Development of a Simulation Framework for 2D-Material MOSFETs: Investigating Cryogenic Behaviors and Enhancing Performance Optimization

by

Yiju Zhao

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2024

© Yiju Zhao 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Jesse Maassen Associate Professor, Dept. of Physics and Atmospheric Science, Dalhousie University
Supervisor(s):	Lan Wei Associate Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
	Youngki Yoon Associate Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal Member:	Dayan Ban Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
	Guoxing Miao Associate Professor, Dept. of Electrical and Computer Engineering, University of Waterloo
Internal-External Member:	Adam Wei Tsen Associate Professor, Dept. of Chemistry, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The relentless pursuit of miniaturization in semiconductor technology has ushered in the era of nanotransistors, which operate at the cutting edge of physical limits. This thesis presents significant advancements in nanoelectronics, focusing on the development of a numerical simulation tool for two-dimensional (2D) material-based Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) operating at cryogenic temperatures and exploring the potentials of materials like germanene (GeH) and hafnium disulfide (HfS_2) for future Complementary Metal-Oxide-Semiconductor (CMOS) technology. Central to this work is an advanced simulation framework capable of accurately predicting and characterizing 2D-based MOSFET behavior under cryogenic conditions, pivotal for cutting-edge applications such as control circuits for quantum computing. Furthermore, my research during PhD involves a comprehensive multi-level simulation approach, addressing substantial challenges in cryogenic device operation and laying a foundational framework for the design and optimization of electronic devices in extreme temperature environments while also advancing the understanding of material-device-circuit interplay for the efficient design of next-generation nanoelectronic circuits.

The simulation framework in the thesis utilizes the non-equilibrium Green's function (NEGF) formalism, incorporating temperature-dependent electrostatics, scattering, dissipation effects, carrier freeze-out, and band tail phenomena. This methodology provides a detailed understanding of the physical processes governing 2D materials in cryogenic environments. Through a rigorous alignment with experimental data, particularly with HfS_2 MOSFETs, the framework validates our model and underscores its potential to enhance cryogenic electronic device performance. This significant achievement marks a stride in using 2D materials for quantum computing applications, offering a sophisticated tool for their design and optimization. In addition, the thesis explores a holistic multi-level simulation approach tailored for 2D material-based nanoelectronics, featuring an in-depth case study on GeH MOSFETs. This segment covers device simulation, physics-based compact modeling, and circuit benchmarking, seamlessly bridging the gap between nanomaterial properties and circuit behavior. By revising the virtual source model to reflect the unique characteristics of 2D-material FETs and conducting HSPICE simulations, we have demonstrated substantial improvements in the energy-delay product by optimizing power supply and threshold voltages. This multi-faceted simulation process deepens the understanding of the synergy between materials, devices, and circuits and paves the way for the efficient design of futuristic nanoelectronic circuits.

Overall, this thesis underscores the critical role of cryogenic simulations in leveraging the full potential of 2D materials for advanced electronic applications, setting the stage for

a new era in cryo-CMOS and beyond. The insights from this research are expected to spark further investigations and developments, pushing the boundaries of what is achievable in nanoelectronics.

Acknowledgements

I am profoundly grateful to my thesis advisors and supervisors, Prof. Lan Wei and Prof. Youngki Yoon, for their exceptional guidance and support throughout my research journey. Their expert insights and relentless encouragement have been instrumental in advancing my work. Their mentorship transcended conventional boundaries, embedding in me a deep-seated appreciation for innovative thinking and rigorous academic discipline. From the early stages of my master's studies, Prof. Yoon was pivotal in steering me toward the cutting edge of my research field, enriching my academic voyage with his profound insights and visionary foresight. Prof. Wei, who co-supervised my journey through the Ph.D. program, demonstrated an unwavering commitment to research and education excellence. Their empathy and support for students and colleagues are a testament to the hallmark of true scholarship. Collaborating with Prof. Wei and Prof. Yoon has catalyzed an extraordinary personal and professional growth journey, facilitating my transition from an aspiring student to a seasoned professional in our field. Their dedication and passion for science have indelibly shaped my professional ethos, instilling lifelong gratitude for their guidance and mentorship.

I am pleased to acknowledge Professors Guoxing Miao, Danyan Ban, Adam Wei Tsen, and Jesse Maassen for serving on my Doctoral Committee and providing valuable comments and suggestions on this work. I am also grateful to Prof. Miao, an exceptional lecturer who guided me in device physics. His insightful teachings have equipped me with a comprehensive understanding of the intricacies of semiconductor physics.

I extend my heartfelt gratitude to my colleague, Demin Yin, for imparting his knowledge of numerical techniques and his invaluable assistance in all aspects of research and life. Special thanks go to Abdulaziz Almutairi and Hyunjae Lee for their guidance on the tight-binding approach to atomistic bandstructure calculations and our enriching discussions in the office. I am thankful for the opportunity to collaborate with Mayuri Sritharan on constructing tight-binding parameters for HfS_2 and to Michael Spinazze for his diligent effort in gathering all the simulation data for the cryogenic simulations. Their collective expertise and support have significantly enhanced the quality and depth of this research. I am also grateful to Gyu Chull Han for his contributions and insightful suggestions regarding the atomistic simulations of MoS_2 transistors, and to Robert KA Bennett for his support in studