F}} - E^v_\beta\right) \ , \quad (3.28)$$

where, in accordance with the definition of probability current (Eq. (3.17)), $\mathbf{S}^{r,v}_{m,\beta}(x,z)$ can be written as:

$$\mathbf{S}^{r,v}_{m,\beta}(x,z) = -\frac{\mathrm{i}\hbar}{2m^{v}_{\mathrm{x}}}\left(\phi^{r,v*}_{m,\beta}\frac{\partial\phi^{r,v}_{m,\beta}}{\partial x} - \phi^{r,v}_{m,\beta}\frac{\partial\phi^{r,v*}_{m,\beta}}{\partial x}\right)\hat{\boldsymbol{x}} \\ - \frac{\mathrm{i}\hbar}{2m^{v}_{\mathrm{z}}}\left(\phi^{r,v*}_{m,\beta}\frac{\partial\phi^{r,v}_{m,\beta}}{\partial z} - \phi^{r,v}_{m,\beta}\frac{\partial\phi^{r,v*}_{m,\beta}}{\partial z}\right)\hat{\boldsymbol{z}} \ .$$

3.3 Self-consistent *Open* System

In the previous section, we discussed how QTBM can be used to determine the current-voltage characteristics of a system under applied bias. To model correctly the behavior of a 'driven' system, we need to determine first the quasi-equilibirum state in which the system is, under the influence of the external drive. This state can later be fed as input to the QTBM simulations, described in Sec. 3.1, to obtain the transport characteristics. In this section, we will describe two schemes to determine this quasi-equilibrium state of the system, driven far from equilibrium. The two schemes are then analyzed and compared to highlight their strengths and weaknesses. In each scheme, we solve the Schrödinger equation and Poisson equation self consistently again, but with *open* boundary conditions, taking help of QTBM to correctly model the electrons flowing into and out of the device.

The two methodologies essentially differ only in how each one models the electron reservoirs leading into the source and drain terminals of the device, resulting in using either Dirichlet [26] or Neumann [32] boundary conditions for solving the Poisson equation. The former condition models the leads more realistically with electric field lines entering or exiting the contacts. The latter scheme, instead, assumes that no electric field is present at the device-contact boundary. Henceforth, we will refer to the two schemes as having Dirichlet and Neumann leads, respectively. We will describe, first, the method to solve the *open* system self consistently with Dirichlet leads and then show how the Neumann leads approximation differs from the first case.

3.3.1 *Open* system Schrödinger equation with Dirichlet leads

As with any QTBM simulations, the traveling modes are first calculated. A self-consistent potential (1-D) is first determined at the drain and source contacts that also preserves charge neutrality at the respective edges [26]. Computationally, this is achieved through two nested loops. In the inner loop, a Dirichlet 'guess' potential is fixed at the center of the respective contact and the Schrödinger and Poisson equations are solved iteratively, until convergence, to obtain the self-consistent total charge (and potential) distribution. The outer iteration then uses this mapping from the center fixed potential to the converged total charge and runs a root-finding algorithm, typically the bisection method (for guaranteed but slower convergence rate) or Newton's method (faster but convergence is conditional), which ultimately drives the total charge to zero. This potential is used to solve Eq. (3.1) (or equivalently Eq. (3.2)) to obtain the φ's.

The next step tackles the important problem of discretizing the continuous energy spectrum that represents the local density-of-states of the device including the infinite leads along the transport direction. It is convenient to use the discretization that samples preferentially states that matter most inside the device, since the charge will be determined mostly by these states. There are several ways reported in literature, to approach this problem. A conventional method is to use dirac delta normalization [24] for the injected wave functions along with a corresponding 2-D density-of-states, accounting for infinite length of the leads. Any discretized energy set will work in this case, provided the range extends sufficiently above the Fermi level of the device. The approach first proposed by Fischetti [21] and later adopted by Laux in QDAME [33, 26], uses the device "normal modes", which are standing wave solutions of the system obtained by employing 'sine-like' and 'cosine-like' boundary conditions along the contact-lead interfaces independently. Our approach is similar in essence to the QDAME method. We solve twice the "closed-system" Schrödinger equation [27] in the device domain, once assuming zero-value Dirichlet and once assuming Neumann boundary conditions along the device-lead interfaces, to double-sample the continuous energy spectrum. The resulting solutions (eigenfunctions) are 'sine-like' and 'cosine-like' at the interfaces and therefore the corresponding eigen-energies form a complete orthogonal basis (injection energies E^v_β) of the whole Hilbert space and any other solution can be represented by a linear combination of these two. The sampling technique is robust enough to successfully discretize the continuous energy spectrum and has provided good convergence and proper device solutions, as will be shown later in Sec. 5.4. An added benefit of this strategy is that a straight forward finite volume ('box') normalization can be used for the open system wavefunctions [21, 26].

The Schrödinger equation is now solved inside the device domain Ω in the same way as extensively described in Sec. 3.1 to get the 2-D wavefunctions $\phi^{s,v}_{m,\beta}$. The waves are 'box' normalized as follows:

$$\int\limits_{\Omega} \mathrm{d}x\mathrm{d}z \left(\sum_{r} \sum_{m=1}^{N^v_s} |\phi^{s,v}_{m,\beta}(x,z)|^2 \right) = \frac{1}{2} \ . \tag{3.29}$$

The use of $\sum_m |\phi_m|^2$ instead of $|\sum_m \phi_m|^2$ signifies that the leads do not inject the traveling modes coherently [26]. The factor of 1/2 is used because the energy spectrum is sampled twice by the 'sine-like' and 'cosine-like' eigenfunctions. The electron density is then calculated by associating with each wavefunction the corresponding occupancy factor along with a 1-D density of states (Eq. (2.15)) and summing over all possible traveling modes for all possible injection states:

$$n(x,z) = \sum_{s} \sum_{v} \sum_{\beta} \sum_{m} \frac{1}{\pi\hbar} \sqrt{\frac{m^v_{\mathrm{y}} k_{\mathrm{B}} T}{2}} F_{-\frac{1}{2}} \left(\frac{E^s_{\mathrm{F}} - E^v_\beta}{k_{\mathrm{B}} T} \right) |\phi^{s,v}_{m,\beta}(x,z)|^2 \ , \tag{3.30}$$

where the symbols hold their usual meanings. Note that the electrons are calculated for different leads r separately using the corresponding Fermi level E^s_{F} associated with the lead and then summed together. The holes are again calculated semiclassically similar to the closed system analysis using Eq. (2.18).

Once all the charges, electrons, holes and ionized dopants, are known, Poisson equation Eq. (2.21) is solved inside the 2-D device domain to obtain the 'new' potential, to be used in the subsequent iteration of the self-consistent simulation. The same matrix **P**, described in Sec. 2.4, is used here as well. The potential at the lead-drain and lead-source interfaces is fixed to the 1-D self-consistent potential, determined at the start of this section, by applying Dirichlet boundary conditions at the respective sections of the Poisson matrix. This step ensures charge neutrality at the device-lead interfaces. Computationally, all the elements of the rows representing the device-lead interfaces on **P** are assigned zeros, except for the diagonal terms which are assigned unity. The self-consistent 1-D potential energy is added to the corresponding rows of the matrix **D** of Eq. (2.22). This fixes the potential energy exactly at the device-edge interface.

Broyden's self-consistent scheme

Similar to the procedure followed in Sec. 2.5, the Poisson equation is not solved directly, instead, a convergence scheme is used that incorporates information about 'past' potentials to determine the 'new' potential. We had formerly described the conventional Newton's iteration scheme to accelerate the convergence of the closed system. The same method has also been found to produce good convergence in the open system case as well. However, acknowledging the increased complexity of the problem at hand, we employ a more sophisticated convergence scheme, the Broyden's method in conjunction with Newton's (or secants) method, to further accelerate the rate of convergence. We start with a semiclassical Jacobian $\mathbf{J_1}$, described by Eq. (2.25), in the first iteration and calculate $\mathbf{V}_2$ using Newton's method. The initial 'guess' $\mathbf{V}_1$ is usually taken as the self-consistent potential obtained from the 'closed' system under identical input conditions. At each subsequent iteration i, the Jacobian $\mathbf{J}$ is updated in the following manner:

$$\mathbf{J}_i = \mathbf{J}_{i-1} + \frac{\Delta\mathbf{F}_i - \mathbf{J}_{i-1}\Delta\mathbf{V}_i}{||\Delta\mathbf{V}_i||^2}\Delta\mathbf{V}_i^{\mathrm{T}} \ , \tag{3.31}$$

where $\mathbf{F}_i = \mathbf{P}\cdot\mathbf{V}_i - \mathbf{D}_i$, $\Delta\mathbf{F}_i = \mathbf{F}_i - \mathbf{F}_{i-1}$ and $\Delta\mathbf{V}_i = \mathbf{V}_i - \mathbf{V}_{i-1}$. We can directly update the inverse of the Jacobian by applying the Sherman-Morrison formula to the above equation, giving us

$$\mathbf{J}_i^{-1} = \mathbf{J}_{i-1}^{-1} + \frac{\Delta\mathbf{V}_i - \mathbf{J}_{i-1}^{-1}\Delta\mathbf{F}_i}{\Delta\mathbf{V}_i^{\mathrm{T}}\mathbf{J}_{i-1}^{-1}\Delta\mathbf{F}_i}\Delta\mathbf{V}_i^{\mathrm{T}}\mathbf{J}_{i-1}^{-1} \ . \tag{3.32}$$

We then proceed with the Newton's method to calculate the 'new' potential $\mathbf{V}_{i+1}$ using $\mathbf{J}_i^{-1}$ in Eq. (2.27). It is important to mention here that computationally it is very expensive to store the full matrix $\mathbf{J}_i^{-1}$ at every iteration step. So instead, Eq. (3.32) is solved recursively at each step. In this manner, only the initial $\mathbf{J}_1^{-1}$ and the vectors $\mathbf{V}_i$, $\mathbf{F}_i$ at each iteration need to be stored. The procedure can be made more efficient by not directly calculating $\mathbf{J}_1^{-1}$ but instead computing $(\mathbf{J}_1^{-1}\cdot\mathbf{F}_i)$ at each iteration and then using any linear system solver to get the new potential through Newton's method, as mentioned before in Sec. 2.5.1.

Inclusion of Drift Vector

The self-consistent solution of the *open* system, obtained in the previous section, exhibits a discontinuity of the electron flux at the lead-device interfaces once the system is driven far from equilibrium, typically with the application of high V_{DS} in the case of FETs. At large currents, where there is substantial transmission of electrons, unphysical electron depletion and accumulation will occur in the originating and receiving contacts, respectively. The issue arises because our boundary conditions assume equilibrium contacts that do not take into consideration the external current flowing from one contact to another through the outer closed electrical circuit connected to the device. Conventional techniques to tackle this problem include adjusting Fermi levels at leads to maintain charge neutrality [34, 16], or injecting a drifted Fermi distribution [22, 26] to achieve charge neutrality or current continuity at the leads. The latter method is physically consistent with current flow in metallic leads or highly doped homogenous systems [34]. Therefore, we will adopt this method, using a drifted wave vector k^r_{D} to maintain current continuity at the leads, where charge neutrality has already been enforced at the lead-device interface by our 'Dirichlet leads' approach. k^s_{D} is determined self-consistently at every iteration using the condition of current continuity between the device and the lead s.

At each self consistent iteration, the total energy $E_{\mathrm{tot}}(k_m, E^v_\beta, k^s_{\mathrm{D}})$ of a wave is now defined as follows:

$$\begin{aligned} E^v_{\mathrm{tot}}(m, E^v_\beta, k^s_{\mathrm{D}}) &= \frac{\hbar^2 (k^{s,v}_{m,\beta} - k^s_{\mathrm{D}})^2}{2m^v_{\mathrm{z}}} + E^{s,v}_m + \frac{\hbar^2 (k^v_{\mathrm{y}})^2}{2m^v_{\mathrm{y}}} \ , \\ &= E^v_\beta + \frac{\hbar^2 k^s_{\mathrm{D}} (k^s_{\mathrm{D}} - 2k^{s,v}_{m,\beta})}{2m^v_{\mathrm{z}}} + \frac{\hbar^2 (k^v_{\mathrm{y}})^2}{2m^v_{\mathrm{y}}} \ . \end{aligned} \tag{3.33}$$

Let us define $E^{m,\beta,s,v}_{\mathrm{D}} = \hbar^2 k^s_{\mathrm{D}}(k^s_{\mathrm{D}} - 2k^{s,v}_{m,\beta})/(2m^v_{\mathrm{z}})$. The device current I^s_{dev} entering lead s is defined using Eq. (3.27) and including the drift term, giving us the expression:

$$\begin{aligned} I^s_{\mathrm{dev}} = \sum_s \sum_v \sum_\beta \sum_{m_s} & \frac{\eta_s e}{2\pi^2 \hbar^2} (m^v_{\mathrm{y}} m^v_{\mathrm{x}})^{1/2} \frac{\Delta E^v_\beta}{(E^v_\beta - E^{s,v}_{m_s})^{1/2}} \\ & \times S^{ss'}_{\mathrm{tot}}(E^v_\beta, m_s, v) F_{-\frac{1}{2}} \left(E_F - E^v_\beta - E^{m,\beta,s,v}_{\mathrm{D}} \right) \cdot (-\hat{\omega}_s) \ , \end{aligned} \tag{3.34}$$

The lead current I^s_{lead} is the product of the drift velocity and total charge, namely ionized impurities, in the leads and can be written as:

$$I^s_{\mathrm{lead}} = -e \frac{\hbar k^s_{\mathrm{D}}}{m_{\mathrm{D}}} \int_0^{d_s} \mathrm{d}\kappa_s N_{\mathrm{D}}(\kappa_s) \ , \tag{3.35}$$

where $1/m_{\mathrm{D}} = (\sum_v 1/m^v_{\mathrm{x}})/(\sum v)$ and N_{D} is the concentration of ionized dopants in the lead. A root finding algorithm, typically bisection method, is used to determine the value of k^v_{D} for which $I^s_{\mathrm{dev}} = I^s_{\mathrm{lead}}$ at each self-consistent iteration. A positive k_{D} implies additional electron injection into the device and a negative k_{D} corresponds to exit of excess electrons out of the device. The determined k_{D} is also incorporated into the calculation of electron charge in the self-consistent loop, affecting the Fermi-Dirac expression of Eq. (3.30).

Inclusion of completely bound electron states

A system can have electronic states that are completely bound inside the system domain and are not connected to any of the contacts. It is not possible to determine these states using QTBM. Such bounded states cannot carry any current but they can still contribute to the electronic charge distribution inside the device and therefore should not be overlooked in self-consistent simulations of *open* systems. Particularly for devices that have quantum confinements deep inside their channels, the impact of these states can be significant. Our approach of calculating eigenstates of the *closed* system at the start of QTBM can be used to tackle this problem to some extent, as we will see. We first determine the injection energies that are lower than the lowest subband energy and label these states as completely bounded. The *closed* system wavefunctions of these states are used to determine their contribution to the total charge, by assigning a Fermi-Dirac distribution and 1-D density of states to each wavefunction and summing over all of them. However, the complication then arises on which Fermi level, drain or source, to choose for these bounded states. We approximate the Fermi level $E^{\mathrm{BD}}_{\mathrm{F}}$ associated with a bound state $\xi^{\mathrm{BD}}(x, z)$ by an average of the Fermi levels E^s_{F} of the different leads, as shown below:

$$E^{\mathrm{BD}}_{\mathrm{F}} = (\sum_r G_s E^s_{\mathrm{F}})/(\sum G_s) \ , \tag{3.36}$$

where $G_s = \int_0^{d_s} \mathrm{d}\kappa_s |\xi^{\mathrm{BD}}(\omega_s = 0, \kappa_s)|^2$. Note that we do not take into account completely bound states at energies higher than the lowest subband. Our assumption is that, at such high energies, the contribution of these bound states will be negligible compared to the current carrying states. The addition of completely bound states in this manner has produced small yet visible difference in our simulations.

3.3.2 *Open* boundary system with Neumann leads

In Sec. 3.3.1, we highlighted a couple of methods by which the issue of excess charge accumulation or depletion at the interfaces of *open* systems is addressed. Here we bring forth another well-known approach to this problem and compare it with our previously described approach, highlighting the fundamental differences as well as the merits and de-merits of each. We call this simulation of *open* system with Neumann leads. Here the self-consistent potential obeying charge neutrality at the edges, discussed in Sec. 3.3.1, is not determined. Instead the 1-D Schrödinger equation is computed at each self-consistent iteration to determine the traveling and evanescent modes. Since a pre-determined potential is not fixed at the lead-device interfaces, the 2-D Poisson equation is solved with zero-derivative Neumann boundary conditions on all sides, except of course at the junction that forms the gate interface. Such a condition not only enforces charge neutrality at the lead interfaces but also forces the electric field to be zero in the same region. This allows the lead conduction band to float as the current flow increases and the issue of current discontinuity at the leads is avoided, rather conveniently. Thus the leads are modeled similar to the ideal 'wires' connecting the different circuit elements of an electrical circuit. The major criticism of this approach is that having zero electric field at the lead interfaces is not a physical situation as a biased device should always have field lines entering or exiting it. On the contrary, the Neumann leads approach is consistent with QTBM theory which is based on the assumption that the potential in the leads is constant along the transport direction.

The Dirichlet leads approach, however, takes into account the existence of the electric field, enforcing only charge neutrality at the lead interfaces. Thus, it is physically consistent but it clashes with the basic assumptions of QTBM. In view of all this, the most fundamentally correct way would be to acknowledge the existence of an electric field (and potential gradient) in the contacts and use something like Airy functions, instead of the simple plane waves, for the QTBM boundary conditions along with Dirichlet leads. However, this would complicate the problem significantly and we would be forsaking the usefulness of a (comparatively) simpler QTBM model. We can conclude that both Dirichlet and Neumann leads approaches have their own strengths and weaknesses. The Neumann leads approach has relatively less computational complexity and a faster convergence rate and therefore we will favor this approach for preliminary tests. To achieve the final output, the Dirichlet leads approach will be given priority as it is physically more consistent. We should mention that the two methods have produced somewhat identical results in most of the systems we have studied so far.

The following few chapters will now present some applications of our ballistic transport model. In Chapter 4, we simulate an explicit quantum phenomenon in the form of negative differential transconductance observed in the transfer characteristics of a novel class CMOS devices, namely lateral quantum-well MOSFETs. Chapter 5 reports the numerical prediction of Fano interference phenomenon in double-gate MOSFETs. Observation of this purely quantum phenomenon in such a realistic CMOS device structure presents a novel and intriguing case. Finally, in Chapter 6, we theoretically analyze the effect on ballistic electron transport when discrete edge roughness is introduced along the dielectric-semiconductor interface in ultra-thin body FETs. The dissipative impact of the scattering of electrons with the surface roughness is obtained through a statistical study over an ensemble of devices with identical structure but with random roughness patterns.

CHAPTER 4

THEORETICAL SIMULATION OF NEGATIVE DIFFERENTIAL TRANSCONDUCTANCE IN LATERAL QUANTUM WELL NMOS DEVICES [1]

Authors – Pratik B. Vyas, Clint Naquin, Hal Edwards, Mark Lee, William G. Vandenberghe, and Massimo V. Fischetti

Department of Electrical and Computer Engineering, EC33

The University of Texas at Dallas

800 West Campbell Road

Richardson, Texas 75080-3021

---

[1] Adapted from [P. B. Vyas, C. Naquin, H. Edwards, M. Lee, W. G. Vandenberghe, and M. V. Fischetti, Journal of Applied Physics 121, 044501 (2017). DOI: 10.1063/1.4974469], with the permission of AIP Publishing.

4.1 Preface

We present a theoretical study of the negative differential transconductance (NDT) recently observed in lateral-quantum-well Si n-channel field-effect transistors [35]. In these devices, p^+ doping extensions are introduced at the source-channel and drain-channel junctions, thus creating two potential barriers that define the quantum well across whose quasi-bound states, resonant/sequential tunneling may occur. Our study, based on our ballistic transport model discussed in Chapter 3, predicts the presence of a sharp NDT in devices with a nominal gate length of 10-to-20 nm at low temperatures ($\sim$10 K). At higher temperatures, the NDT weakens and disappears altogether as a result of increasing thermionic emission over the p^+ potential barriers. In larger devices (with gate length of 30 nm or longer), the NDT cannot be observed because of the low transmission probability and small energetic spacing (smaller than $k_{\mathrm{B}}T$) of the quasi-bound states in the quantum well. We speculate that the inability of the model to predict the NDT observed in 40 nm gate-length devices may be due to an insufficiently accurate knowledge of the actual doping profiles. On the other hand, our study shows that NDT suitable for novel logic applications may be obtained at room temperature in devices of the current or near-future generation (sub-10 nm node), provided an optimal design can be found that minimizes thermionic emission (requiring high p^+ potential-barriers) and punch-through (that meets the opposite requirement of potential-barriers low enough to favor the tunneling current).

The contents of this chapter are taken with permission from "Theoretical simulation of negative differential transconductance in lateral quantum well nMOS devices" [Journal of Applied Physics 121, 044501 (2017)] Copyright (2017) AIP Publishing. The authors are Pratik B. Vyas, Clint Naquin, Hal Edwards, Mark Lee, William G. Vandenberghe, and Massimo V. Fischetti. I developed the simulation tool, generated and analyzed the result, and wrote majority of the manuscript. Dr. Clint Naquin and Dr. Hal Edwards provided details regarding the profile of the devices under study and took part in interactive discussions. Dr. Mark Lee was the co-principal investigator of the project along with Dr. Massimo V. Fischetti. He supervised the experimental side of the project, played a major role in correlating experimental and simulated data, and took active part in all the project meetings. Dr. William G. Vandenberghe provided valuable guidance in developing the simulation tool, recommended test cases, analyzed the results and assisted with writing the manuscript. Dr. Massimo V. Fischetti supervised the theoretical study of the devices, provided critical guidance in understanding and developing the physical models, analyzed the results, and helped with writing and reviewing the manuscript. Portions of the manuscript have been excluded to maintain the flow of the dissertation.

4.2 Introduction

One little explored route towards achieving the very aggressive high frequency and high sensitivity goals outlined in the International Technology Roadmap for Semiconductors (ITRS)[11] for the silicon semiconductor industry is to move beyond conventional semi-classical device physics. This can be done by integrating explicitly quantum mechanical transport into industrial Si complementary metal-oxide-semiconductor (CMOS) technology. This route would enable Si CMOS to emulate the path of III-V devices that incorporate transport through electronic states that are localized and whose energy is discretized by quantum confinement. Esaki diodes[36, 37] and resonant tunneling diodes (RTDs)[38, 39, 40] undoubtedly constitute the best-known examples of devices based on III-V compound semiconductors that have concretely realized this idea. These instances have yielded new capabilities in very high speed and very low-noise applications [12, 13, 14], a level of performance that is beyond the reach of conventional semi-classical devices. However, Si CMOS quantum devices must be fabricated within the standardized progression of industrial process nodes (now approaching the 7 nm node at the high end of the performance scale, but still commonly relying on the processing technology of the 28 or 22 nm nodes) to guarantee economically scalable production. This restriction rules out the two primary methods of fabricating quantum structures in III-V devices, hetero-epitaxial layer growth and electron-beam lithography, therefore rendering more difficult the possibility of entering the quantum-transport regime in Si CMOS devices. Recently, however, tantalizing hints of quantum transport have been observed in Si CMOS devices using lateral quantum wells defined by ion implantation[41, 35]. Here, we present a theoretical study of quantum electron transport in these devices, emphasizing the limitations of both their practical realization as well as of our understanding of the basic physical processes involved. At the same time, we also show that this type of quantum transport may be achieved in practice and indicate possible future promising paths that the technology may follow.

4.3 Device Description and Experimental Observations

An explicit experimental demonstration of quantum transport in Si n-channel MOS (nMOS) transistors fabricated using an industrially standard 45 nm-node process technology has been recently reported [41, 35]. These nMOS devices have a lateral quantum well (QW) built into the surface channel. This is obtained by reversing the ion-implantation dopant polarity of the shallow source/drain (S/D) extensions (pSDE) from the standard n-type (for an nMOS transistor) to p-type, as sketched in Fig. 4.1.

The p-type extensions create an energy barrier for electrons between the n^{++} S/D and the surface channel beneath the gate. A two-dimensional electron quantum well (QW) is formed when the gate voltage V_{GS} is large enough to invert the channel between the p-type extensions. The depth of the QW can be controlled by the source/gate bias V_{GS}. Explicit evidence of quantum transport in these QW nMOS devices was shown in the form of a negative differential transconductance (NDT). This occurs when the drain-source current (I_{DS}) behaves non-monotonically, so that $g_{\mathrm{m}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{GS}} < 0$. This has been observed

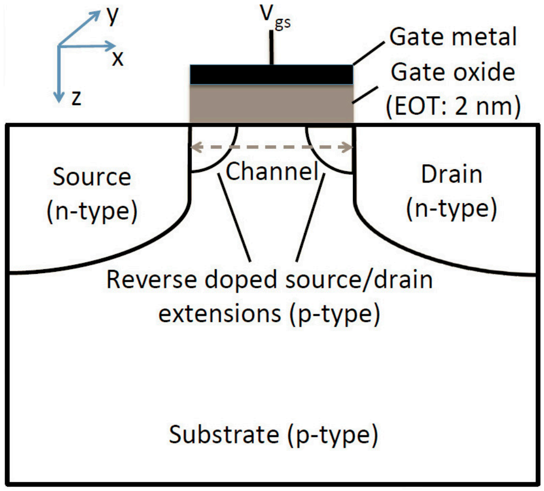

Figure 4.1: Schematic cross-section of the lateral QW nMOSFET [1].

only in QW nMOS devices, but not in standard nMOS devices fabricated on the same chip as experimental controls, thus showing that indeed the quantized electronic states in the lateral QW play a fundamental role in controlling the electronic transport. Whereas NDT is an expected signature of direct or sequential tunneling through discrete QW bound states [38, 42], several quantitative aspects of the reported NDT indicate that the detailed physical mechanism causing the NDT is more complicated than a straightforward QW tunneling phenomenon. Among these apparently anomalous features is the observation of only a small number of NDT peaks (no more than 3) observed in any given device with a width and V_{GS}-separation much larger than expected, given the nominal QW lengths. Poorly understood is also the need to apply a positive body current or voltage bias, that introduces bipolar-like operational characteristics, in order to observe the NDT [35].

4.4 Theoretical Formulation

The phenomenon that we wish to see and which is the motivation behind this whole project, is the presence of NDT in the current-voltage characteristics as a consequence of resonant or sequential tunneling in the channel of the QW nMOS device. The objective of the two-dimensional confinement created in the channel is to produce bound electronic states in

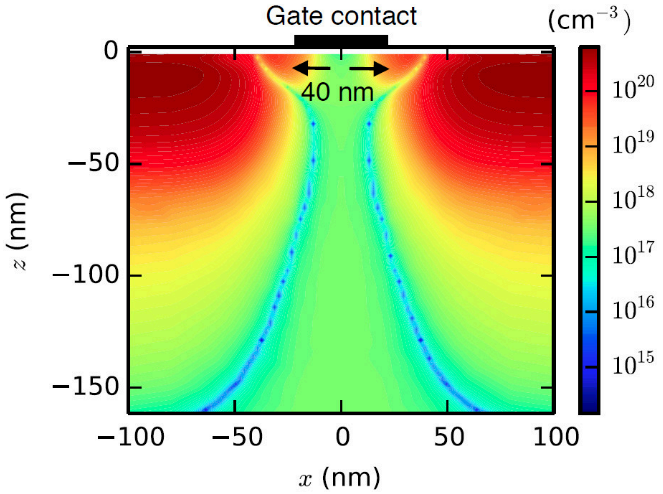

Figure 4.2: Magnitude of the net doping profile of the device with a nominal gate length of 40 nm. The white region at the top represents the gate oxide [1].

that region. Electrons injected into the channel at energies equal to these bound states will undergo resonant tunneling that occurs with a very high transmission coefficient. The magnitude of the transmission coefficient depends on the relative height of the potential barriers. The gate bias acts as a control for the energy of the bound states. At certain gate biases, the energy of these bound states will coincide with the Fermi energy of the electrons in the source region. Therefore, at the appropriate gate bias, the source-to-drain current, I_{DS}, should exhibit sharp peaks, as the electrons that contribute most to the current have a high resonant transmission coefficient. Scaling the device from its nominal size (printed gate length of about 40 nm) to the smaller dimensions we have considered (30, 20, and 10 nm gate length) has been done following conventional scaling laws [43, 44] as strictly as possible. Obvious exceptions had to be made regarding the scaling of the doping concentrations, since their high values in the 40 nm device cannot be realistically increased with scaling demands. Moreover, the SiO_2-equivalent thickness of the gate insulator (EOT) has been kept fixed at 2 nm. This is not a crucial parameter for the application of interest here. Crucial, instead, is the scaling of the doping profiles of the p-type substrate, S/D regions, and of the pSDEs: Preventing short-channel effects (mainly punch-through) in short devices, requires increasing the p-type substrate doping. Since the peak pSDE doping is already quite large in the 40 nm device, the higher scaled substrate doping results in a reduced height of the pSDE potential barriers in shorter devices. This does result in the desired boost of the tunneling current across the QW and in a reduced punch-through current. On the other hand, it also results in the undesired occurrence of a large thermionic current (over the pSDE barriers) that hides the (resonant) tunneling current. The problem caused by this narrow 'design window' will be discussed in Sec. 4.5.

Our simulation method proceeds in two steps. First, we solve the two-dimensional (2D) Schrödinger equation under *closed* boundary condition self consistently with the Poisson equation, as described in Chapter 2. This gives quantitative information about the energetic positions of the confined states for various gate-bias conditions and provides the equilibrium electrostatic potential. As a second step, electron transport via tunneling through the confined states – the effect that gives rise to the NDT – is studied by solving the Schrödinger equation with *open* boundary conditions, as detailed in Sec. 3.1. This allows us to calculate the current-voltage characteristics. We perform this calculation step by using the self-consistent potential obtained from the previous step. Indeed, self-consistent effects between the Schrödinger equation with open boundary conditions and the Poisson equation are not expected to alter the potential profile significantly at the low source-to-drain bias, V_{DS}, of interest (of the order of a few tens of mV). This choice of low V_{DS}, also employed in the experimental observations[41, 35], ensures that electrons remain in near-equilibrium conditions, so that the use of the self-consistent solution obtained for the closed system constitutes an excellent approximation also for the open system. Thus, having obtained the electrostatic potential, we calculate the current by applying a small shift V_{DS} of a few mV between the Fermi levels in the source and drain.

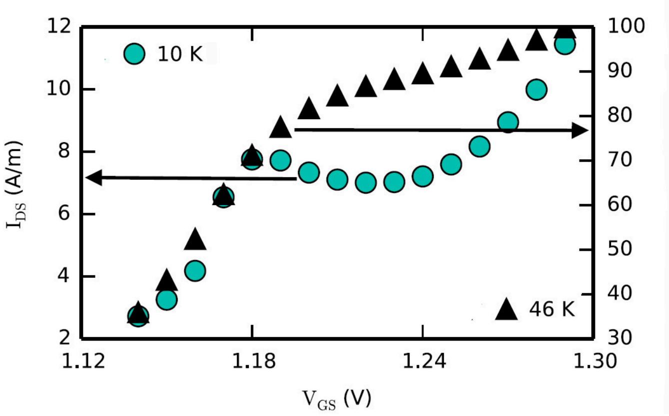

Figure 4.3: Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics at 46 K (black triangles) and 10 K (cyan circles) for the 10 nm device with $V_{\mathrm{DS}} = 10$ mV [1].

4.5 Simulation Results and Discussion

We present now the current-voltage ($I_{\mathrm{DS}} - V_{\mathrm{GS}}$) characteristics of the 10 nm, 20 nm, and 40 nm devices in Figs. 4.3, 4.4, and 4.5, respectively. The 10 nm device exhibits NDT at V_{GS}= 1.18 V at a temperature of 10 K. The NDT is reduced to a small 'kink' in the $I_{\mathrm{DS}}-V_{\mathrm{GS}}$ characteristics as the temperature is raised to 46 K, as shown in Fig. 4.3, anddisappears altogether at room temperature. A similar behavior is seen in the 20 nm device: The NDT seen at V_{GS}= 1.93 V at 10 K (Fig. 4.4), but not at room temperature. A 'kink' is seen in the current voltage characteristics of the 30 nm device at 10 K (not shown), but no defined NDT is detected. Finally, the 40 nm device does not exhibit any NDT peaks at any temperature (Fig. 4.5). Note that in all devices as the temperature increases the current I_{DS} increases exponentially, as a result of thermionic emission over the pSDE potential barriers (discussed later).

To confirm that the NDT seen at low temperatures is indeed the result of resonant tunneling through the two-dimensional quantum well created by the pSDEs, we plot in Fig. 4.6 the average local density-of-states (LDOS) for all devices. The contour plots show the two dimensional LDOS computed along the length (x direction) of the device, averaged

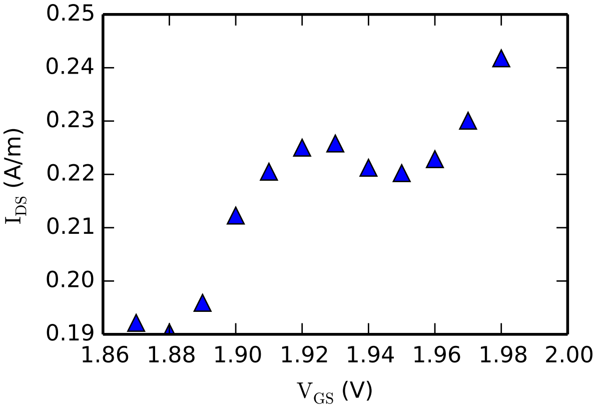

Figure 4.4: Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics for the 20 nm device at 10 K. $V_{\mathrm{DS}} = 1$ mV [1].

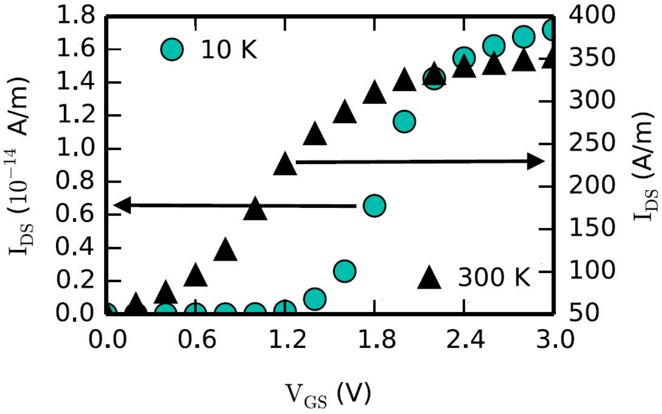

Figure 4.5: Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics at 10 K (cyan circle) and 300 K (black triangle) for the 40 nm device with $V_{\mathrm{DS}} = 10$ mV. Negligible current is seen at 10 K. Thermionic emission is therefore the major cause of the vastly larger current observed at 300 K [1].

over a thin 'vertical' region (z direction). The energy scale identifies different injection energies. The LDOS is shown only for injection from the source (left contact). In Fig. 4.7 we show the transmission coefficient as a function of injection energy for the traveling modes that exhibits the best resonant behavior (wherever applicable). Only the LDOS and transmission coefficient *vs.* total injection energy plots corresponding to the two Si ellipsoidal energy valleys having the longitudinal mass in the z direction ('unprimed subbands') are shown in Figs. 4.6 and 4.7, since these are the only valleys that exhibit NDT.

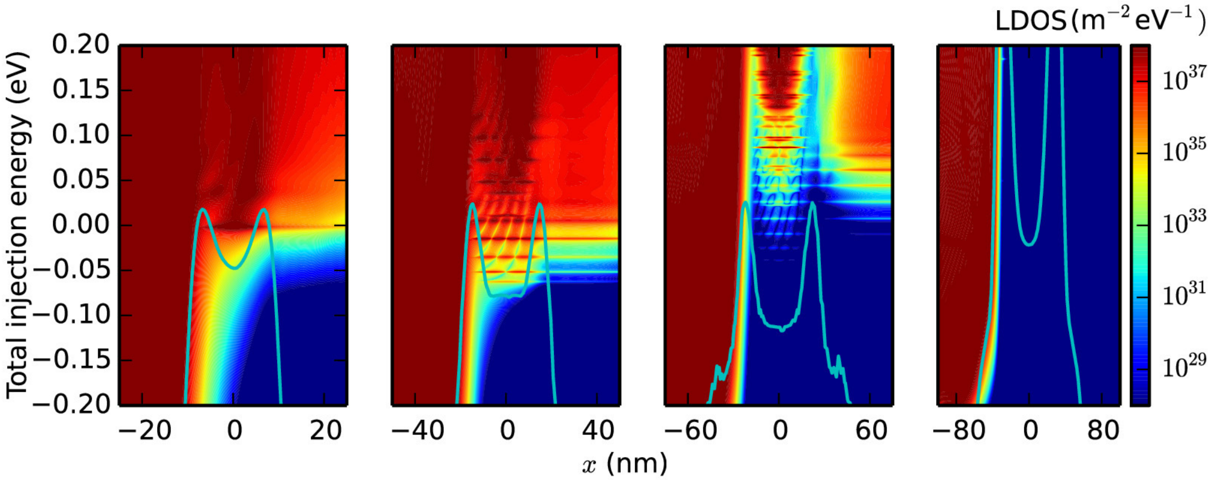

Figure 4.6: Average LDOS in the channel for 10 nm, 20 nm, 30 nm, and 40 nm devices (from left to right) at 10 K. The cyan colored lines represent the potential-energy profile at the semiconductor/gate-insulator interface in each device. The energies are measured with respect to the Fermi energy in the source contact. The bias conditions are $V_{\mathrm{GS}} = 1.18$ V, 1.93 V, 2.2 V, and 2.2 V, respectively (left to right) [1].

The dark 'streaks' in the middle of the channel seen in Fig. 4.6 (first and second frames from the left, respectively) show the presence of the quasi-bound states in the 10 nm and 20 nm devices. The corresponding peaks seen in the transmission coefficient (first and second frames from the left in Fig. 4.7, respectively) at those energies confirm that resonant tunneling across the QW is indeed the origin of the NDT. More so, the NDT occurs at a gate bias for which the energy of the first bound state (10 nm device) or fourth bound state (20 nm device) crosses the Fermi energy of the electron gas in the source region. The LDOS for the 30 nm device (third frame from the left in Fig. 4.6) shows closely-spaced bound states and no quasi-bound states can be seen in the LDOS of the 40 nm device (right-most frame in Fig. 4.6). The corresponding transmission coefficient (two rightmost frames in Fig. 4.7), show closely-spaced peaks with very low transmission coefficients (30 nm) or no peaks (40 nm), explaining the absence of any NDT seen in the current-voltage characteristics. We should also emphasize that all devices are affected by a significant punch-through current. The severity of this problem is reduced in the shorter devices, thanks to the reduced height of the

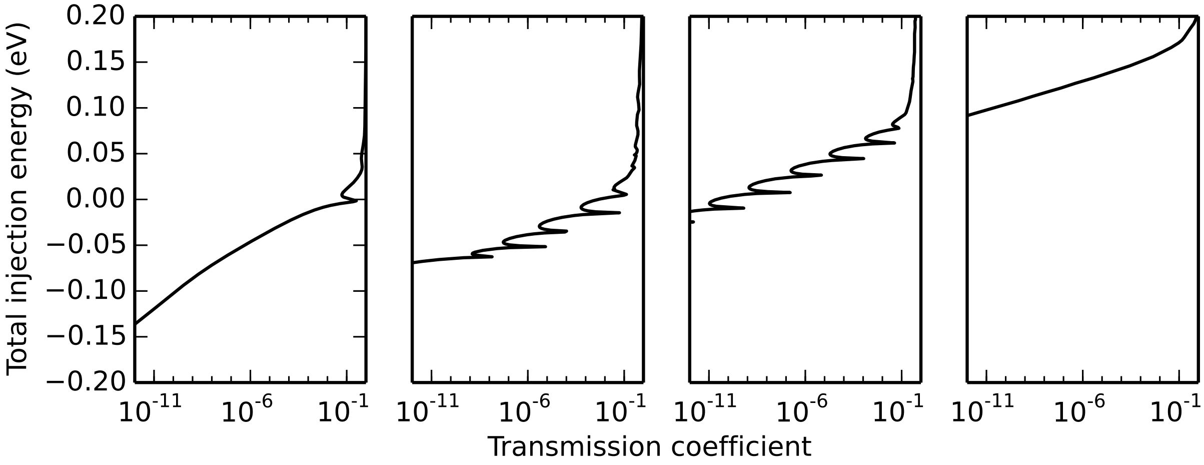

Figure 4.7: Transmission coefficient *vs.* injection energy for a particular traveling mode in the 10 nm, 20 nm , 30 nm and 40 nm devices, respectively. The traveling mode energies E_m^r are, from left to right, -0.83 eV, -0.49 eV, -0.26 eV and -0.04 eV. These energies are chosen since they exhibit the best resonant behavior in the respective devices (wherever applicable). $V_{\text{GS}} = 1.18$ V, 1.93 V, 2.2 V, and 2.2 V, respectively (left to right) [1].

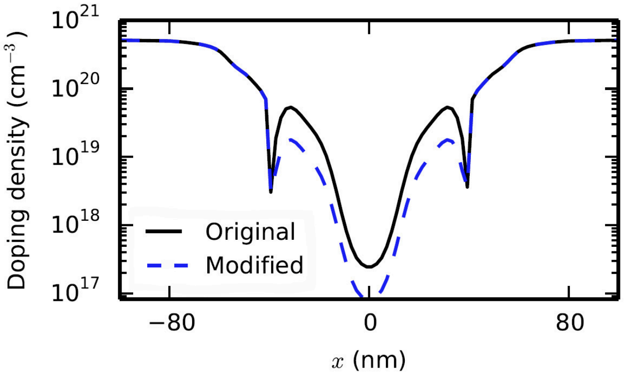

Figure 4.8: Magnitude of the modified net doping profile along the Si/gate-insulator interface of the 40 nm device compared to the original doping profile. The pSDE's have been broadened and their doping concentration has been reduced, to increase the conduction via resonant tunneling [1].

potential pSDE barriers, a result of device scaling, as we have mentioned above. However, this comes at the price of a larger thermionic current.

The absence of any NDT peak (or 'kink') in the 40 nm device at any temperature can be explained from the behavior seen in the smaller devices. Going from the 10 nm to the 30 nm

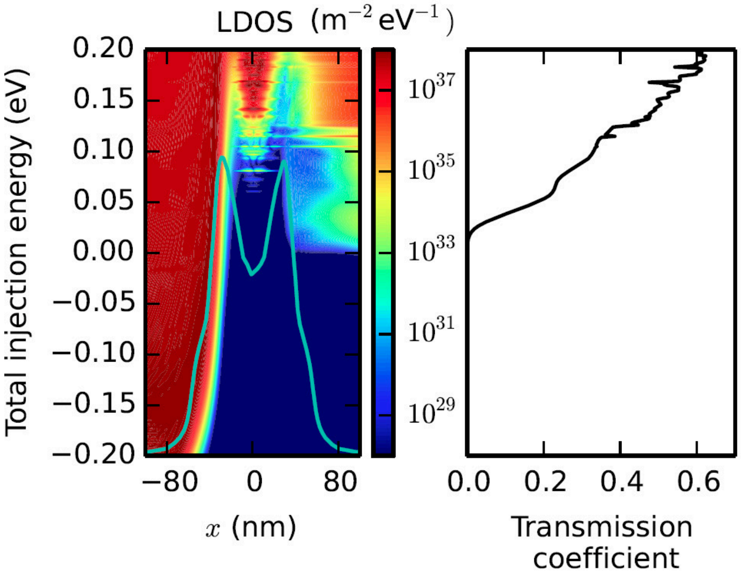

Figure 4.9: Left: Average LDOS in the channel of the 40 nm device with the modified doping profile (shown in Fig. 4.8), at 10 K. The cyan colored line represents the potential-energy profile at the semiconductor/gate-insulator interface. Right: Transmission coefficient *vs.* total injection energy for the particular traveling mode that exhibits the highest resonant transmission. The energy E_m^r of the traveling mode is -0.16 eV and the gate bias $V_{\mathrm{GS}} = 2.0$ V [1].

device, the transmission coefficient peaks (Fig. 4.7) become sharper and lower in magnitude. This can be expected, since scaling the device results in the pSDE potential barriers to become narrower and reduced in height. Narrower barriers cause a broadening of the peaks in the transmission coefficient, as the electron lifetimes in these bound states becomes shorter as a result of the higher probability of 'leaking' out. The transmission coefficient scales inversely and exponentially with the height of the barriers. Already in the 30 nm device, the transmission peaks are very sharp, spanning an extremely small energy-width of the order of 10^{-7} eV, and have low transmission coefficients of the order $10^{-12} - 10^{-7}$. This implies that pSDE barriers in the 40 nm device are too opaque to allow any significant current through the channel of the device. Therefore, no NDT is seen theoretically and the current mostly consists of punch-through current. Moreover, the bound states are energetically grouped closely together. This results in a quasi-continuum of states, a situation that is far from ideal for the generation of NDT at any finite temperature.

The discrepancy between the theoretical predictions and the experimentally observed device characteristics can only be explained by assuming that the doping profile used in the simulations is significantly different from the actual doping profile of the fabricated device. This is a common phenomenon not unlikely to occur in this particular case. As mentioned before, the doping profile used in our theoretical calculations is generated by Sentaurus® TCAD process-simulation tool. The actual doping profile of the fabricated device is generally somewhat different from the TCAD result because of inaccuracies in the modeling of implant dopant diffusion, even if the device is made using industrial CMOS processing standards. Moreover, in this case, the design of the pSDEs is a significant departure from the conventional CMOS design. These sharp, heavily-doped, and highly localized p-type regions may easily be broadened by lateral diffusion enhanced by the defects (especially vacancies) caused by the ion implantation. This is an effect that is notoriously difficult to predict accurately.

In order to assess the sensitivity of the NDT on details of the doping profile of the pSDE, we have modified the original simulated doping profile of the 40 nm device by reducing the peak p-doping by a factor of 3 and by widening laterally the barriers by a factor of 1.8. This is done to roughly simulate the broadening of the pSDEs due to lateral diffusion. The modified doping profile and the corresponding LDOS distribution for a specific gate bias and applied V_{DS} are shown in Fig. 4.8 and Fig. 4.9, respectively. Interestingly, the modified doping profile results in the appearance of bound states, as indicated by the peaks of the transmission coefficient having decent magnitude, and in the occurrence of resonant tunneling. This confirms that the device characteristics are very sensitive to small changes of the profile of the pSDEs. Therefore, we speculate that the devices may behave as intended, with the occurrence of resonant tunneling along the channel through the quasi-bound states in the lateral QW, provided the pSDE doping differs, and not too appreciably, from the original design specifications and from the Sentaurus® simulated profile.

More generally, in order to produce NDT, tunneling through the quasi-bound states in the lateral QW must occur with a probability large enough to overcome the undesired 'leakage paths' of thermionic emission over the pSDE barriers and/or punch-through 'around them'. Peaks of the transmission coefficient in the range of 10^{-2} or larger are required to result in NDT. Clear examples are the first bound state in the 10 nm device and the fourth bound state in the 20 nm device, both giving rise to peaks in the NDT at 10 K. On the contrary, the 3 lowest-energy bound states (20 nm device) have much lower transmission coefficients and do not result in any significant tunneling current, even though the energy spacing between them might be sufficient enough to produce NDT at 10 K. On the contrary, in the 30 nm device, the transmission coefficient peaks are much weaker, of the order of $10^{-12} - 10^{-7}$, and no NDT is seen at low temperature, since the current flows in punch-through, or at high temperatures, because the current now is dominated by thermionic emission over the barriers.

The important question that we need to answer is: What is the maximum gate length for which NDT could be observed at room temperature? And what device design may be required to reach this goal? Performing simulations with different gate lengths, doping profiles, at different temperatures, or even considering different alternative device structures, is an almost impossible task as we scale devices down to the 5 or 7 nm gate length. However, the results we have presented so far allow us to formulate an 'educated guess'. We have already noticed that NDT can be observed when $k_{\mathrm{B}}T \leq DE - dE$, where DE is the spacing of the quasi-bound state in the lateral QW (that is: the energetic spacing of the peaks in the transmission coefficient shown in Fig. 4.6) and dE is the full broadening of the confined states. We have also emphasized the fact that our device-scaling procedure results in a different width and height of the pSDE barriers at different channel lengths. Specifically, Fig. 4.7 shows that DE increases as the devices are scaled, from $\approx$ 20 meV for the 20 nm to 37 meV for the 10 nm device. Also, in these devices $dE \approx 3$ meV (20 nm) and 5 meV (10 nm device). Therefore, at 300 K, 25 meV $= k_{\mathrm{BT}} < DE - dE \approx 32$ meV for the 10 nm device. Thus, it is the thermionic 'leakage path' that hides the expected NDT. Assuming $DE \sim 1/L^2$, where L is the gate length, DE would be as large as 75 or 150 meV in devices scaled to 7 and 5 nm, respectively. This would be more than sufficient to ensure the occurrence of NDT at 300 K. However, as we have already remarked, the possibility of having current flowing via thermionic emission over low pSDE-barriers and via punch-through around high pSDE-barriers would have to be minimized. The latter requirement likely demands that we move from a 'bulk' MOSFET design to ultra-thin-body (UTB) silicon-on-insulator (SOI) structures. This is required by scaling to the 5 nm gate-length, regardless of the particular goal we have in mind. This is already 'conventional' Si CMOS VLSI technology and no major difficulties should be expected. Much more difficult is the constraint posed by the necessity of reducing the thermionic leakage path. This would require narrow and high pSDE barriers. This would have the welcome effect of reducing the energetic width of the quasi-bound states, dE, that scales as $dE \sim \exp(-\alpha W)$, where W is the width of the pSDE barriers, assumed to be proportional to L with proper scaling (α is a quantity that depends on the energy of the quasi-bound-state in the QW and on the electron effective mass in the gap, m_x). Clearly, a barrier-width W of the order of a few nm is hard to envision as achievable by ion implantation. However, the use of larger, slower-diffusion acceptor impurities, such as In in place of B, and use of doping techniques not relying on ion implantation and subsequent rapid-thermal annealing (RTA) steps, such as the low-temperature epitaxy employed for thin-base Si/SiGe epitaxial bipolar transistors [45], could provide a solution if implemented in a 'horizontal' epitaxy, such as the technology used to regrow S/D regions in uniaxially-strained-Si p-channel MOSFETs [46].

We should finally observe that our simulations have been performed assuming ballistic quantum transport. Introducing scattering into the simulation, for example via a Master

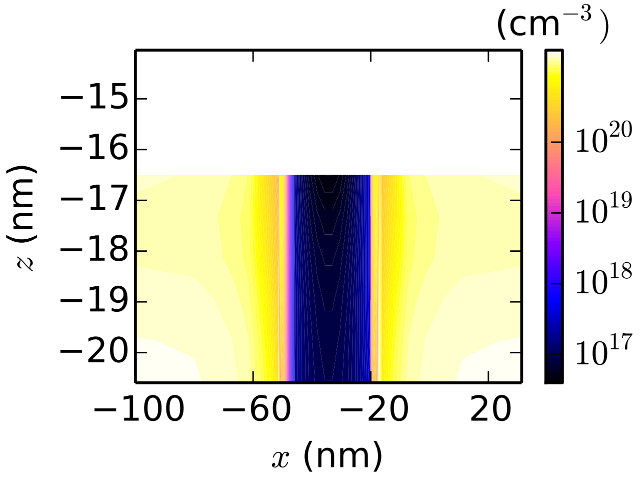

Figure 4.10: Magnitude of the net doping profile of the SOI QW nMOS device with a nominal gate length of 20 nm. The white region at the top represents the gate oxide.

equation [22] or a NEGF approach [47], would lead to a larger broadening dE of the resonant states. Moreover, nonparabolic corrections to the electron dispersion, ignored here, would reduce the energetic separation DE of the transmission resonances. Therefore, our estimates should be considered moderately optimistic, but only ‘moderately’ so.

4.6 Silicon-on-Insulator QW CMOS

As mentioned above, to observe NDT, we need an optimized device design that has potential barriers high enough to prevent thermionic emission, but at the same time, transparent enough to allow sufficient resonant current to pass through. This idea gave us the motivation to simulate a silicon-on-insulator (SOI) QW nMOS device that attempts to meet the above requirements. The doping profile of the device is shown in Fig. 4.10.

The device is found to produce extremely sharp resonant transmission peaks having energy spacing of the order of 1e-7 eV. This is a direct consequence of the high potential barriers in the channel and the absence of body current. Detecting such sharp resonances and

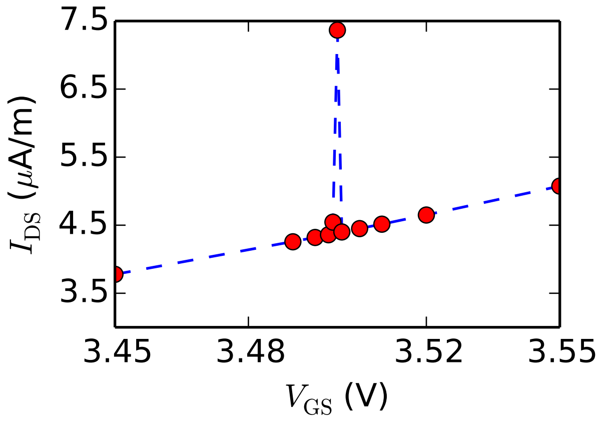

Figure 4.11: Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ characteristics for the 20 nm SOI nMOS at 46 K. $V_{\mathrm{DS}} = 1$ mV

accurately modeling the resulting current using our traditional *open* boundaries Schrödinger solver is time consuming and computationally very heavy. An adaptive integration method is used to modify the Schrödinger solver in an effort to make it more efficient in detecting sharp resonances. The method is as follows: The Schrödinger equation with *open* boundaries is solved for two different energy meshes, one being finer than the other, but both over the same energy range. If the difference between the currents obtained from the two simulations is larger than a pre-determined error estimate, the finer energy range is sub-divided into two or more sub-ranges and the former step is repeated for each of these sub-ranges. The process goes on until the error estimate is satisfied and the final current is obtained as the recursive sum of the current over all the sub-ranges. Using the method on a 20 nm SOI QW nMOS device we are able to observe NDT at a higher temperature (46 K), although room temperature simulations have not produced the desired result. The current-voltage characteristics of the simulated SOI device has been shown in Fig. 4.11.

CHAPTER 5

SIMULATION OF QUANTUM CURRENT IN DOUBLE GATE MOSFETS: VORTICES IN ELECTRON TRANSPORT [1]

Authors – Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti

Department of Electrical and Computer Engineering, EC33

The University of Texas at Dallas

800 West Campbell Road

Richardson, Texas 75080-3021

---

[1] ©[2018] IEEE. Adapted, with permission, from Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti, “Simulation of Quantum Current in Double Gate MOSFETs: Vortices in Electron Transport”, International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Austin, TX, pp. 1-4 (2018).

5.1 Preface

Quantum simulation of electronic transport in double gate (DG) field-effect transistors (FETs) and FinFETs is usually deemed to be required as the devices are scaled to the nanometer length-scale. Here, we present results obtained using our ballistic-transport model, described in Chapter 3 in these devices. Our quantum simulations show the presence of quasi bound electronic states in the channel and Fano-interference phenomenon in the transport behavior of ultra-thin body (UTB) Si DG MOSFETs. Vortices in electron wave-functions are also reported at energies at which transmission zeros (antiresonance) occur.

The contents of this chapter are taken with permission from "Simulation of Quantum Current in Double Gate MOSFETs: Vortices in Electron Transport" [2018 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Austin, TX, pp. 1-4 (2018)] Copyright (2018) IEEE. The authors are Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti. I developed the simulation tool, generated and analyzed the result, and wrote the manuscript. Dr.. Maarten L. Van de Put provided valuable guidance in developing the simulation tool, helped with analyzing the results as well as improve the manuscript content. Dr. Massimo V. Fischetti provided critical guidance in developing the physical model, understanding the observed phenomenon, and helped to improving the manuscript content. Portions of the manuscript have been excluded to maintain the flow of the dissertation.

5.2 Introduction

Ultra-thin body (UTB) double gate (DG) MOSFETs have a strong potential to overcome short channel effects [48] and thus have superior scalability in comparison to conventional MOSFETs. Moreover, these device structures provide a significant improvement in performance [49] in terms of low subthreshold slope, high ON current and high switching speed.

This makes them very attractive for current and near future generations of silicon (Si) semiconductor devices. Simulation of these devices is usually done using semi-classical approaches based on the solution of Boltzmann transport equation (BTE) [50, 8], moments of BTE [51] or compact models augmented by quantum corrections [52, 53]. However, a full quantum treatment of these devices is necessary, since at such small dimensions explicit quantum effects would become observable. To this end, we use our simulation tool, described in Chapter 3, to model ballistic quantum transport in these devices.

The most striking result that we obtain is the occurrence of the Fano-interference phenomenon [54] in the simulated UTB DG FETs. Bowen et al. [55] have shown that the Fano resonance-antiresonance line shapes can be accurately represented by poles and zeros, respectively, of the inverse of the retarded Green's function representing the system Hamiltonian (tight-binding, in their study) connected to infinite reservoirs. They have presented an efficient numerical method, based on a shift-and-invert non-symmmetric (SINS) Lanczos algorithm, to locate the poles and zeros, mentioned above, in single-barrier GaAs/AlAs/GaAs heterostructures. Fano interference has also been previously predicted and/or experimentally reported in the optical absorption spectra of impurities in crystals [56], quantum waveguides [57], and coupled quantum dot systems [58, 59]. Our observation of this resonance in a realistic CMOS device structure thus presents a novel and interesting case. Additionally, the simulated device exhibits only symmetric antiresonance 'dips' in electron transmission, contrary to the characteristic asymmetric Fano resonance-antiresonance line-shape observed in all the former cases. Moreover, vortices in current density are seen at energies at which antiresonance occurs. Such vortices have been previously reported in quantum simulations of semiconductor devices, but only in the presence of deviations from ideality – scattering with discrete dopant atoms [60] or tapered and bent semiconductor channels [61].

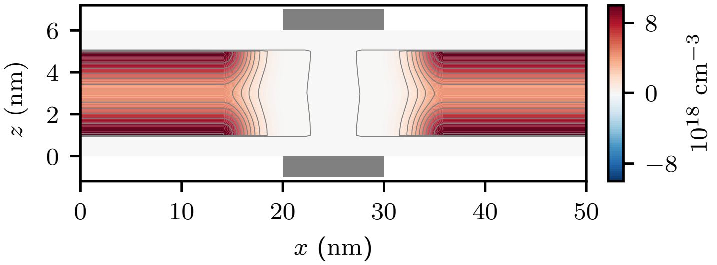

Figure 5.1: Net doping profile of the 10 nm UTB DG nMOS that we have studied. The white regions at the top and bottom represent the 1 nm thick gate oxide while the grey patches are used to highlight the position of the gate terminals.

5.3 Device Description

We simulate the transfer characteristics of a Si (UTB) DG nMOS with channel length of 10 nm (Fig. 5.1). The device (simulation region) is 4 nm thick (τ_{Si}) and 50 nm long with symmetric 1 nm (EOT $\approx$ 0.3 nm) oxide at each gate. The channel is lightly p-type doped ($\approx 10^{15}$ cm^{-3}), while the highly doped n-type source and drain regions are modeled using a dual Gaussian profile with peaks located at the two oxide-semiconductor interfaces. The device behavior is observed under the application of equal gate bias (V_{GS}) at the two gates with a low drain-to-source bias ($V_{\mathrm{DS}} \approx 20$ mV). The channel orientation is taken along the [110] direction, following the general trend in VLSI technology.

We solve the Schrödinger and Poisson equations self-consistently in the 2-D plane of the device to analyze the device behavior under applied bias. We have used Neumann boundary conditions on the electrostatic potential, as described in Sec. 3.3.2 to avoid charge imbalance at the edges. Although physically inconsistent, the method is computationally less expensive and provides results similar to other methods described in Sec. 3.3.1.

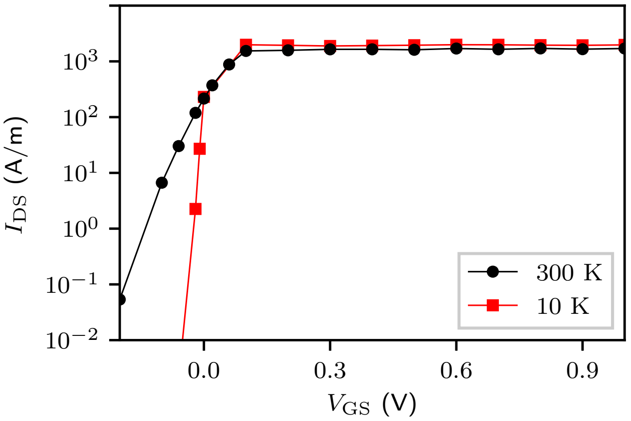

Figure 5.2: I_{DS}-V_{GS} characteristics of a 10 nm UTB DG nMOS at 10 K and 300 K. V_{GS} is measured with respect to the flat band voltage of the device. V_{DS} =10 mV.

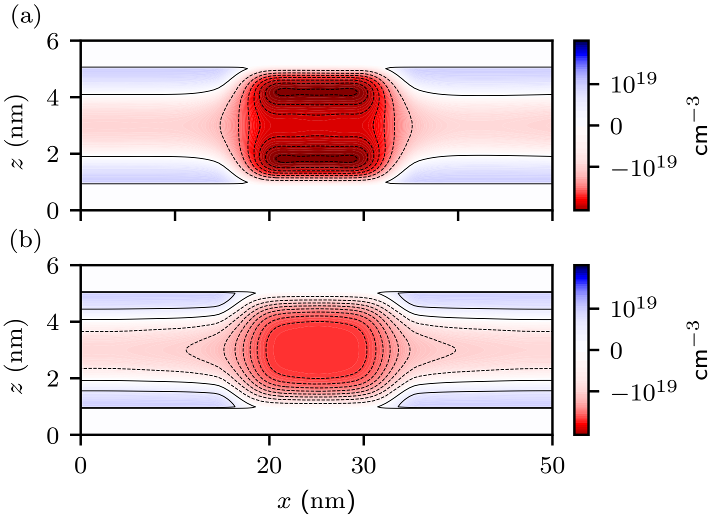

Figure 5.3: Total charge distribution in a 10 nm UTB DG nMOS at 300 K. (a) The dark red regions show the creation of two separate inversion channels deep inside saturation, whereas in (b) volume inversion is seen in the linear region of operation. V_{GS}=0.6 V, V_{DS} =10 mV.

5.4 Simulation Results

Fig. 5.2 shows the simulated I_{DS}-V_{GS} characteristics of a 10 nm DG nMOS at 10 K and 300 K with equal V_{GS} applied at both gates. Characteristic CMOS behavior is observed with fast switching action represented by a low subthreshold slope (≈ 64 mV/dec at 300 K). For device operation deep inside saturation, the charge-distribution plot in Fig. 5.3(a) shows the occurrence of channel inversion, whereas volume inversion is observed in the linear region of operation (Fig. 5.3(b)). The current-density distribution, plotted in Fig. 5.5(a) for a V_{GS} deep inside saturation, illustrates the path followed by the current. Electrons are injected in a single centered beam at the source (drain), splitting into two when flowing through the inversion channels, to finally merge at the drain (source).

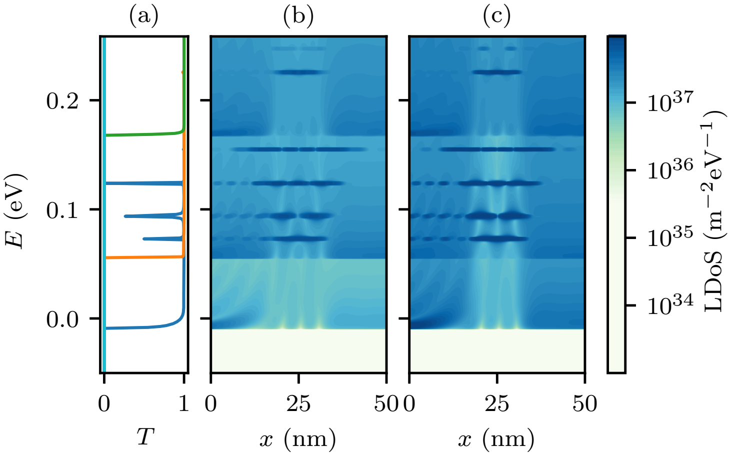

Figure 5.4: Transmission coefficient T *vs.* injection energy (a) for current injection from the source contact in a 10 nm UTD DG nMOS at 300 K. The different colored lines correspond to the different injected subbands. The energies are measured with respect to source Fermi level. The LDoS distribution averaged over a cross-sectional thickness of roughly 1.3 nm in the top channel (b) and middle of device (c) at 300 K for injection from the source contact. V_{GS} =1.6 V, V_{DS} =10 mV.

Fig. 7.5 shows the average LDoS distribution along a cross-section of one of the channels and the center of the device for different injection energies (from the source contact). The darker regions exhibit the presence of quasi bound states created within the channel region as a result of the 2-D and field-induced confinement. An interesting feature is the presence of sharp dips in the transmission coefficient (T) observed at these bound-state energies, as shown in Fig. 7.5(a). Conventionally, a sharp peak in transmission is expected at the resonant energies. On the contrary, the dips in transmission signify occurrence of antiresonances

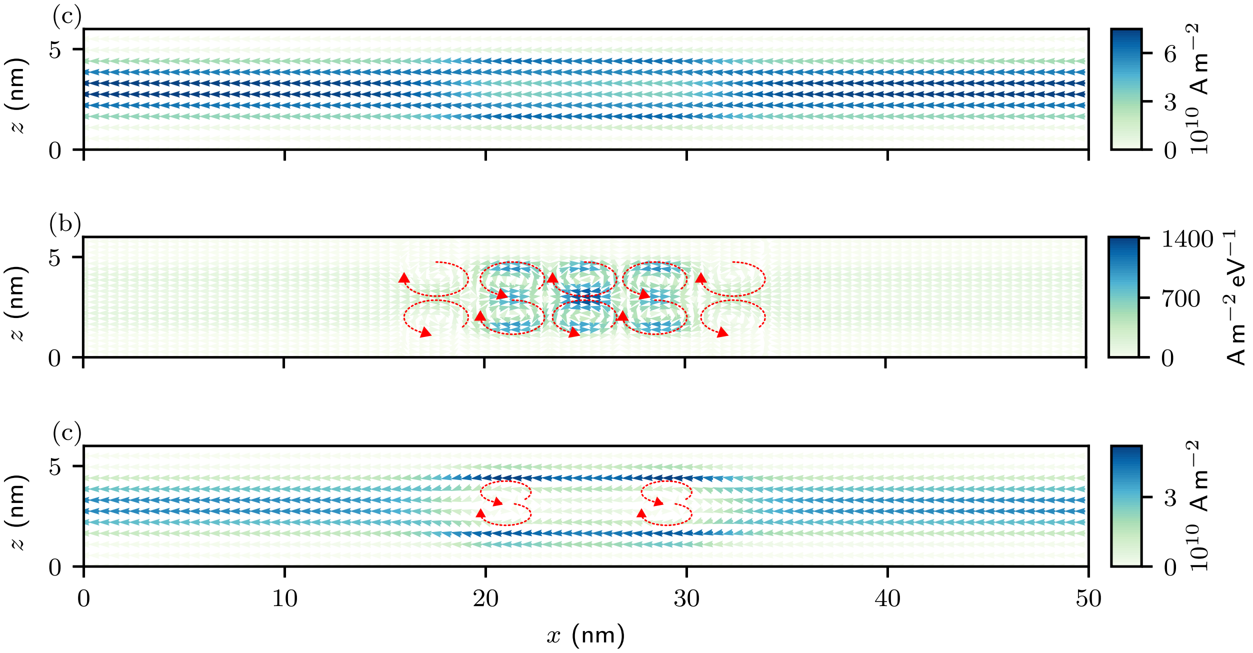

Figure 5.5: (a) Current density distribution in a 10 nm UTB DG nMOS at 300 K. (b) Current density distribution resolved for a single injection energy at 300 K. The energy is chosen to have the value at which a sharp dip in transmission occurs due to destructive interference at resonance. V_{GS} =0.6 V, V_{DS} =10 mV. (c) Current density distribution in the same device at 10 K. V_{GS} =0.6 V, V_{DS} =10 mV. The red arrows highlight the direction of the vortices. The plots are stretched to match the aspect ratio of the device.

caused by the interaction, or configuration interaction as termed by Fano [54], between quasi-bound states in the channel and the continuum of injected states from the source and drain. To understand the phenomenon qualitatively, we take into consideration Fano's argument [54] which states that waves transmitted at resonating frequencies undergo a phase shift as well as a change of magnitude. Indeed, in the DG nMOS, we observe this resonance for electrons injected into the two degenerate quasi-bound states in the two inversion channels: The two paths undergo opposite phase-shifts that brings them out-of-phase, resulting in destructive interference. Hence we see the antiresonance dips of the transmission probability at the bound-state energies. The fact that antiresonance is seen only in the presence of channel inversion gives further confirmation of our interpretation. Also, as mentioned before in Sec. 5.2, we observe symmetric transmission zeros in our case, contrary to the asymmetric resonance-antiresonance line shapes reported in other studies [56, 57, 58]. This is because the DG FET structure in saturation mode of operation is analogous to a system of coupled oscillators in which both oscillators (inversion channels in our case) are driven by an external force (V_{DS} in our case), as compared to only one driven oscillator in the latter studies.

Theoretically, one expects to see drops in the total drain current at those values of the gate bias for which the Fermi level of the device crosses the energy of one of the antiresonance-producing bound states. However, as Fig. 7.5(a) illustrates, the antiresonance features are extremely sharp and thermal smearing prevents them from appearing in the current-voltage (I-V) characteristics of Fig. 5.2 at 300 K, and even at 10 K.

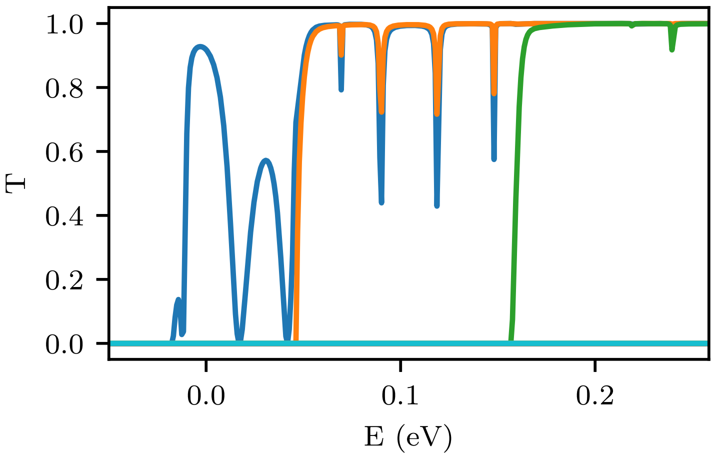

Figure 5.6: Antiresonance features in the transmission spectra broaden in the presence of an asymmetric gate bias. The potential difference between the gates is 0.4 V. Temperature = 300 K, V_{DS} =10 mV.

Moreover, circulations are seen in the current density resolved for the individual injection energies (from the source contact) at which antiresonance occurs, as illustrated in Fig. 5.5(b). The even number of vortices, formed as a result of destructive interference, leads to negligible transmission of current at the resonating energies, while momentum conservation forces almost all the electrons to reflect back to the injecting lead. It is important to mention here that the total current density at 300 K does not exhibit vortices, while faint vortices persist in the total current density at 10 K, as highlighted in Fig. 5.5(c). At higher temperatures, the wider energy window resulting from a Fermi Dirac distribution masks the contribution of the resonating states to the total current density.

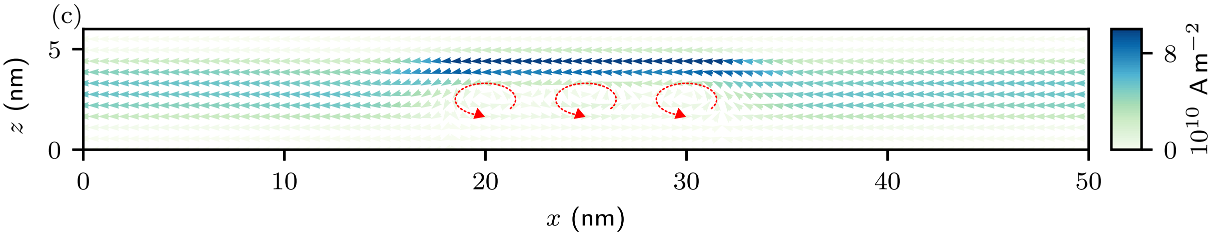

Figure 5.7: Current density distribution in the device with asymmetric gate bias at 300 K. The red arrows highlight the direction of vortices in the current density. Potential difference between gates is 0.4 V for the latter case.

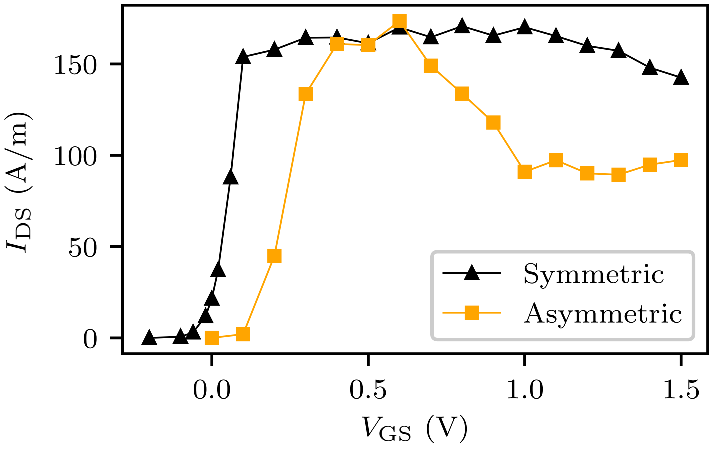

Figure 5.8: Comparison of the transfer characteristics of the device with equal (symmetric) and unequal (asymmetric) gate biases applied at 300 K. Potential difference between gates is 0.4 V for the latter case.

When asymmetry is introduced by applying an unequal gate bias in this case, the transmission dips broaden and vortices in the current appear even at room temperature, as can be seen in Figs. (5.6) and (5.7), respectively. At the same time, there is a noticeable decline in the drain current (Fig. (5.8)) in presence of the unequal gate biases, demonstrating that asymmetry hinders device performance. We can also conjecture that, under the right conditions, the Fano phenomenon in these FETs can be observed experimentally.

CHAPTER 6

QUANTUM MECHANICAL STUDY OF IMPACT OF SURFACE ROUGHNESS ON ELECTRON TRANSPORT IN ULTRA- THIN BODY SILICON FETS [1]

Authors – Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti

Department of Electrical and Computer Engineering, EC33

The University of Texas at Dallas

800 West Campbell Road

Richardson, Texas 75080-3021

---

[1] ©[2018] IEEE. Adapted, with permission, from Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti, "Quantum Mechanical Study of Impact of Surface Roughness on Electron Transport in Ultra- Thin Body Silicon FETs", IEEE 13th Nanotechnology Materials and Devices Conference (NMDC), Portland, OR, pp. 1-4 (2018).

6.1 Preface

We present a theoretical study – based on the effective-mass approximation – of scattering of electrons with microscopic roughness at the oxide-semiconductor interface for ultra-thin body (UTB) double gate (DG) and silicon-on-insulator (SOI) MOSFETs. Discrete edge roughness is introduced along the oxide-semiconductor interface by locally adding or removing 1-2 silicon atomic layers. The random nature of the interface roughness is characterized by an exponential autocovariance. The dissipative impact of surface roughness scattering is obtained by averaging over a sample space of devices with identical structure but with different random interface-roughness patterns. Our results show significant variations in the ON current among the 'roughened' samples along with the introduction and impact of quantum confinement effects caused by interface roughness. Finally, we report the dependence of these results on the semiconductor/channel thickness.

The contents of this chapter are taken with permission from "Quantum Mechanical Study of Impact of Surface Roughness on Electron Transport in Ultra-Thin Body Silicon FETs" [2018 IEEE 13th Nanotechnology Materials and Devices Conference (NMDC), Portland, OR, pp. 1-4 (2018)] Copyright (2018) IEEE. The authors are Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti. I developed the simulation tool, generated and analyzed the result, and wrote the manuscript. Dr. Maarten L. Van de Put provided valuable guidance in developing the simulation tool, helped with analyzing the results as well as improving the article content. Dr. Massimo V. Fischetti provided critical guidance in developing the physical model, understanding the observed phenomenon, and helped to improve the article content. Portions of the manuscript have been excluded to maintain the flow of the dissertation.

6.2 Introduction

The nature of the interface between the gate insulator and semiconductor may significantly affect the electronic properties of the semiconductor devices as they are scaled down aggressively to the nanometer regime. The impact may be more pronounced for promising device structures like the double gate (DG) FET, silicon-on-insulator (SOI) FET, FinFETs and likewise with ultra-thin bodies (UTB). Scattering of electrons with microscopic roughness at the oxide-semiconductor interface has been modeled at different levels of approximation to provide a quantitative picture of this impact. Statistical analysis of the effect of surface roughness (SR) scattering has been done by modeling geometrically roughened interfaces of Si FETs using classical drift-diffusion[62] and Monte Carlo methods[63] and of Si nanowires (NW) using Non-equilibrium Green's function (NEGF)[64] and tight-binding[65] based approaches. A first-principles study based on density functional theory (DFT) coupled with solution of the linearized Boltzmann transport equation to model the impact of atomic-scale roughness on electron mobilities in strained silicon (Si) inversion layers has also been reported[66, 67]. Fischetti et al. [68] has proposed a local empirical pseudopotential based approach to calculate SR scattering rates that aims to achieve the physical accuracy of the ab-initio models without the added computational cost. Also pervasive in literature are models [69, 70, 71] that treat fluctuations of the oxide barrier at the oxide-semiconductor interface as a perturbation potential. The idea was first proposed by Ando [29] for the case of bulk Si CMOS and later extended to thin bodies [72] and SOI FETs [73, 74].

Here we follow the former 'ab-initio' method of modeling surface roughness as random physical shifts of the oxide-semiconductor boundary along the length of semiconductor channel. The randomness of the interface-roughness patterns follows an exponential autocorrelation, in agreement with a critical statistical study done by Goodnick et al. [75] for Si-SiO$_2$ interface. Dissipation due to SR scattering is determined by averaging over a statistical ensemble of devices with identical structure but random interface-roughness patterns. The ballistic transport of electrons in these devices is modeled using our simulation tool described in Chapter 3.

6.3 Device Description and Simulation of Interface Roughness

We analyze the transport characteristics of Si UTB DG and SOI nMOS with channel lengths of 10 nm in the presence of SR scattering. The devices (simulation region) are 4 nm wide (W) and 50 nm long with lightly p-type doped ($\approx 10^{15}$ cm^{-3}) channel and symmetric 2.1 nm (equivalent oxide thickness EOT $\approx$ 0.7 nm) oxide at the top and bottom interfaces. The highly doped n-type source and drain regions are modeled using Gaussian profiles with peaks located at the oxide-semiconductor interfaces, as shown in Fig. 6.1. The device behavior is observed at room temperature under the application of different gate biases (V_{GS}) (equal gate biases applied at the two gates for the DG nMOS case) and constant drain-to-source bias ($V_{\mathrm{DS}} \approx 100$ mV). The range of V_{GS} is chosen such that it includes sufficient bias points in all three regions of operation — accumulation, inversion and saturation. The channel orientation for both device structures is taken along the [110] direction while the oxide-semiconductor interface is assumed along the [100] plane, following general trends in VLSI technology.

The exact nature of the Si-SiO$_2$ interface is still an open topic of debate. Traditionally, a Gaussian autocovariance has been used to characterize the interface. However, Goodnick et al. [75] has shown that an exponential autocovariance more closely resembles the nature of the interface and therefore we have used the latter for our simulations. The corresponding power spectrum $\langle|\Delta_{\mathrm{Q}}|^2\rangle$ is given by the expression:

$$\langle|\Delta_{\mathrm{Q}}|^2\rangle = \frac{\pi\Delta^2\Lambda^2}{(1+q^2\Lambda^2/2)^{\frac{3}{2}}} \tag{6.1}$$

where Δ is the root-mean-square (rms) roughness, Λ is the correlation length and q is the scattered wave vector from the Fourier transform of the interface shifts. For this work,

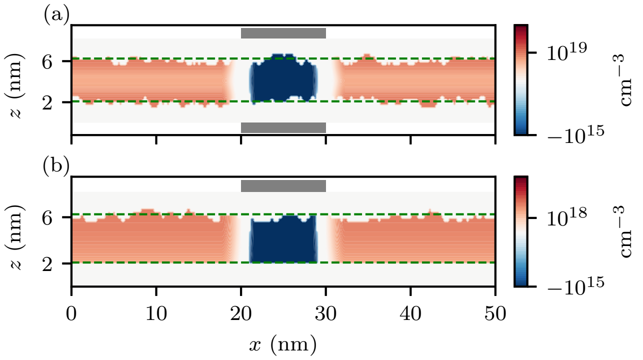

Figure 6.1: Doping profile of (a) DG and (b) SOI nMOS with surface roughness. The white region at the top and bottom represents the gate oxide while the grey patches are used to highlight position of the gate terminals. Green dashed lines are used to mark the oxide-semiconductor interface of the ideal devices.

we have taken $\Delta = 0.3$ nm and $\Lambda = 1.5$ nm as reported in [75]. A complex function is assigned to each interface mesh point, having magnitude equal to the square root of the power spectrum Eq. (7.22) and a randomly generated complex phase. The autocorrelation function characterizing the interface roughness is obtained by doing inverse Fourier transform of the resulting complex function. The interface-roughness pattern is obtained by discretizing the autocorrelation function into steps, such that each discretized value corresponds to the addition or removal of $0 - 2$ silicon atomic layers. The step size is assumed equal to the thickness of a Si atomic layer (or equivalently distance between two Si lattice points) oriented along the [100] plane (≈ 0.27 nm) and the step length is taken equal to the distance between consecutive lattice points along the [110] direction (≈ 0.19 nm). The doping distributions of 4 nm wide DG and SOI nMOS with interface roughness are shown in Fig. 6.1. Separate roughness patterns that are anti-correlated with respect to each other are applied to the top and bottom oxide-semiconductor interfaces of the DG nMOS. For the SOI nMOS, the bottom oxide-semiconductor interface is assumed perfectly smooth, reflecting the relatively uniform formation of the buried oxide layer using the well-known SIMOX process. Finally, we comment on the use of a 2-D device model: The averaging effect of projecting the complete

2-D oxide-roughness on a 1-D cross-section will only lead to a weakening of the impact of SR scattering [75]. Therefore our use of a 2-D device model is a reasonable approximation for studying the phenomenon of SR scattering.

We solve the *open* system using the method described in Sec. 3.3.2. Though physically inconsistent, the method is computationally less expensive and provides results similar to other methods described in Sec. 3.3.1. A statistical ensemble of 50 'roughened' samples are simulated using this model for each device structure, to demonstrate the dissipative impact of SR scattering.

Table 6.1: Statistical Analysis

FET structure	**I_{ON}** (A/m)			**V_T** (V)			**SS$_{min}$** (mV/dec)			**W_{eff}** (nm)			**W_{min}** (nm)		
	Ideal	Mean	SD	Ideal	Mean	SD	Ideal	Mean	SD	Ideal	Mean	SD	Ideal	Mean	SD
4 nm DG	1869	1185	154	1.05	1.06	0.005	64.1	63.7	0.58	4	2.62	0.237	4	3.057	0
4 nm SOI	1836	1424	164	1.0	1.02	0.003	80.4	80	0.87	4	3.17	0.255	4	3.61	0
8 nm DG	4001	3448	202	1.0	1.065	0.006	77.4	77	0.35	8	7.08	0.31	8	7.128	0

6.4 Simulation Results

Simulation results are shown for 4 nm and 8 nm wide DG FETs as well as 4 nm wide SOI FET. Fig. 6.2(a) compares the transfer characteristics of the three devices with (dashed lines) and without interface roughness (solid lines), focussing primarily on the 'ON' state of the transistor. A notable decrease in the mean ON current I_{ON} is observed in the roughened devices (see also Table 6.1) when compared to their ideal counterparts. The percentage decrease (≈37%) is highest for the 4 nm DG FET while the impact is relatively weaker in 4 nm SOI (≈22%) and 8 nm DG (≈13.8%) FETs. A reduced impact of SR scattering in the SOI case can be expected due to presence of roughness at one interface only. An interesting point to note is that, although the percentage decrease in current is much smaller for the 8 nm DG FET compared to its 4 nm counterpart, the net decrease in current is similar for both. To investigate this further, we plot the current-density distribution in Fig. 6.3 for the

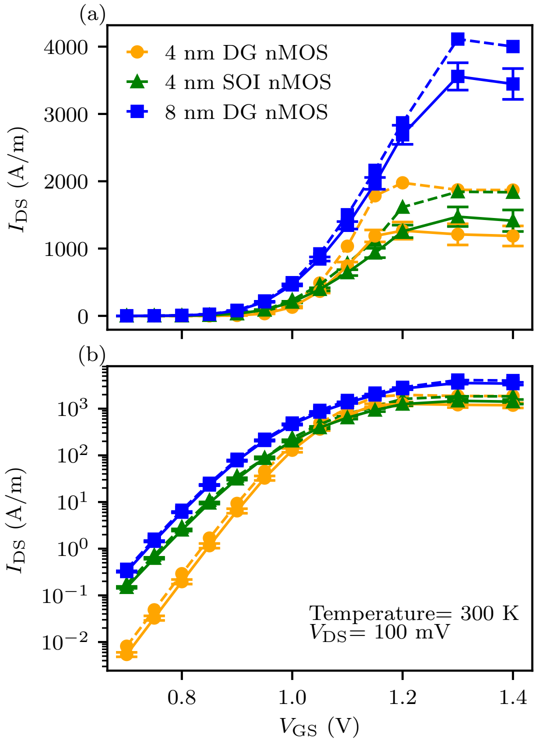

Figure 6.2: Transfer characteristics of the simulated devices in (a) linear and (b) semi-log scale to highlight the impact of SR scattering in above-threshold and subthreshold regions, respectively. The dashed lines represent ideal device behavior while solid lines along with error-bars are used to plot statistical mean and standard deviation of the behavior of 'roughened' samples, respectively.

two devices under identical bias conditions. We can see that, even though the current flows deep within the body of the 4 nm device, it is still affected by the interface roughness. On the contrary, this is not observed in the wider device, since only the current near the interface is affected by the roughness. Thus, for channels wider than 4 nm, decrease in current becomes independent of channel thickness due to SR scattering becoming a predominantly superficial effect. Fig. 6.3(a) also shows presence of vortices (highlighted) in the current.

The subthreshold behavior of the device remains unaffected by the presence of interface roughness, as can be seen in Fig. 6.2(b). The shift in threshold voltage V_{T} and minimum subthreshold slope ($\mathrm{SS}_{\mathrm{min}}$), also evident in Table 6.1, is minimal in the presence of SR scattering for all three simulated structures. Thus, SR scattering impacts electron transport only in the presence of a substantial current.

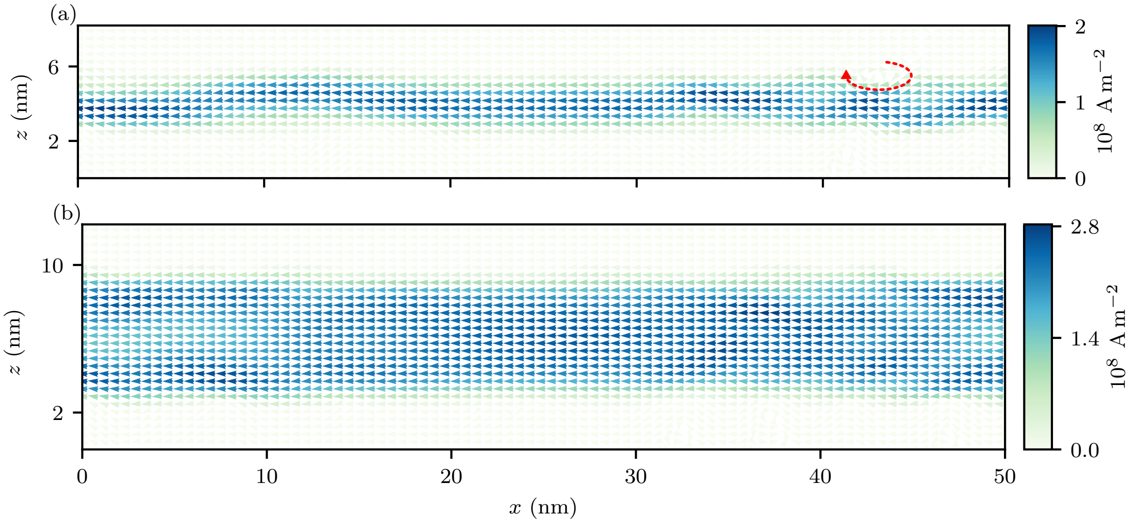

Figure 6.3: Current density distribution in (a) 4 nm and (b) 8 nm DG nMOS at 300 K. The plots show that SR scattering mostly affects the electron transport close to the oxide-semiconductor interface. A vortex in the current density is highlighted by a red curved arrow. V_{GS} =0.95 V, V_{DS} =100 mV.

Next, we define the effective width W_{eff} of a 'roughened' sample as the equivalent width of an ideal sample that will exhibit identical current-voltage (I-V) characteristics as the former. The quantity is determined by fitting the I-V plot of a roughened sample to that of an ideal device and assuming linear scaling of the current with the device width. Classically, the minimum cross-sectional width W_{min} bottlenecks the total current flowing through the device and therefore its value should be close to W_{eff} of the sample. However, Table 6.1 shows that W_{eff} is substantially smaller than W_{min} for the 4 nm devices. The apparent discrepancy can be explained as the result of quantum confinement effects. Random contractions of channel width, caused by interface roughness, increase vertical quantum confinement and consequently the spacing between individual electronic subbands as well. The net effect is a further reduction of current flow, as now, less subbands conduct. The significant reduction of this effect for the wider 8 nm device lays further foundation to the argument.

Finally, we see that the SOI FET, compared to DG FET of similar size, has higher mean I_{ON} in the presence of interface roughness (almost identical I_{ON} in ideal case) and is less affected by SR scattering (mean I_{ON} closer to ideal value). Thus, whereas it exhibits a slightly higher $\mathrm{SS}_{\mathrm{min}}$, the SOI FET appears as a better candidate, compared to DG FET, for UTB structures when SR scattering is concerned.

CHAPTER 7

DISSIPATIVE QUANTUM TRANSPORT: MASTER EQUATION APPROACH

7.1 Introduction

In an *open* system driven far from equilibrium, scattering between electronic states is a major factor that drives the system towards steady state. Thus, the introduction of dissipation into our purely ballistic approach should have a noticeable impact, providing us a more realistic picture of the device behavior. Here, we follow a less conventional approach [21, 22] based on the Pauli master equation (PME) to treat scattering. This Markovian class of Master equations is used to describe the transition between quantum states and, thereby, the time evolution of an irreversible *open* system. In similar fashion, the PME can be used to express the time evolution of the density matrix of a device connected to external reservoirs. Coupling between the device and reservoirs, which are assumed to be in thermal equilibrium, is established phenomenologically. Additionally, Ref. 21, following Van Hove [76], shows that for a device smaller than the phase coherence length, the injected electrons are highly delocalized and therefore the off-diagonal terms of the density matrix, that are responsible for the interference between the injected states, can be neglected. This assumption is vital to enable us to numerically model relatively large realistic *open* systems efficiently. We determine the scattering states that diagonalize the density matrix by solving the *open* system Schrödinger equation in the two-dimensional (2-D) plane of the device, as described in Sec. 3.1. We also employ our technique described in detail in Sec. 3.3.1 to model accurately the continuous density-of-states of the *open* system. Transition probabilities between the quantum states, determined using Fermi's golden rule, are then incorporated into the PME to obtain the steady-state distribution of electrons away from the purely ballistic picture. It is important to note that we consider the system to be at steady-state, even in the presence of an external force (applied bias) causing exchange of particles with the environment. The Schrödinger equation and the PME are solved self-consistently with the Poisson equation to obtain the final state of the system.

It is worth mentioning that the natural basis of *open* system wavefunctions the PME makes use of, contrary to plane waves used in the BTE, allows us to include quantum mechanical processes like tunneling and confinement effects. At the same time, these solutions contain the requisite information regarding the electric-field distribution in the device. Therefore, the assumption of the BTE, that the external electric field must vary slowly over the electron mean free path, is no longer required, giving us a more practical picture of the system without the semiclassical concept of point-like electrons. Thus the PME could correctly and efficiently describe electronic transport in sub-50 nm devices, considering the fact that the electron dephasing length can be of the order of 50 nm for Si devices at room temperature.

In this work, we incorporate the collisions between electrons and ions, excited away from equilibrium, namely phonons into our transport studies. Up until now we have used the Born-Oppenheimer approximation assuming that the ions are frozen in their equilibrium positions and only interact with electrons through the Hartree potential. We also study the impact of scattering of electrons with roughness present at dielectric-semiconductor. We start by giving a general description of the PME approach and then provide details on how we determine the scattering rates for the different scattering mechanisms considered here, namely electron-phonon and SR scattering, and their impact on the semiconductor-device behavior. As an application, dissipative transport in ultra-thin body (UTB) FETs is simulated.

7.1.1 Scattering between electronic states: Solution of Pauli master equation

We solve the PME using the *open* system wavefunctions $\phi^{s,v}_{m,\beta}$ (considering only the non-evanescent modes) obtained in Sec. 3.1. We start with the Louiville-von Neumann equation of motion for the single-electron reduced density matrix ρ of the system [21], having traced over the phonons:

$$\frac{\partial \rho(t)}{\partial t} = \frac{i}{\hbar}[\rho, H] + \left(\frac{\partial \rho}{\partial t}\right)_{\mathrm{res}} \quad , \tag{7.1}$$

where the second term on the right accounts for the effect of the reservoirs. $\boldsymbol{H}$ here is the electron, ion, and electron-ion Hamiltonian. The density matrix ρ on the basis of the eigenstates of the Hamiltonian $\boldsymbol{H}$ can be written as:

$$\rho_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}} = \left|m\beta k_{\mathrm{y}}\right\rangle \left\langle m'\beta' k'_{\mathrm{y}}\right| \ , \tag{7.2}$$

A general formulation of dissipative transport should include time evolution of both the diagonal and off-diagonal terms of the density-matrix. Solving the transport equation for the full density-matrix is a daunting task. However, in Refs. 21 and 22 it was shown that the off-diagonal terms of the density matrix can be ignored under certain cases and Eq. (7.1) can be simplified to the following expression for the PME:

$$\frac{\partial \rho_{m\beta k_{\mathrm{y}}}}{\partial t} = \sum_{m\beta k_{\mathrm{y}} \neq m'\beta' k'_{\mathrm{y}}} \left[W_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}}\, \rho_{m'\beta' k'_{\mathrm{y}}} - W_{m'\beta' k'_{\mathrm{y}},m\beta k_{\mathrm{y}}}\, \rho_{m\beta k_{\mathrm{y}}}\right] + \left(\frac{\partial \rho_{m\beta k_{\mathrm{y}}}}{\partial t}\right)_{\mathrm{res}} \quad , \tag{7.3}$$

where $W_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}}$ represents the probability per unit time for an electron to make a transition from state $\left|m'\beta' k'_{\mathrm{y}}\right\rangle$ to state $\left|m\beta k_{\mathrm{y}}\right\rangle$. The symbol $\rho_{m\beta k_{\mathrm{y}}}$ is used to represent the diagonal matrix element $\rho_{m\beta k_{\mathrm{y}},m\beta k_{\mathrm{y}}}$ for simplicity. The expression for $W_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}}$ is given by Fermi's golden rule:

$$W_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}} = \frac{2\pi}{\hbar}|\alpha|^2 | \left\langle m\beta k_{\mathrm{y}}\middle|H'\middle|m'\beta' k'_{\mathrm{y}}\right\rangle |^2\, \delta\Big(E_{m\beta k_{\mathrm{y}}} - E_{m'\beta' k'_{\mathrm{y}}}\Big) \ , \tag{7.4}$$

where H' is defined as $H_{\mathrm{int}} = \alpha H'$, α being a dimensionless constant representing the strength of the interaction. The calculation of W accounts for all intravalley and inervalley transitions. Account is made for all transitions between states injected from different leads s, also termed back-scattering, which forms a major scattering mechanism for electrons. However, to avoid notational complexity, a separate "lead index" will not be introduced explicitly to reflect this mechanism.

We will not go in depth into the derivation of the above equation, as Ref. 22 describes it in detail. However, it is important to mention the assumptions taken in order to obtain Eq. (7.3) and to understand the conditions of its validity. The off-diagonal terms of the density matrix are responsible for the quantum interference and consequently localization of plane waves into wave packets. Devices with active regions much smaller than the dephasing length of electrons will see the incoming waves as totally delocalized and therefore the off-diagonal terms are not injected by the contacts. We will assume the limiting case of contacts injecting wavefunctions of the form given by the QTBM, Eq. (3.4). Moreover, we will restrict ourselves to the weak scattering limit ($\alpha^2 \ll 1$), so that the off-diagonal terms of the density matrix are not generated within the device in the course of reaching steady-state. Under the premise that the contacts inject plane waves, the reservoir interaction term of Eq. (7.3) can be solved phenomenologically [22]. The reservoirs attempt to restore charge neutrality by injecting electrons, with traveling wavevectors $k^{r,v}_{m,\beta}$ through the corresponding leads r, into the states that are transmitted and absorbed by these ideal contacts. Mathematically, this can be expressed as:

$$\left(\frac{\partial \rho_{m\beta k_{\mathrm{y}}}}{\partial t}\right)_{\mathrm{res}} = \frac{\hbar k^{s,v}_{m,\beta}}{m^v_{\mathrm{x}}} |A^{s,v}_{\beta}|^2 \Big[f_{\mathrm{FD}}(E^v_\beta + E_{\mathrm{y}}) - \rho_{m\beta k_{\mathrm{y}}, m\beta k_{\mathrm{y}}} \Big] \,, \tag{7.5}$$

where the term outside the parentheses on the the right hand side describes the group velocity of the incident electron wave, $|A^{s,v}_{\beta}|^2$ is the normalization constant used for $\phi^{s,v}_{m,\beta}$ and f_{FD} is the Fermi-Dirac distribution describing the occupation of the states at thermal equilibrium in the leads. At steady-state, Eq. (7.3) becomes:

$$\begin{aligned} \sum_{m\beta k_{\mathrm{y}} \neq m'\beta' k'_{\mathrm{y}}} & \Big[W_{m\beta k_{\mathrm{y}}, m'\beta' k'_{\mathrm{y}}}\, \rho_{m'\beta' k'_{\mathrm{y}}} - W_{m'\beta' k'_{\mathrm{y}}, m\beta k_{\mathrm{y}}}\, \rho_{m\beta k_{\mathrm{y}}} \Big] \\ &= -\frac{\hbar k^{s,v}_{m,\beta}}{m^v_{\mathrm{x}}} |A^{r,v}_{\beta}|^2 \Big[f_{\mathrm{FD}}(E^v_\beta + E_{\mathrm{y}}) - \rho_{m\beta k_{\mathrm{y}}} \Big] \,. \end{aligned} \tag{7.6}$$

This is a linear system consisting of M such equations, where M is the total number of distinct electronic states $|m\beta k_{\mathrm{y}}\rangle$ (states injected from different leads counted separately) or, equivalently, the rank of the diagonal density-matrix. Any generic linear system solver can be used to obtain the diagonal terms of the density-matrix.

Electron-Phonon scattering

We consider intravalley and intervalley scattering processes mediated by acoustic and optical phonons using Fermi's golden rule. The transition probability for intravalley scattering from the initial state $\left|m'\beta' k'_{\mathrm{y}}\right\rangle$ to the final state $|m\beta k_{\mathrm{y}}\rangle$ in presence of acoustic phonons of all possible wavevectors $\boldsymbol{q}$ and angular frequencies ω_q can be defined as:

$$W^{\mathrm{ac}}_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}} = \frac{2\pi}{\hbar} \sum_{k'_{\mathrm{y}}} \frac{1}{L_{\mathrm{y}}^2} \frac{\hbar \Delta_{\mathrm{ac}}^2 q^2}{2 d_{\mathrm{c}} \omega_q} \left(\langle N_q \rangle + \frac{1}{2} \pm \frac{1}{2} \right) \times \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \left| \left\langle m\beta k_{\mathrm{y}} \middle| \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} \middle| m'\beta' k'_{\mathrm{y}} \right\rangle \right|^2 \delta\Big(E_{\beta,k_{\mathrm{y}}} - E_{\beta',k'_{\mathrm{y}}} \mp \hbar\omega_q \Big) \ , \quad (7.7)$$

where Δ_{ac} is the acoustic-phonon deformation potential, d_{c} is the mass density of the crystal, $\langle N_q \rangle$ is the Bose-Einstein distribution and the plus and minus signs refer to phonon emission and absorption, respectively. The factor of $\frac{1}{L_{\mathrm{y}}^2}$ stems from a finite volume normalization of the electron plane wave along y, having taken a finite device width of L_{y} in that direction and then imposing periodic boundary conditions as $L_{\mathrm{y}} \to \infty$. The q dependence of the interaction term adds to the complexity of the calculations. To get rid of the q dependence, we assume the elastic, high-temperature, and equipartition approximations, and approximate the dispersion of the acoustic phonons in Si as $\omega_q = c_{\mathrm{s}} q$, where c_{s} is the sound velocity. Equation (7.7) then simplifies to:

$$W^{\mathrm{ac}}_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}} = \sum_{k'_{\mathrm{y}}} \frac{\pi}{L_{\mathrm{y}}^2} \frac{\Delta_{\mathrm{ac}}^2 \mathrm{k_B} T}{\hbar d_{\mathrm{c}} c_{\mathrm{s}}^2} \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \psi^{s,v*}_{m,\beta}(\boldsymbol{r}) \psi^{s,v}_{m,\beta}(\boldsymbol{r'}) \times \mathrm{e}^{\mathrm{i}\boldsymbol{q}(\boldsymbol{r}-\boldsymbol{r'})} \psi^{s,v}_{m',\beta'}(\boldsymbol{r}) \psi^{s,v*}_{m',\beta'}(\boldsymbol{r'}) \, \delta\left(E_{\beta,k_{\mathrm{y}}} - E_{\beta',k'_{\mathrm{y}}} \right)$$

Integrating over all $\boldsymbol{q}$, this expression can be rewritten as:

$$\begin{aligned}
W^{\mathrm{ac}}_{m\beta k_{\mathrm{y}},m'\beta'k'_{\mathrm{y}}} &= \sum_{k'_{\mathrm{y}}} \frac{\pi}{L_{\mathrm{y}}^2}\frac{\Delta_{\mathrm{ac}}^2 \mathrm{k_B} T}{\hbar d_{\mathrm{c}} c_{\mathrm{s}}^2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \psi^{r,v*}_{m,\beta}(\boldsymbol{r})\psi^{s,v}_{m,\beta}(\boldsymbol{r}')\mathrm{e}^{\mathrm{i}\boldsymbol{q}(\boldsymbol{r}-\boldsymbol{r}')}\psi^{s,v}_{m',\beta'}(\boldsymbol{r})\psi^{s,v*}_{m',\beta'}(\boldsymbol{r}') \\
&\qquad\qquad \times \delta\left(\boldsymbol{r}-\boldsymbol{r}'\right)\delta\left(E_{\beta,k_{\mathrm{y}}} - E_{\beta',k'_{\mathrm{y}}}\right) \\
&= \sum_{k'_{\mathrm{y}}} \frac{\pi}{L_{\mathrm{y}}^2}\frac{\Delta_{\mathrm{ac}}^2 \mathrm{k_B} T}{\hbar d_{\mathrm{c}} c_{\mathrm{s}}^2} \int \mathrm{d}^3 r \left|\psi^{s,v}_{m,\beta}(\boldsymbol{r})\right|^2 \left|\psi^{s,v}_{m',\beta'}(\boldsymbol{r})\right|^2 \delta\left(E_{\beta,k_{\mathrm{y}}} - E_{\beta',k'_{\mathrm{y}}}\right) \\
&= \sum_{k'_{\mathrm{y}}} \frac{\pi}{L_{\mathrm{y}}^2}\frac{\Delta_{\mathrm{ac}}^2 \mathrm{k_B} T}{\hbar d_{\mathrm{c}} c_{\mathrm{s}}^2} \int_0^{L_{\mathrm{y}}} \mathrm{d}y \iint \mathrm{d}x\,\mathrm{d}z \left|\phi^{s,v}_{m,\beta}(x,z)\right|^2 \left|\phi^{s,v}_{m',\beta'}(x,z)\right|^2 \delta\left(E_{\beta,k_{\mathrm{y}}} - E_{\beta',k'_{\mathrm{y}}}\right)
\end{aligned}$$

The integral over the (x,z) plane is required only over the device domain since our set of *open* system wavefunctions already contain information regarding the infinite leads, as discussed in Sec. 3.3.1. Also, assuming periodic boundary conditions along y, the wavevectors k_{y} can be taken as multiples as $2\pi/L_{\mathrm{y}}$ and $\Delta k_{\mathrm{y}} = 2\pi/L_{\mathrm{y}}$. The final expression for the transition probability in presence of acoustic phonons then becomes:

$$\begin{aligned}
W^{\mathrm{ac}}_{m\beta k_{\mathrm{y}},m'\beta'k'_{\mathrm{y}}} = \frac{\Delta_{\mathrm{ac}}^2 \mathrm{k_B} T}{2\hbar d_{\mathrm{c}} c_{\mathrm{s}}^2} &\iint \mathrm{d}x\,\mathrm{d}z \left|\phi^{s,v}_{m,\beta}(x,z)\right|^2 \left|\phi^{s,v}_{m',\beta'}(x,z)\right|^2 \\
&\times \sum_{k'_{\mathrm{y}}} \Delta k'_{\mathrm{y}} \delta\Big(E^v_\beta - E^v_{\beta'} + E_{k_{\mathrm{y}}} - E_{k'_{\mathrm{y}}}\Big)\ , \quad (7.8)
\end{aligned}$$

where the summation over k'_{y} amounts to the DoS at $E_{k_{\mathrm{y}}} = E^v_\beta - E^v_{\beta'} + E_{k_{\mathrm{y}}}$.

In similar fashion, the intravalley transition probability for processes assisted by optical phonons can be expressed as:

$$\begin{aligned}
W^{\mathrm{op}}_{m\beta k_{\mathrm{y}},m'\beta'k'_{\mathrm{y}}} = \frac{(D_{\mathrm{t}}K)^2}{2 d_{\mathrm{c}} \omega_{\mathrm{op}}} \left(\langle N_{\mathrm{op}}\rangle + \frac{1}{2} \pm \frac{1}{2}\right) &\iint \mathrm{d}x\,\mathrm{d}z \left|\phi^{s,v}_{m,\beta}(x,z)\right|^2 \left|\phi^{s,v}_{m',\beta'}(x,z)\right|^2 \\
&\times \sum_{k'_{\mathrm{y}}} \Delta k'_{\mathrm{y}} \delta\Big(E^v_\beta - E^v_{\beta'} + E_{k_{\mathrm{y}}} - E_{k'_{\mathrm{y}}} \pm \hbar\omega_{\mathrm{op}}\Big)\ , \quad (7.9)
\end{aligned}$$

where $D_{\mathrm{t}}K$ is the optical phonon deformation potential. Equation (7.9) accounts for both emission and absorption of an optical phonon of energy $\hbar\omega_{\mathrm{op}}$, where the angular frequency

ω_{op} is assumed to be constant. A similar expression describes also intervalley scattering mediated by optical phonons.

The approximations made here are good enough to provide a decent quantitative idea of the impact of electron-phonon scattering in nanoscale Si devices. A comprehensive treatment of the full phonon dispersion is beyond the scope of this work.

Surface-roughness scattering

In this chapter, we model SR scattering as a perturbation away from ballistic transport, contrary to the statistical ab-initio approach followed in Chapter 6. To this end we follow Ando's model [29], but we also account for the change of the wavefunction, an effect that can be significant when dealing with UTB devices and, so, strong quantum confinement.

We start by introducing the surface-roughness as a perturbation of the potential barrier at the oxide-semiconductor interface within the first Born approximation. Let the oxide-semiconductor interfaces lie at $z = 0$ (bottom oxide interface) and $z = t_{\mathrm{s}}$, where t_{s} is the semiconductor thickness. In this case, the barrier potentials due to the bottom and top gate oxides can be written as:

$$\begin{aligned} V_{\mathrm{ox}}^{\mathrm{T}}(x,y,z) =& V_{\mathrm{b}}\Theta(-z) \\ V_{\mathrm{ox}}^{\mathrm{B}}(x,y,z) =& V_{\mathrm{b}}\Theta(z-t_{\mathrm{s}}) \ , \end{aligned}$$

where V_{b} is the oxide barrier which can typically be approximated by the band gap of the oxide. Let $\Delta^{\mathrm{B}}(x,y)$ and $\Delta^{\mathrm{T}}(x,y)$ represent variations of the bottom and top oxide interfaces, respectively, from the perfect interface plane. The perturbing potential for SR scattering can therefore be written as:

$$\begin{aligned} V_{\mathrm{SR}} = V_{\mathrm{SR}}^{\mathrm{B}} + V_{\mathrm{SR}}^{\mathrm{T}} = (V_{\mathrm{ox}}^{\mathrm{B}}(x,y,z+\Delta^{\mathrm{B}}(x,y)) - V_{\mathrm{ox}}^{\mathrm{B}}(x,y,z)) \\ + (V_{\mathrm{ox}}^{\mathrm{T}}(x,y,z+\Delta^{\mathrm{T}}(x,y)) - V_{\mathrm{ox}}^{\mathrm{T}}(x,y,z)) \ . \quad (7.10) \end{aligned}$$

We assume that scattering with the top and bottom interface is uncorrelated and therefore can be treated separately. Assuming $L_\mathrm{y} \to \infty$, the matrix element for scattering with the bottom oxide interface can be expressed as:

$$\langle m\beta k_\mathrm{y}| V_\mathrm{SR}^\mathrm{B} |m'\beta' k'_\mathrm{y}\rangle = \frac{1}{L_\mathrm{y}}\Bigg[\int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z \phi_{m,\beta}^{s,v*}(x,z) \mathrm{e}^{-\mathrm{i}k_\mathrm{y}y}\, V_\mathrm{ox}^\mathrm{B}(x,y,z+\Delta^\mathrm{B}(x,z)) \\ \times \phi_{m',\beta'}^{s,v}(x,z)\mathrm{e}^{\mathrm{i}k'_\mathrm{y}y} - \int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z\, \phi_{m,\beta}^{s,v*}(x,z)\mathrm{e}^{-\mathrm{i}k_\mathrm{y}y}\, V_\mathrm{ox}^\mathrm{T}(x,y,z)\phi_{m',\beta'}^{s,v}(x,z)\mathrm{e}^{\mathrm{i}k'_\mathrm{y}y}\Bigg] . \quad (7.11)$$

Transforming coordinates of the first integral from z to $z' = z + \Delta^\mathrm{B}(x, y)$ and replacing z' with z, we have:

$$\langle m\beta k_\mathrm{y}| V_\mathrm{SR}^\mathrm{B} |m'\beta' k'_\mathrm{y}\rangle = \frac{1}{L_\mathrm{y}}\Bigg[\int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z \phi_{m,\beta}^{s,v*}(x,z-\Delta^\mathrm{B}(x,z)) \mathrm{e}^{-\mathrm{i}k_\mathrm{y}y}\, V_\mathrm{ox}^\mathrm{B}(x,y,z) \\ \times \phi_{m',\beta'}^{s,v}(x,z-\Delta^\mathrm{B}(x,z))\mathrm{e}^{\mathrm{i}k'_\mathrm{y}y} - \int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z\, \phi_{m,\beta}^{s,v*}(x,z)\mathrm{e}^{-\mathrm{i}k_\mathrm{y}y}\, V_\mathrm{ox}^\mathrm{B}(x,y,z)\phi_{m',\beta'}^{s,v}(x,z)\mathrm{e}^{\mathrm{i}k'_\mathrm{y}y}\Bigg] . \quad (7.12)$$

This apparently simple transformation of axes is a significant step. Indeed the scattering-matrix element is now defined in terms of a shift of the electron wavefunction, rather than of the potential. Ando's model neglects changes in the wavefunction due to SR by invoking "the electric quantum limit" for inversion layers in large semiconductor devices. However, we do not have the same liberty for the case of modern SOI and UTB devices since any change in the boundary conditions affects the wavefunction. Moreover, we use the complete (2-D) device solutions (wavefunctions) in an attempt to improve upon the already existent models [29, 72, 73, 74] which consider solutions of the Schrödinger equation only along the 1-D cross-section of the device and plane waves elsewhere. Thus our model should theoretically provide a more realistic picture, properly taking into account the quantum confinement and other non-local quantum effects associated with SR scattering [77].

Expanding Eq. (7.12) to first order in $\phi^{r,v}_{m,\beta}$, we obtain:

$$\langle m\beta k_{\mathrm{y}}|\, V^{\mathrm{B}}_{\mathrm{SR}}\, |m'\beta' k'_{\mathrm{y}}\rangle = -\frac{V_{\mathrm{b}}}{L_{\mathrm{y}}}\int \mathrm{d}x \int \mathrm{d}y \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \Delta^{\mathrm{B}}(x,z) \mathrm{e}^{\mathrm{i}(k'_{\mathrm{y}}-k_{\mathrm{y}})y} \left[\frac{\partial \phi^{r,v*}_{m,\beta}(x,z)}{\partial z} \phi^{r,v}_{m',\beta'}(x,z) \right.$$
$$\left. + \phi^{r,v*}_{m,\beta}(x,z) \frac{\partial \phi_{m',\beta'}(x,z)}{\partial z} \right] , \quad (7.13)$$

where t_{ox} is the oxide thickness. The surface-roughness patterns $\Delta^{\mathrm{B}}(x,z)$ are random in nature and obey an exponential or Gaussian autocovariance [75]. The Fourier transform of the roughness pattern $\Delta^{\mathrm{B}}(x,z)$ can be taken equal to $\sqrt{L_{\mathrm{x}}L_{\mathrm{y}}}\Delta_{q_{\mathrm{x}},q_{\mathrm{y}}}\mathrm{e}^{\mathrm{iR}(q_{\mathrm{x}},q_{\mathrm{y}})}$, $\langle|\Delta_{q_{\mathrm{x}},q_{\mathrm{y}}}|\rangle^2$ being the power spectrum of the roughness (discussed in detail in Sec. 7.2) and $\sqrt{L_{\mathrm{x}}L_{\mathrm{y}}}$ is the normalization constant. $\mathrm{e}^{\mathrm{iR}(\mathbf{q})}$ (R being a random number generator) is the Fourier transform of white noise that is introduced to generate the random pattern of the roughness. The roughness pattern $\Delta^{\mathrm{B}}(x,z)$ can be written as:

$$\Delta^{\mathrm{B}}(x,z) = \iint \mathrm{d}q_{\mathrm{x}}\, \mathrm{d}q_{\mathrm{y}} \sqrt{L_{\mathrm{x}}L_{\mathrm{y}}}\Delta_{q_{\mathrm{x}},q_{\mathrm{y}}}\mathrm{e}^{\mathrm{iR}(q_{\mathrm{x}},q_{\mathrm{y}})} \times \mathrm{e}^{\mathrm{i}q_{\mathrm{x}}x}\mathrm{e}^{\mathrm{i}q_{\mathrm{y}}y} \quad (7.14)$$

Inserting this expression into Eq. (7.13) and performing the integration over y, we obtain:

$$\langle m\beta k_{\mathrm{y}}|\, V^{\mathrm{B}}_{\mathrm{SR}}\, |m'\beta' k'_{\mathrm{y}}\rangle = -\frac{2\pi V_{\mathrm{b}}}{L_{\mathrm{y}}}\int \mathrm{d}x \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \left[\frac{\partial \phi^{s,v*}_{m,\beta}(x,z)}{\partial z} \phi^{s,v}_{m',\beta'}(x,z) \right.$$
$$\left. + \phi^{s,v*}_{m,\beta}(x,z) \frac{\partial \phi^{s,v}_{m',\beta'}(x,z)}{\partial z} \right] \iint \mathrm{d}q_{\mathrm{x}}\, \mathrm{d}q_{\mathrm{y}} \sqrt{L_{\mathrm{x}}L_{\mathrm{y}}}\Delta_{q_{\mathrm{x}},q_{\mathrm{y}}}\mathrm{e}^{\mathrm{iR}(q_{\mathrm{x}},q_{\mathrm{y}})}\mathrm{e}^{\mathrm{i}q_{\mathrm{x}}x}\delta(k'_{\mathrm{y}} - k_{\mathrm{y}} + q_{\mathrm{y}}) , \quad (7.15)$$

Let us define:

$$\Phi(x,z) = \left[\frac{\partial \phi^{s,v*}_{m,\beta}(x,z)}{\partial z} \phi^{s,v}_{m',\beta'}(x,z) + \phi^{s,v*}_{m,\beta}(x,z) \times \frac{\partial \phi^{s,v}_{m',\beta'}(x,z)}{\partial z} \right] .$$

The transition probability $W^{\mathrm{SR}}_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}}$, as well as the scattering rates associated with SR scattering, do not depend on the randomness of the roughness patterns. Therefore, instead of taking the absolute square of the matrix element Eq. (7.12) to determine the transition probability, we use the expectation value of the absolute square of the matrix element given by Eq. (7.15):

$$\left\langle \left| \langle m\beta k_{\mathrm{y}} | \, V_{\mathrm{SR}}^{\mathrm{B}} \, | m'\beta' k'_{\mathrm{y}} \rangle \right|^2 \right\rangle = \frac{4\pi^2 V_{\mathrm{b}}^2 L_{\mathrm{x}}}{L_{\mathrm{y}}} \int \mathrm{d}x \int \mathrm{d}x' \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z' \Phi(x,z) \Phi(x',z')^* \\ \times \int \mathrm{d}q_{\mathrm{x}} \int \mathrm{d}q_{\mathrm{x}'} \, \mathrm{e}^{\mathrm{i}q_{\mathrm{x}}x} \mathrm{e}^{-\mathrm{i}q'_{\mathrm{x}}x'} \left\langle \Delta_{q_{\mathrm{x}},k_{\mathrm{y}}-k'_{\mathrm{y}}} \Delta^*_{q_{\mathrm{x}'},k_{\mathrm{y}}-k'_{\mathrm{y}}} \right\rangle \left\langle \mathrm{e}^{\mathrm{iR}(q_{\mathrm{x}})} \mathrm{e}^{\mathrm{iR}(q'_{\mathrm{x}})} \right\rangle \ . \tag{7.16}$$

Now the power spectrum of white noise $\left\langle \mathrm{e}^{\mathrm{iR}(q_{\mathrm{x}})} \mathrm{e}^{\mathrm{iR}(q'_{\mathrm{x}})} \right\rangle$ is $\mathrm{C}\delta(q_{\mathrm{x}} - q'_{\mathrm{x}})$, where C is a constant. Using this fact in Eq. (7.16) and taking $\mathrm{C} = L_{\mathrm{x}}$, we obtain:

$$\left\langle \left| \langle m\beta k_{\mathrm{y}} | \, V_{\mathrm{SR}}^{\mathrm{B}} \, | m'\beta' k'_{\mathrm{y}} \rangle \right|^2 \right\rangle = \frac{4\pi^2 V_{\mathrm{b}}^2}{L_{\mathrm{y}}} \int \mathrm{d}q_{\mathrm{x}} \left\langle \left| \Delta_{q_{\mathrm{x}},k_{\mathrm{y}}-k'_{\mathrm{y}}} \right|^2 \right\rangle \left| \int \mathrm{d}x \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \, \Phi(x,z) \right|^2 \ . \tag{7.17}$$

Expressing Φ in terms of ϕ, the transition probability for SR scattering can be written as:

$$W^{\mathrm{SR}}_{m\beta k_{\mathrm{y}}, m'\beta' k'_{\mathrm{y}}} = \frac{2\pi}{\hbar} \sum_{k'_{\mathrm{y}}} \frac{4\pi^2 V_{\mathrm{b}}^2}{L_{\mathrm{y}}} \int \mathrm{d}q_{\mathrm{x}} \left\langle \left| \Delta_{q_{\mathrm{x}},k_{\mathrm{y}}-k'_{\mathrm{y}}} \right|^2 \right\rangle \\ \times \left| \int \mathrm{d}x \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \left[\frac{\partial \phi^{s,v*}_{m,\beta}(x,z)}{\partial z} \phi^{s,v}_{m',\beta'}(x,z) + \phi^{s,v*}_{m,\beta}(x,z) \right. \right. \\ \left. \left. \times \frac{\partial \phi^{s,v}_{m',\beta'}(x,z)}{\partial z} \right] \right|^2 \delta\left(E^v_{\beta} - E^v_{\beta'} + E_{k_{\mathrm{y}}} - E_{k'_{\mathrm{y}}} \right) \ . \tag{7.18}$$

Using the same discretization for the k_{y}'s described in Sec. 7.1.1, Eq. (7.18) becomes:

$$W^{\mathrm{SR}}_{m\beta k_{\mathrm{y}}, m'\beta' k'_{\mathrm{y}}} = \frac{4\pi^2 V_{\mathrm{b}}^2}{\hbar} \int \mathrm{d}q_{\mathrm{x}} \left\langle \left| \Delta^{\mathrm{B}}_{q_{\mathrm{x}},k_{\mathrm{y}}-k'_{\mathrm{y}}} \right|^2 \right\rangle \\ \times \left| \int \mathrm{d}x \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \left[\frac{\partial \phi^{s,v*}_{m,\beta}(x,z)}{\partial z} \phi^{s,v}_{m',\beta'}(x,z) + \phi^{s,v*}_{m,\beta}(x,z) \right. \right. \\ \left. \left. \times \frac{\partial \phi^{s,v}_{m',\beta'}(x,z)}{\partial z} \right] \mathrm{e}^{\mathrm{i}q_{\mathrm{x}}x} \right|^2 \sum_{k'_{\mathrm{y}}} \Delta k'_{\mathrm{y}} \delta\Big(E^v_{\beta} - E^v_{\beta'} + E_{k_{\mathrm{y}}} - E_{k'_{\mathrm{y}}} \Big) \tag{7.19}$$

Focussing on the limits for the integral over q_{x}, our sampling frequency (angular), given by $2\pi/\Delta x$, is limited by the spacing interval Δx set by our discrete mesh. From Nyquist theorem, the maximum frequency (q_{x}) of the roughness pattern that our sampling method can capture will be half of the sampling frequency, *i.e*, $\pi/\Delta x$. The integral over q_{x} in Eq. (7.19) can then be performed numerically within the range $[-\pi/\Delta x, \pi/\Delta x]$. The transition probability $W^{\mathrm{SR}}_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}}$ from the state $\left|m'\beta' k'_{\mathrm{y}}\right\rangle$ to the state $\left|m\beta k_{\mathrm{y}}\right\rangle$ therefore be written as:

$$W^{\mathrm{SR}}_{m\beta k_{\mathrm{y}},m'\beta' k'_{\mathrm{y}}} = \frac{4\pi^2 V_{\mathrm{b}}^2}{\hbar} \int_{-\pi/\Delta x}^{\pi/\Delta x} \mathrm{d}q_{\mathrm{x}} \left\langle \left| \Delta^{\mathrm{B}}_{q_{\mathrm{x}},k_{\mathrm{y}}-k'_{\mathrm{y}}} \right|^2 \right\rangle \left| \int \mathrm{d}x \int_{-t_{\mathrm{ox}}}^{0} \mathrm{d}z \left[\frac{\partial \phi^{s,v*}_{m,\beta}(x,z)}{\partial z} \phi^{s,v}_{m',\beta'}(x,z) \right.\right.$$
$$\left.\left. + \, \phi^{s,v*}_{m,\beta}(x,z) \frac{\partial \phi^{s,v}_{m',\beta'}(x,z)}{\partial z} \right] \mathrm{e}^{\mathrm{i}q_{\mathrm{x}}x} \right|^2 \sum_{k'_{\mathrm{y}}} \Delta k'_{\mathrm{y}} \delta\Big(E^{v}_{\beta} - E^{v}_{\beta'} + E_{k_{\mathrm{y}}} - E_{k'_{\mathrm{y}}} \Big) \tag{7.20}$$

A similar expression can be derived for scattering with the top interface.

Electron charge density in presence of dissipation

The different scattering mechanisms we have discussed are independent of each other. Therefore their transition rates can be summed together and incorporated into the PME to determine their net impact on electron transport. The density-matrix elements $n(x, y, z)$ obtained after solving Eq. (7.6) can be used to determine the electron-density distribution at steady-state.

$$n(x,z) = \frac{1}{\pi} \sum_{r} \sum_{v} \sum_{\beta} \sum_{m} \sum_{k_{\mathrm{y}}} \Delta k_{\mathrm{y}} \, \rho_{m\beta k_{\mathrm{y}}} \left| \phi^{s,v}_{m,\beta}(x,z) \right|^2 \tag{7.21}$$

Self-consistent scheme

The Schrödinger, PME and Poisson equations are solved self-consistently in the 2-D plane of the device to study dissipative quantum transport. The Newton's method, described in Sec. 2.5, is used to accelerate the convergence. We use Neumann boundary conditions on the electrostatic potential to maintain charge neutrality at the contact edges. Although physically inconsistent, this method is computationally less expensive and provides results similar to the method using Dirichlet boundaries, described in Sec. 3.3.1. Once we have obtained the self-consistent solutions of the *open* system and, by using the PME, determined their corresponding occupations, we can calculate different current transport parameters, namely, the transmission coefficients for electrons injected at different energies, the local density-of-states (LDOS), the current-density distribution and the total drain current [1], as described in Sec. 3.2.

7.2 Dissipative Transport in UTB FETs

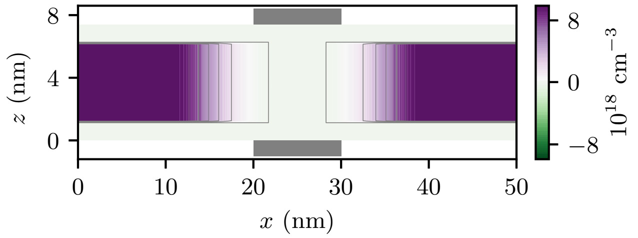

Figure 7.1: Net doping profile of the simulated UTB DG nMOS. The white regions at the top and bottom represent the 1.2 nm-thick gate oxide, while the grey patches are used to highlight the position of the gate contacts.

To demonstrate the capability of our dissipative quantum-transport model, we simulate the electrical behavior of a 5 nm-thick and 10 nm-long nMOS with lightly p-type doped ($\approx 10^{15}$ cm^{-3}) channel as shown in Fig. 7.1. The device dimensions are set in accordance with the end-of-2027 roadmap goals set by the IRDS [78]. A symmetric 1.2 nm-thick SiO_2 gate insulator is assumed for the top and bottom gates. This assumption is consistent with the reported fabrication procedures [79, 80] where a 1 nm-thick SiO_2 interface layer is grown before high-κ deposition. The exact nature of the roughness present at the Si-SiO_2 interface is still an open topic of debate. Traditionally it has been characterized using a Gaussian autocorrelation function. However, a critical statistical study by Goodnick et al. [75] shows that an exponential autocovariance resembles the nature of the interface more closely. Therefore, we have used this model in our calculation (Eq. (7.20)) for SR scattering. The power spectrum $\langle|\Delta_{q_\mathrm{x},q_\mathrm{y}}|\rangle^2$, in this case, is given by the expression:

$$\langle|\Delta_{q_\mathrm{x},q_\mathrm{y}}|^2\rangle = \frac{\pi\Delta_\mathrm{rms}^2\Lambda^2}{(1+(q_\mathrm{x}^2+q_\mathrm{y}^2)\Lambda^2/2)^{\frac{3}{2}}} \ , \tag{7.22}$$

where Δ_{rms} is the root-mean-square (rms) roughness, Λ is the correlation length and $\boldsymbol{q}$ is the scattered wave vector, in the (x, y) plane, from the Fourier transform of the roughness. For this work, we have taken the $\Delta = 0.3$ nm and $\Lambda = 1.5$ nm, following Ref. 75.

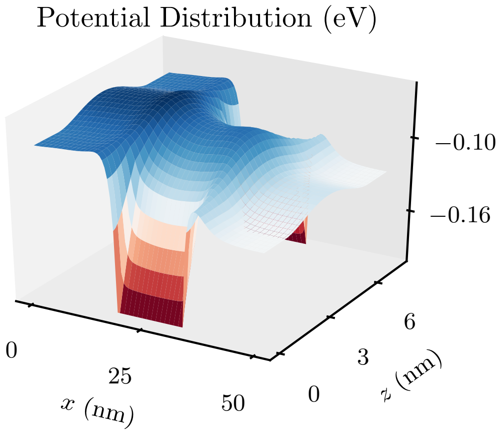

Figure 7.2: Simulated potential-energy distribution in the 5 nm UTB DG nMOS plotted for reference. The potential energy is measured with respect to the source fermi level.

7.2.1 Simulation of different scattering phenomena

Intravalley scattering with acoustic and optical phonons is modeled by employing the following values for the isotropic deformation potentials, $\Delta_{\text{ac}} = 9$ eV, and $D_{\text{t}}K = 2.2 \times 10^8$ eV/cm, respectively, following Ref. 6. $\hbar\omega_{\text{op}} = 60$ meV is taken for the optical phonons. For intervalley scattering, we only consider g-type processes with optical phonons, which should be adequate enough to ascertain the impact of electron-phonon scattering without going into elaborate details. We choose to use the parameters given by Canali et al. [81] favoring a weaker g-type scattering, since they have been shown [82] to correspond well with experimental mobility values for bulk and strained Si. Finally, the SR scattering rates given by Eq. (7.20) are inversely dependent on q and can be neglected for short-wavelengths (large q) corresponding to state transitions between Si valleys. Therefore, we only consider intravalley scattering for SR.

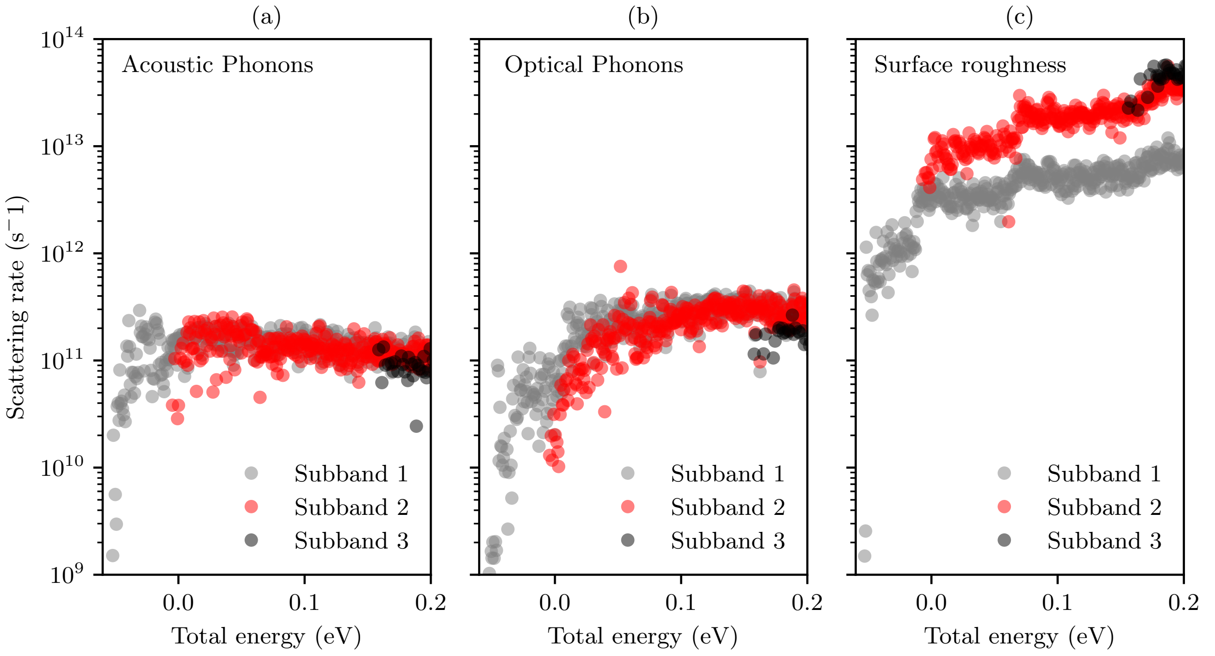

Figure 7.3: Scattering rates as a function of total electron-energy simulated for the different scattering mechanisms. V_{GS} =0.2 V, V_{DS} = 100 mV.

The transition rates for the different scattering phenomenon within the device domain are given by the out-scattering terms $W_{m'\beta' k'_{\mathrm{y}}, m\beta k_{\mathrm{y}}}$ of the PME (Eq. (7.3)). The total scattering rates for scattering with acoustic phonons, optical phonons and SR, $\sum_{m'\beta' k'_{\mathrm{y}}} W_{m'\beta' k'_{\mathrm{y}}, m\beta k_{\mathrm{y}}}$ are shown in Fig. 7.3 as a function of the total electron-energy. The figure(s) shows rates averaged over all states $|m\beta k_{\mathrm{y}}\rangle$ with the same subband index m (represented by the different colored lines) and injection energy E^v_β but having different k_{y}. These scattering rates are calculated for the individual delocalized wavefunctions. The figure clearly shows that the scattering rates for SR are at least an order of magnitude higher than the rates for scattering with both acoustic and optical phonons, thereby indicating a stronger impact of SR on device performance, as expected. A more quantitative discussion of the impact of this process on device performance will be addressed later when dealing with the electrical behavior of the device in Sec. 7.2.2. Note here that the scattering rates, especially in the case of phonons, are much smaller than the reported values for bulk Si. The situation has

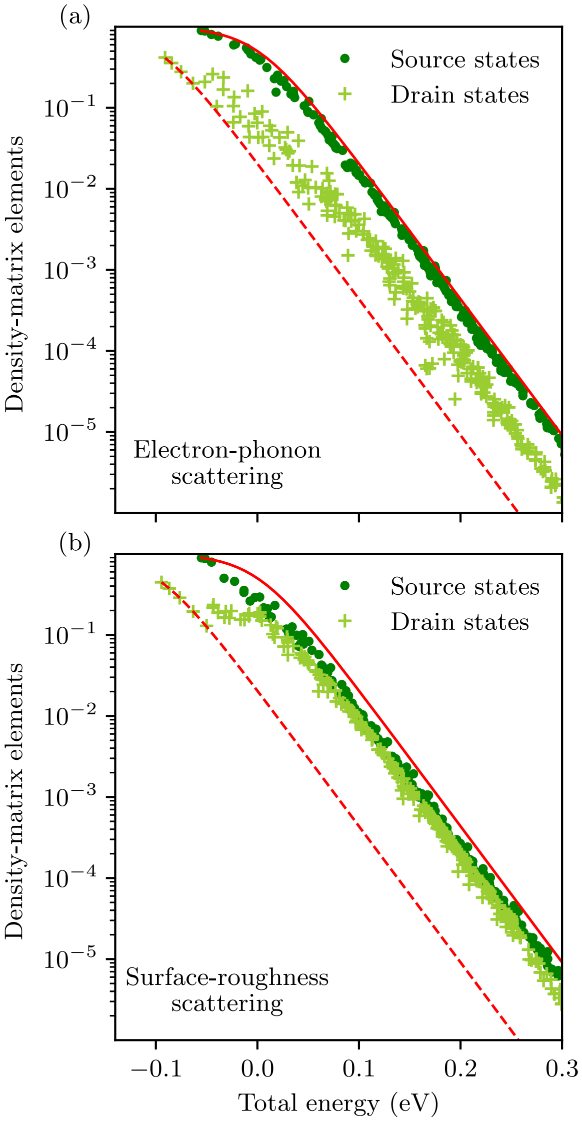

Figure 7.4: Diagonal elements of the density-matrix, as a function of total electron-energy, representing the final occupation of states in presence of (a) electron-phonon scattering, and (b) SR scattering. The solid and dashed red lines represent the ballistic occupation of states injected from the source and the drain, respectively.

been explicitly reported for the case of Si inversion layers in micrometer-scale MOSFETs [83]. The discrepancy arises because we take into account the overlap integral between 2-D device solutions pertaining to the initial and final states when computing the corresponding transition rates, thus limiting scattering to ”mutually favorable” states allowed by the device environment. From a more physical point of view, the scattering rates calculated in this way account for the characteristics of the device environment—namely electric field distribution, carrier concentration, device structure, thereby presenting a more realistic and quantifiable picture.

Figure 7.4 shows the diagonal elements of the density-matrix in presence of acoustic phonons, optical phonons and SR, separately, which gives us an idea about the final occupation of electronic states as a function of their energy $E_{m\beta k_y}$. The solid lines represent the ballistic distribution, while the color dots represent the distribution in the presence of different scattering mechanisms for injections from both the source and drain. The considerable shift in the final occupation of a state injected from source/drain from the ballistic occupation controlled by the Fermi level of the corresponding electrode (drain/ source) shows that back-scattering is the dominant mechanism of dissipation in all the cases presented.

7.2.2 Transport Characteristics

In this section we investigate the impact of the different scattering mechanisms on the transport behavior of the FETs we have considered. The density-of-states representing the continuous energy spectrum of the infinitely long leads is assigned a spatial dependence by factoring in the electron probability-density distribution in the device. This gives us the local density-of-states (LDOS), which is an important parameter for analyzing the device behavior. At the same time, we also incorporate the occupation factors determined by the PME in an attempt to understand how the different scattering mechanisms affect the LDOS. Thus the quantity plotted in Fig. 7.5 resembles the spectral function commonly used in NEGF calculations. The expected broadening/smearing in energy of the electronic states in presence of dissipation can be clearly seen from these plots. Back-scattering of electrons results in the build up of space charge in the channel of the device. As a result, the potential-energy in the channel is higher in the presence of scattering than in the ballistic case. This can be seen in Fig. 7.5 where the dashed lines represent the potential-energy distribution along the middle of the device.

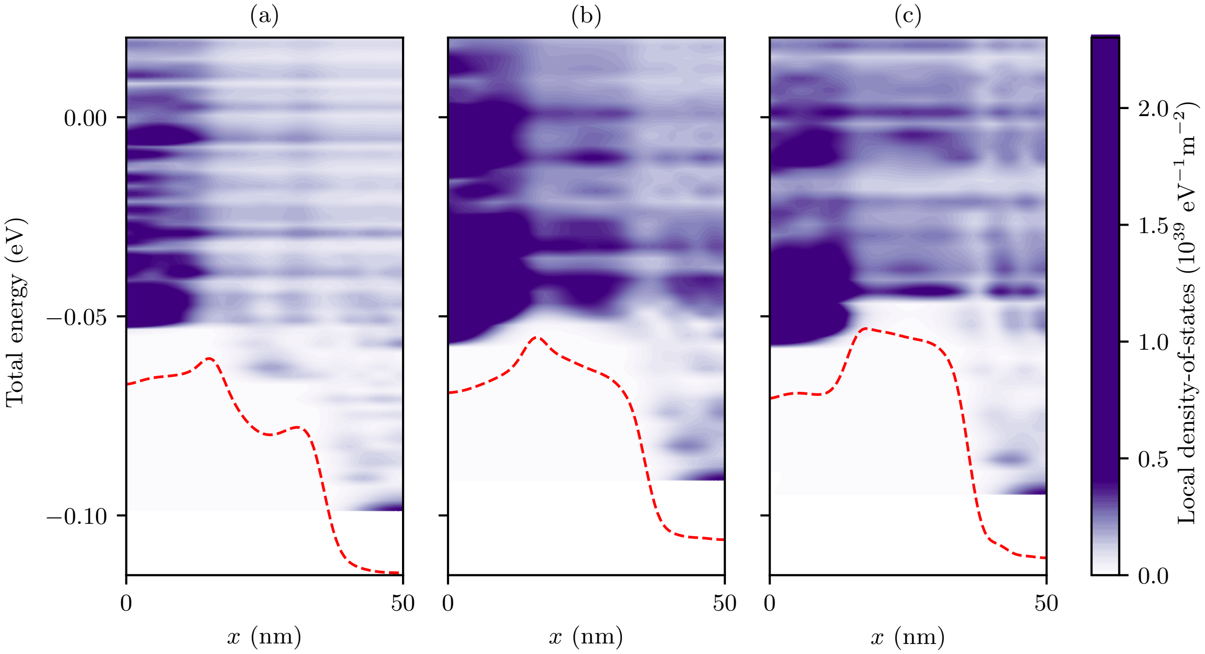

Figure 7.5: LDOS distribution in (a) ballistic case, (b) in presence of phonon (optical and acoustic) scattering, and (c) in presence of scattering with surface roughness. The LDOS is averaged over a cross-sectional thickness of roughly 1 nm along the middle of the device. The red dashed line represents the potential distribution cut across the device mid-section. The energies are measured with respect to the source Fermi level. V_{GS} =0.2 V, V_{DS} = 100 mV.

Moreover, semi-classical physical observables, the main information provided by Monte Carlo simulations, like the kinetic energy and drift velocity of electrons, can also be extracted from our simulations. To study and compare the effect of the different scattering mechanisms, a spatial dependence is assigned to the expectation value of these observables (quantum operators), following Ref. 21. The general definition of the expectation value of an operator O can be written as:

$$\langle O \rangle = \frac{1}{Z} \sum_{m\beta k_{\mathrm{y}}} \rho_{m\beta k_{\mathrm{y}}} \langle m\beta k_{\mathrm{y}} | O | m\beta k_{\mathrm{y}} \rangle \ , \tag{7.23}$$

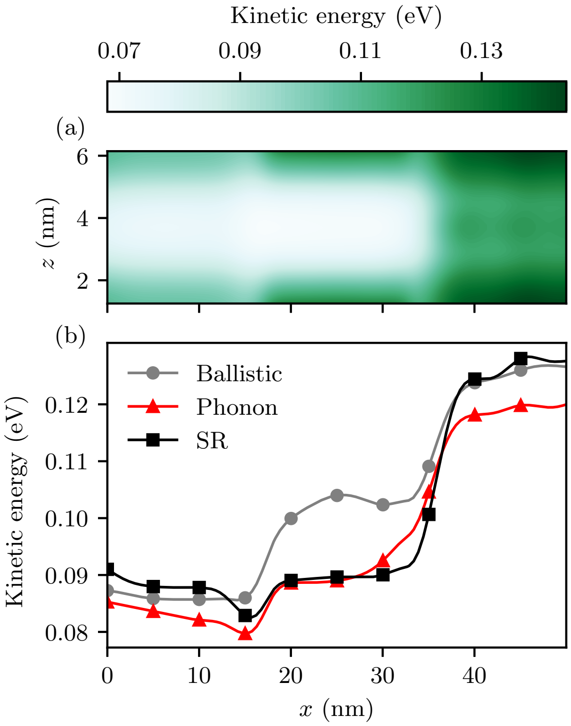

Figure 7.6: (a) Spatial distribution of the electron kinetic energy in presence of SR scattering. Kinetic-energy distributions for the ballistic and electron-phonon are largely similar to the above, except in the channel. The latter is highlighted in (c) which shows the kinetic energy averaged over the cross-sectional thickness along the length of the device for the different scattering mechanisms.

where Z is the partition function expressed as $\sum_{m\beta k_\mathrm{y}} \rho_{m\beta k_\mathrm{y}} \langle m\beta k_\mathrm{y}|m\beta k_\mathrm{y}\rangle$. This expression can be extended to provide an arbitrary yet useful definition of a spatially dependent expectation value of an observable O, $O(x,y)$ [21, 22]:

$$O(x,y) = \frac{1}{Z}\sum_{m\beta k_\mathrm{y}} \rho_{m\beta k_\mathrm{y}} \langle m\beta k_\mathrm{y}|x,z\rangle \langle x,z|O|m\beta k_\mathrm{y}\rangle \ . \tag{7.24}$$

For the kinetic energy per electron E_KE, we then have:

$$\begin{aligned} E_\mathrm{KE}(x,z) &= \frac{1}{Z}\sum_{m\beta k_\mathrm{y}} \rho_{m\beta k_\mathrm{y}} \phi^{s,v*}_{m,\beta}(x,z) \left[\frac{1}{2m_\mathrm{e}}\left(-\mathrm{i}\hbar\nabla\right)^2\right] \phi^{s,v}_{m,\beta}(x,z) \\ &= \frac{1}{Z}\sum_{m\beta k_\mathrm{y}} \rho_{m\beta k_\mathrm{y}} [E^v_\beta + E_{k_\mathrm{y}} - V(x,z)] \left|\phi^{s,v}_{m,\beta}(x,z)\right|^2 \ . \end{aligned} \tag{7.25}$$

Here we have utilized the fact that the wavefunctions $\phi_{m,\beta}^{s,v}$ are eigensolutions of the 2-D Schrödinger equation and m_{e} represents the effective mass tensor. Figure 7.6 shows the kinetic energy distribution, calculated as described above, in the presence of SR scattering and its average over vertical cross sections along the device length compared with the result obtained when considering ballistic transport or scattering only with phonons. It can be seen that there is marginal distinction in kinetic energy for the different cases, except in the inverted channel region. Here, the back-scattering of electrons raises the potential energy, as shown in Fig. 7.5, and simultaneously lowers the kinetic energy in the channel.

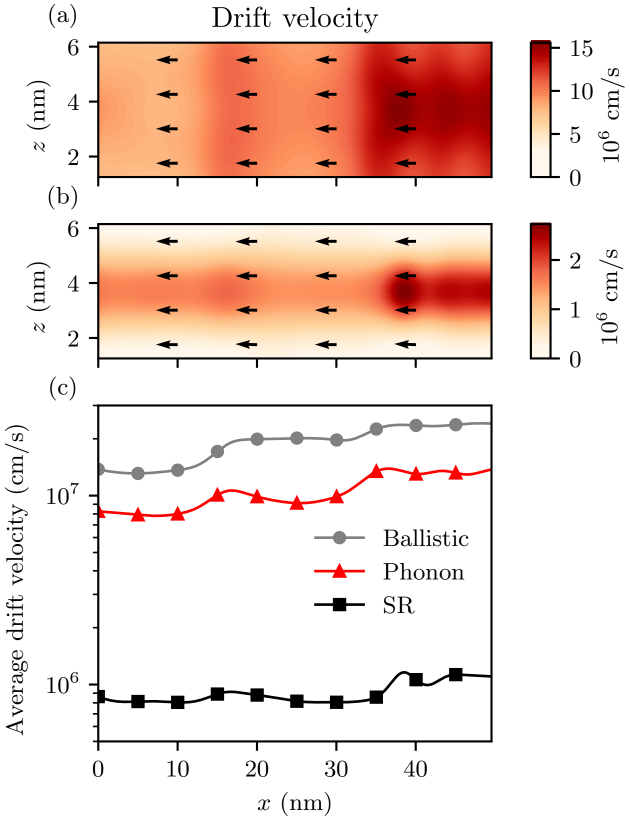

Figure 7.7: Spatial distribution of the electron drift velocity in the presence of (a) electron-phonon scattering and (b) SR scattering. (c) Comparative plot of the drift velocity, averaged over the cross-sectional thickness along the length of the device, in presence of different scattering mechanisms.

Semi-classically, the drift velocity v_{d} can be viewed as $v_{\mathrm{d}} = j/(\mathrm{q_e}n)$, where j is the current density and n is the carrier concentration. Extending this to the quantum regime, j transforms into the probability current given by $j = \mathrm{q_e}\hbar/m_{\mathrm{e}}\,\mathrm{Re}\left(\phi^{s,v*}_{m,\beta}\nabla\phi^{s,v}_{m,\beta}\right)$. Using the semi-classical expression and following the procedure of Eq. (7.24), we can now define the spatially-dependent expectation value of the drift velocity as:

$$v_{\mathrm{d}} = \frac{1}{Z}\frac{\hbar}{m_{\mathrm{e}}}\sum_{m\beta k_{\mathrm{y}}} \rho_{m\beta k_{\mathrm{y}}}\mathrm{Re}\left(\phi^{s,v*}_{m,\beta}(x,z)\nabla\phi^{s,v}_{m,\beta}(x,z)\right) \ , \tag{7.26}$$

where the partition function Z acts as the quantum equivalent of the carrier concentration n. Calculations have been limited to the (x, z) plane of the device since this is the region of interest. Figures 7.7(a) and (b) compare the drift velocity distributions when accounting for scattering with phonons and SR, respectively. We can observe that there is a stark decrease in the drift velocity in presence of surface roughness as compared to the phonons case and also the ballistic situation displayed in the comparison plot of Fig. 7.7(c). At the same time, Fig. 7.7(b) shows that transport is affected mostly near the semiconductor-dielectric interface, highlighting the interfacial nature of SR scattering [77].

Finally, we study the transfer characteristics of the device in the presence of the different scattering mechanisms. We can obtain a clear quantitative assessment of the negative impact of surface roughness in UTB devices from the 'ON' state behavior highlighted in Fig. 7.8(a). The subthreshold characteristics, as can be seen from Fig 7.8, remain relatively unaffected. Interestingly, there is also notable decrease in drain current in the presence of scattering with phonons even in devices of such small dimensions. However, scattering with surface roughness dominates among the dissipative processes and is the principal factor limiting device performance. Comparing our results with statistical ab-initio simulations of SR in similar device structures [84, 77], we see a much more pronounced impact of SR primarily because we explicitly take into account the 2-D nature of the interface, contrary to a purely 1-D [77] or a projected 1-D [84] roughness pattern assumed in the cited articles. Moreover,

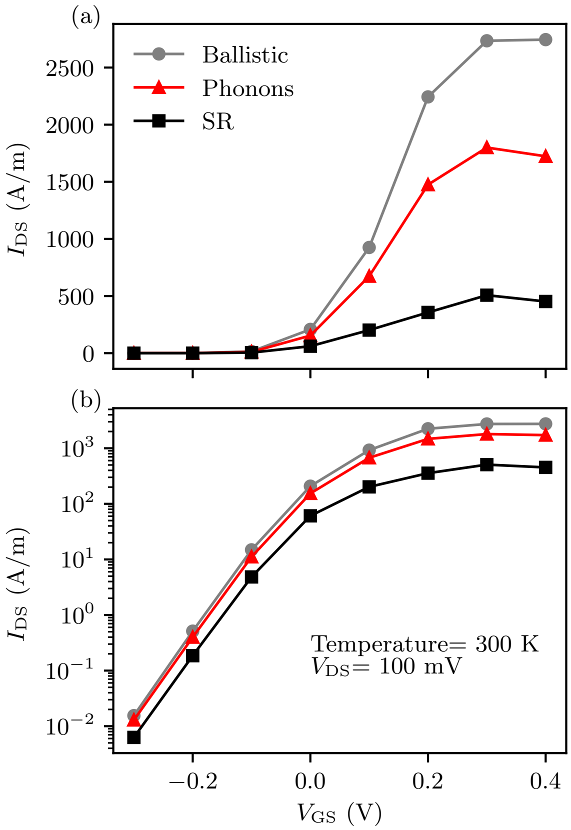

Figure 7.8: Transfer characteristics of the simulated 5 nm UTB DG nMOS in (a) linear and (b) semi-log scale focussing on the above-threshold and subthreshold regions, respectively.

this allows us to include transitions between states having different momenta or k_{y} in the out-of-plane direction, while the ab-initio cases, mentioned above, implicitly assume conservation of this out-of-plane momenta, thus limiting the scattering of electrons.

CHAPTER 8

CONCLUSION

In summary we have developed and demonstrated a novel and efficient approach to simulate quantum transport, both ballistic and dissipative, in realistic semiconductor devices. The method allows us to analyze the internal aspects of the device and the factors affecting its performance. In addition, the significance of our theoretical model is highlighted by the varied applications described in this work.

In Chapter 4, our ballistic transport simulations show that lateral-quantum-well Si nMOSFETs exhibit negative differential transconductance at cryogenic temperatures and for gate lengths shorter than 20 nm. Extrapolating from our results, devices with a gate length of 10 nm and lower should exhibit a sharp NDT signature even at room temperature, provided there is sufficient mitigation of the thermionic and punch-through currents. The former plays a crucial role in suppressing the NDT, the latter is favored over the tunneling current for high pSDE barriers. We have argued that alternative device designs (UTB SOI devices and/or epitaxial pSDE barriers) with sub-10 nm gate length are required to observe NDT at room temperature. Discussing processing and fabrication issues related to these scaled devices is a problem that transcends the scope of this work. However, the theoretical prediction of NDT in 10 nm devices leaves room for a moderate optimism. We have also shown the strong dependence of the resonant tunneling current on details of the doping profile of the pSDE pockets. This suggests that our inability to explain the NDT observed in 40 nm gate-length QW nMOS devices is likely due to the uncertainty of the actual doping profiles. This observation also bolsters our optimism.

In Chapter 5, we simulate ballistic electron transport in UTB DG nMOS. The simulated UTB DG nMOS exhibits Fano interference which results in the formation of vortices in the electron current at cryogenic temperatures. Thermal smearing prevents the phenomenon from manifesting itself in the I-V characteristics of the device at higher temperatures. Moreover, when asymmetry is introduced by applying an unequal gate bias in this case, the antiresonance features broaden and vortices in the current appear even at room temperature. We this conjecture that this quantum phenomenon can be observed at the macroscopic scale under the right experimental conditions.

In Chapter 6, we model the impact of surface roughness (SR) on electronic transport in UTB FETs. Here we adopt a statistical ab-initio approach to simulating the SR scattering. Our simulations show that SR scattering considerably affects the above-threshold behavior of UTB FETs, while subthreshold device parameters remain largely unaffected. We also show that quantum confinement plays a notable role in augmenting the impact of SR scattering. Finally, we demonstrate the (in)dependence of SR scattering on channel thickness.

Finally, in Chapter 7, we have demonstrated the capability of our simulation tool to model dissipative quantum transport in realistic semiconductor devices. As an application of our method, the impact of electron-phonon and SR scattering has been studied for UTB FETs. An extension of the Ando's model has been described to properly include quantum confinement and non-local effects associated with surface roughness. Our results show that, even in nanoscale devices, electron transport is predominantly dissipative (non-ballistic). Particularly, scattering of electrons with surface-roughness is the prime source of dissipation, drastically reducing the drain current (ballistic) by almost an order of magnitude.

In the future, our PME based model can be readily extended to include other important scattering mechanisms, like electron-electron and impurity scattering [85]. Moreover, the approximations for the phonon dispersion assumed here can be relaxed by treating the full phonon band dispersion. Finally, extending our effective-mass based model to a full-band model would enable us to accurately model electron transport in nanostructures like nanoribbons and nanowires as well as semiconducting materials beyond Si.

REFERENCES

[1] P. B. Vyas, C. Naquin, H. Edwards, M. Lee, W. G. Vandenberghe, and M. V. Fischetti. Theoretical simulation of negative differential transconductance in lateral quantum well nmos devices. *Journal of Applied Physics*, 121(4):044501, 2017.

[2] F. Gamiz, J. A. Lopez-Villanueva, J. B. Roldan, J. E. Carceller, and P. Cartujo. Monte carlo simulation of electron transport properties in extremely thin soi mosfet's. *IEEE Transactions on Electron Devices*, 45(5):1122–1126, May 1998.

[3] J. Guo and M. S. Lundstrom. A computational study of thin-body, double-gate, schottky barrier mosfets. *IEEE Transactions on Electron Devices*, 49(11):1897–1902, Nov 2002.

[4] K. K. Bhuwalka, S. Sedlmaier, A. K. Ludsteck, C. Tolksdorf, J. Schulze, and I. Eisele. Vertical tunnel field-effect transistor. *IEEE Transactions on Electron Devices*, 51(2):279–282, Feb 2004.

[5] M. C. Lemme, T. J. Echtermeyer, M. Baus, and H. Kurz. A graphene field-effect device. *IEEE Electron Device Letters*, 28(4):282–284, April 2007.

[6] C. Jacoboni and L. Reggiani. The monte carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Rev. Mod. Phys.*, 55:645–705, Jul 1983.

[7] M. V. Fischetti and S. E. Laux. Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects. *Phys. Rev. B*, 38:9721–9745, Nov 1988.

[8] J. Saint-Martin, A. Bournel, F. Monsef, C. Chassat, and P. Dollfus. Multi sub-band monte carlo simulation of an ultra-thin double gate MOSFET with 2d electron gas. *Semiconductor Science and Technology*, 21(4):L29–L31, feb 2006.

[9] B. Meinerzhagen and W. L. Engl. The influence of the thermal equilibrium approximation on the accuracy of classical two-dimensional numerical modeling of silicon submicrometer mos transistors. *IEEE Transactions on Electron Devices*, 35(5):689–697, May 1988.

[10] R. Thoma, A. Emunds, B. Meinerzhagen, H. J. Peifer, and W. L. Engl. Hydrodynamic equations for semiconductors with nonparabolic band structure. *IEEE Transactions on Electron Devices*, 38(6):1343–1353, Jun 1991.

[11] International technology roadmap for devices and systems 2017 edition. `https://irds.ieee.org/roadmap-2017`.

[12] T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C. D. Parker, and D. D. Peck. Resonant tunneling through quantum wells at frequencies up to 2.5 thz. *Applied Physics Letters*, 43(6):588–590, 1983.

[13] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi. Terahertz semiconductor-heterostructure laser. *Nature*, 417:156 EP –, 05 2002.

[14] O. Astafiev, S. Komiyama, T. Kutsuwa, V. Antonov, Y. Kawaguchi, and K. Hirakawa. Single-photon detector in the microwave range. *Applied Physics Letters*, 80(22):4250–4252, 2002.

[15] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, Jun 1932.

[16] W. R. Frensley. Boundary conditions for open quantum systems driven far from equilibrium. *Rev. Mod. Phys.*, 62:745–791, Jul 1990.

[17] M. L. Van de Put, B. Sorée, and W. Magnus. Efficient solution of the wigner-liouville equation using a spectral decomposition of the force field. *Journal of Computational Physics*, 350:314 – 325, 2017.

[18] L. P. Kadanoff and G. Baym. *Quantum Statistical Mechanics.* W. A. Benjamin Inc., New York, 1962.

[19] L. V. Keldysh. Diagram technique for nonequilibrium processes. *Zh. Eksp. Teor. Fiz.*, 47:1515, 1964.

[20] S. O. Koswatta, S. Hasan, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov. Nonequilibrium green's function treatment of phonon scattering in carbon-nanotube transistors. *IEEE Transactions on Electron Devices*, 54(9):2339–2351, Sep. 2007.

[21] M. V. Fischetti. Theory of electron transport in small semiconductor devices using the pauli master equation. *Journal of Applied Physics*, 83(1):270–291, 1998.

[22] M. V. Fischetti. Master-equation approach to the study of electronic transport in small semiconductor devices. *Phys. Rev. B*, 59:4901–4917, Feb 1999.

[23] M. L. Van de Put, M. V. Fischetti, and W. G. Vandenberghe. Scalable Atomistic Simulations of Quantum Electron Transport using Empirical Pseudopotentials. *arXiv e-prints*, page arXiv:1903.00548, Mar 2019.

[24] J. Fang, W. G. Vandenberghe, B. Fu, and M. V. Fischetti. Pseudopotential-based electron quantum transport: Theoretical formulation and application to nanometer-scale silicon nanowire transistors. *Journal of Applied Physics*, 119(3):035701, 2016.

[25] C. S. Lent and D. J. Kirkner. The quantum transmitting boundary method. *Journal of Applied Physics*, 67(10):6353–6359, 1990.

[26] S. E. Laux, A. Kumar, and M. V. Fischetti. Analysis of quantum ballistic electron transport in ultrasmall silicon devices including space-charge and geometric effects. *Journal of Applied Physics*, 95(10):5545–5582, 2004.

[27] F. Stern. Self-consistent results for n-type si inversion layers. *Phys. Rev. B*, 5:4891–4899, Jun 1972.

[28] J. M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Phys. Rev.*, 97:869–883, Feb 1955.

[29] T. Ando, A. B. Fowler, and F. Stern. Electronic properties of two-dimensional systems. *Rev. Mod. Phys.*, 54:437–672, Apr 1982.

[30] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, Inc., New York, 1972.

[31] G. Dahlquest and A. Bjorck. *Numerical Methods.* Prentice-Hall, Englewood Cliffs, N. J., 1974.

[32] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom. nanomos 2.5: A two-dimensional simulator for quantum transport in double-gate mosfets. *IEEE Transactions on Electron Devices*, 50(9):1914–1925, Sept 2003.

[33] S. E. Laux, A. Kumar, and M. V. Fischetti. Ballistic fet modeling using qdame: quantum device analysis by modal evaluation. *IEEE Transactions on Nanotechnology*, 1(4):255–259, Dec 2002.

[34] W. Pötz. Self-consistent model of transport in quantum well tunneling structures. *Journal of Applied Physics*, 66(6):2458–2466, 1989.

[35] C. Naquin, M. Lee, H. Edwards, G. Mathur, T. Chatterjee, and K. Maggio. Gate length and temperature dependence of negative differential transconductance in silicon quantum well metal-oxide-semiconductor field-effect transistors. *Journal of Applied Physics*, 118(12):124505, 2015.

[36] L. Esaki. New phenomenon in narrow germanium $p-n$ junctions. *Phys. Rev.*, 109:603–604, Jan 1958.

[37] L. Esaki. Discovery of the tunnel diode. *IEEE Transactions on Electron Devices*, 23(7):644–647, Jul 1976.

[38] L. L. Chang, L. Esaki, and R. Tsu. Resonant tunneling in semiconductor double barriers. *Applied Physics Letters*, 24(12):593–595, 1974.

[39] J. R. Söderström, D. H. Chow, and T. C. McGill. New negative differential resistance device based on resonant interband tunneling. *Applied Physics Letters*, 55(11):1094–1096, 1989.

[40] L. F. Luo, R. Beresford, and W. I. Wang. Interband tunneling in polytype gasb/alsb/inas heterostructures. *Applied Physics Letters*, 55(19):2023–2025, 1989.

[41] C. Naquin, M. Lee, H. Edwards, G. Mathur, T. Chatterjee, and K. Maggio. Negative differential transconductance in silicon quantum well metal-oxide-semiconductor field effect/bipolar hybrid transistors. *Applied Physics Letters*, 105(21):213507, 2014.

[42] F. Beltram, F. Capasso, J. F. Walker, and R. J. Malik. Memory phenomena in heterojunction structures: Evidence for suppressed thermionic emission. *Applied Physics Letters*, 53(5):376–378, 1988.

[43] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc. Design of ion-implanted mosfet's with very small physical dimensions. *IEEE Journal of Solid-State Circuits*, 9(5):256–268, Oct 1974.

[44] G. Baccarani, M. R. Wordeman, and R. H. Dennard. Generalized scaling theory and its application to a 1/4 micrometer mosfet design. *IEEE Transactions on Electron Devices*, 31(4):452–462, Apr 1984.

[45] D. L. Harame, J. H. Comfort, J. D. Cressler, E. F. Crabbe, J. Y.-C. Sun, B. S. Meyerson, and T. Tice. Si/sige epitaxial-base transistors. ii. process integration and analog applications. *IEEE Transactions on Electron Devices*, 42(3):469–482, 1995.

[46] S. E. Thompson, M. Armstrong, C. Auth, M. Alavi, M. Buehler, R. Chau, S. Cea, T. Ghani, G. Glass, T. Hoffman, C. H. Jan, C. Kenyon, J. Klaus, K. Kuhn, Zhiyong Ma, B. Mcintyre, K. Mistry, A. Murthy, B. Obradovic, R. Nagisetty, Phi Nguyen, S. Sivakumar, R. Shaheed, L. Shifren, B. Tufts, S. Tyagi, M. Bohr, and Y. El-Mansy. A 90-nm logic technology featuring strained-silicon. *IEEE Transactions on Electron Devices*, 51(11):1790–1797, Nov 2004.

[47] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic. Single and multiband modeling of quantum electron transport through layered semiconductor devices. *Journal of Applied Physics*, 81(12):7845–7869, 1997.

[48] R. H. Yan, A. Ourmazd, and K. F. Lee. Scaling the si mosfet: from bulk to soi to bulk. *IEEE Transactions on Electron Devices*, 39(7):1704–1710, July 1992.

[49] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini, and T. Elewa. Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance. *IEEE Electron Device Letters*, 8(9):410–412, Sep. 1987.

[50] D. J. Frank, S. E. Laux, and M. V. Fischetti. Monte carlo simulation of a 30 nm dual-gate mosfet: how short can si go? In *1992 International Technical Digest on Electron Devices Meeting*, pages 553–556, Dec 1992.

[51] R. Granzner, V. M. Polyakov, F. Schwierz, M. Kittler, R. J. Luyken, and W. R˙ Simulation of nanoscale mosfets using modified drift-diffusion and hydrodynamic models and comparison with monte carlo results. *Microelectronic Engineering*, 83(2):241 – 246, 2006.

[52] G. Baccarani and S. Reggiani. A compact double-gate mosfet model comprising quantum-mechanical and non-static effects. In *1999 International Conference on Simulation of Semiconductor Processes and Devices. SISPAD'99 (IEEE Cat. No.99TH8387)*, pages 11–14, Sep. 1999.

[53] D. Munteanu, J. Autran, X. Loussier, S. Harrison, R. Cerutti, and T. Skotnicki. Quantum short-channel compact modelling of drain-current in double-gate mosfet. *Solid-State Electronics*, 50(4):680 – 686, 2006. Special Issue: Papers Selected from the 35th European Solid-State Device Research Conference - ESSDERC'05.

[54] U. Fano. Effects of configuration interaction on intensities and phase shifts. *Phys. Rev.*, 124:1866–1878, Dec 1961.

[55] R. C. Bowen, W. R. Frensley, G. Klimeck, and R. K. Lake. Transmission resonances and zeros in multiband models. *Phys. Rev. B*, 52:2754–2765, Jul 1995.

[56] M. D. Sturge, H. J. Guggenheim, and M. H. L. Pryce. Antiresonance in the optical spectra of transition-metal ions in crystals. *Phys. Rev. B*, 2:2459–2471, Oct 1970.

[57] W. Porod, Z. Shao, and C. S. Lent. Resonance-antiresonance line shape for transmission in quantum waveguides with resonantly coupled cavities. *Phys. Rev. B*, 48:8495–8498, Sep 1993.

[58] K. Kobayashi, H. Aikawa, A. Sano, S. Katsumoto, and Y. Iye. Fano resonance in a quantum wire with a side-coupled quantum dot. *Phys. Rev. B*, 70:035319, Jul 2004.

[59] M. L. L. de Guevara, F. Claro, and P. A. Orellana. Ghost fano resonance in a double quantum dot molecule attached to leads. *Phys. Rev. B*, 67:195335, May 2003.

[60] M. J. Gilbert and D. K. Ferry. Vorticity and quantum interference in ultra-small soi mosfets. *IEEE Transactions on Nanotechnology*, 4(3):355–359, May 2005.

[61] S. E. Laux, A. Kumar, and M. V. Fischetti. Does circulation in individual current states survive in the total current density? *Journal of Computational Electronics*, 2(2):105–108, Dec 2003.

[62] A. Asenov, S. Kaya, and J. H. Davies. Intrinsic threshold voltage fluctuations in decanano mosfets due to local oxide thickness variations. *IEEE Transactions on Electron Devices*, 49:112 – 119, Jan 2002.

[63] C. Riddet, A. R. Brown, C. L. Alexander, J. R. Watling, S. Roy, and A. Asenov. 3-d monte carlo simulation of the impact of quantum confinement scattering on the magnitude of current fluctuations in double gate mosfets. *IEEE Transactions on Nanotechnology*, 6(1):48–55, Jan 2007.

[64] J. Wang, E. Polizzi, A. Ghosh, S. Datta, and M. Lundstrom. Theoretical investigation of surface roughness scattering in silicon nanowire transistors. *Applied Physics Letters*, 87(4):043101, 2005.

[65] M. Luisier, A. Schenk, and W. Fichtner. Atomistic treatment of interface roughness in si nanowire transistors with different channel orientations. *Applied Physics Letters*, 90(10):102103, 2007.

[66] M. H. Evans, X.-G. Zhang, J. D. Joannopoulos, and S. T. Pantelides. First-principles mobility calculations and atomic-scale interface roughness in nanoscale structures. *Phys. Rev. Lett.*, 95:106802, Sep 2005.

[67] G. Hadjisavvas, L. Tsetseris, and S. T. Pantelides. The origin of electron mobility enhancement in strained mosfets. *IEEE Electron Device Letters*, 28(11):1018–1020, Nov 2007.

[68] M. V. Fischetti and S. Narayanan. An empirical pseudopotential approach to surface and line-edge roughness scattering in nanostructures: Application to si thin films and nanowires and to graphene nanoribbons. *Journal of Applied Physics*, 110(8):083713, 2011.

[69] R. E. Prange and T. Nee. Quantum spectroscopy of the low-field oscillations in the surface impedance. *Phys. Rev.*, 168:779–786, Apr 1968.

[70] K. Moors, B. Sore, and W. Magnus. Modeling surface roughness scattering in metallic nanowires. *Journal of Applied Physics*, 118(12):124307, 2015.

[71] E. B. Ramayya, D. Vasileska, S. M. Goodnick, and I. Knezevic. Electron transport in silicon nanowires: The role of acoustic phonon confinement and surface roughness scattering. *Journal of Applied Physics*, 104(6):063711, 2008.

[72] S. Jin, M. V. Fischetti, and T. Tang. Modeling of surface-roughness scattering in ultrathin-body soi mosfets. *IEEE Transactions on Electron Devices*, 54(9):2191–2203, Sep. 2007.

[73] T. Ishihara, K. Uchida, J. Koga, and S. Takagi. Unified roughness scattering model incorporating scattering component induced by thickness fluctuations in silicon-on-insulator metal–oxide–semiconductor field-effect transistors. *Japanese Journal of Applied Physics*, 45(4B):3125–3132, apr 2006.

[74] D. Esseni. On the modeling of surface roughness limited mobility in soi mosfets and its correlation to the transistor effective field. *IEEE Transactions on Electron Devices*, 51(3):394–401, March 2004.

[75] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek. Surface roughness at the si(100)-sio$_2$ interface. *Phys. Rev. B*, 32:8171–8186, Dec 1985.

[76] L. van Hove. Quantum-mechanical perturbations giving rise to a statistical transport equation. *Physica*, 21(1):517 – 540, 1954.

[77] P. B. Vyas, M. L. Van de Put, and M. V. Fischetti. Quantum mechanical study of impact of surface roughness on electron transport in ultra- thin body silicon fets. In *2018 IEEE 13th Nanotechnology Materials and Devices Conference (NMDC)*, pages 1–4, Oct 2018.

[78] International roadmap for semiconductors 2.0, 2015 edition. `http://www.itrs2.net/itrs-reports.html`.

[79] V. V. Iyengar, A. Kottantharayil, F. M. Tranjan, M. Jurczak, and K. De Meyer. Extraction of the top and sidewall mobility in finfets and the impact of fin-patterning processes and gate dielectrics on mobility. *IEEE Transactions on Electron Devices*, 54(5):1177–1184, May 2007.

[80] M. J. H. van Dal, N. Collaert, G. Doornbos, G. Vellianitis, G. Curatola, B. J. Pawlak, R. Duffy, C. Jonville, B. Degroote, E. Altamirano, E. Kunnen, M. Demand, S. Beckx, T. Vandeweyer, C. Delvaux, F. Leys, A. Hikavyy, R. Rooyackers, M. Kaiser, R. G. R. Weemaes, S. Biesemans, M. Jurczak, K. Anil, L. Witters, and R. J. P. Lander. Highly manufacturable finfets with sub-10nm fin width and high aspect ratio fabricated with immersion lithography. In *2007 IEEE Symposium on VLSI Technology*, pages 110–111, June 2007.

[81] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta. Electron drift velocity in silicon. *Phys. Rev. B*, 12:2265–2284, Sep 1975.

[82] M. V. Fischetti, F. Gámiz, and W. Hänsch. On the enhanced electron mobility in strained-silicon inversion layers. *Journal of Applied Physics*, 92(12):7320–7324, 2002.

[83] M. V. Fischetti and S. E. Laux. Monte carlo study of electron transport in silicon inversion layers. *Phys. Rev. B*, 48:2244–2274, Jul 1993.

[84] A. Cresti, M. G. Pala, S. Poli, M. Mouis, and G. Ghibaudo. A comparative study of surface-roughness-induced variability in silicon nanowire and double-gate fets. *IEEE Transactions on Electron Devices*, 58(8):2274–2281, Aug 2011.

[85] M. V. Fischetti and W. G. Vandenberghe. *Advanced Physics of Electron Transport in Semiconductors and Nanostructures*. Springer Intl., Cham, Switzerland, 2016.

BIOGRAPHICAL SKETCH

Pratik B. Vyas is a doctoral student in the field of Electrical Engineering (EE) at The University of Texas at Dallas (UT Dallas). Under the supervision of Dr. Massimo Fischetti, his current research is focussed on the simulation of quantum transport to perform predictive analysis of the behavior of realistic semiconductor devices. He has developed simulation tools programmed in Python to achieve the same.

Previously, he received his Bachelor of Science in Electronics and Communications Engineering in 2012 from the Techno India University in Kolkata, India. Thereafter, he joined the Department of Electrical and Computer Engineering at UT Dallas in 2013 for the master's program in EE and obtained his degree in May 2015. During this time, he also conducted experiments under the supervision of Dr. Chadwin Young to optimize design flow of MOS capacitors with transition metal dichalcogenide (TMD) channel materials.

CURRICULUM VITAE

Pratik B. Vyas

Contact Information:

Department of Electrical and
Computer Engineering
The University of Texas at Dallas
800 W. Campbell Rd.
Richardson, TX 75080-3021, U.S.A.

Phone: (214) 681-7082
Email: `pbv130130@utdallas.edu`

Educational History:

B.S., Electronics and Communications Engineering, Techno India, 2012
M.S., Electrical Engineering, The University of Texas at Dallas, 2015
Ph.D., Electrical Engineering, The University of Texas at Dallas, 2019

Theoretical Study of Quantum Transport in Realistic Semiconductor Devices
Ph.D. Dissertation
Department of Electrical and Computer Engineering, The University of Texas at Dallas
Advisors: Dr. Massimo V. Fischetti

Employment History:

Research Assistant, The University of Texas at Dallas, August 2014 – present
Graduate Assistant, The University of Texas at Dallas, December 2013 – August 2014

Publications:

P. B. Vyas, C. Naquin, H. Edwards, M. Lee, W. G. Vandenberghe, and M. V. Fischetti, "Theoretical Simulation Of Negative Differential Transconductance In Lateral Quantum Well nMOS Devices", J. Appl. Phys. **121**, 044501 (2017)

Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti, "Simulation of Quantum Current in Double Gate MOSFETs: Vortices in Electron Transport", 2018 SISPAD, Austin, TX, 1 (2018)

Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti, "Quantum Mechanical Study of Impact of Surface Roughness on Electron Transport in Ultra-Thin Body Silicon FETs", 2018 IEEE 13th NMDC, Portland, OR, 1 (2018)

P. Zhao, P. B. Vyas, S. McDonnell, P. Bolshakov-Barrett, A. Azcatl, C. L. Hinkle, P. K. Hurley, R. M. Wallace, and C. D. Young, "Electrical characterization of top-gated molybdenum disulfide metal-oxide-semiconductor capacitors with high-k dielectrics", Microelectron. Eng. **147**, 2015

Conference presentations:

Pratik B. Vyas, Maarten L. Van de Put, and Massimo V. Fischetti, ”Impact of Interface Roughness on Quantum Transport in Double Gate MOSFETs”, IEEE SISC 2018

P. B. Vyas, C. Naquin, M. Lee, W. G. Vandenberghe, and M. V. Fischetti, ”Simulation Of Negative Differential Transconductance From Devices Fabricated Using Conventional CMOS Technology”, IOP IWCN 2017

P. B. Vyas, S. McDonnell, P. Bolshakov-Barrett, A. Azcatl, C. Hinkle, R. M. Wallace, and C. D. Young, ”Simple Test Structure for the Electrical Characterization of MoS2 Interface Functionalization with High-k Dielectrics”, IEEE SISC 2014

Community Involvement:

Reviewer for IEEE Transactions on Electron Devices and Journal of Computational Electronics
Member of Materials Research Society (MRS) UT Dallas chapter
Student member of IEEE and IEEE Young Professionals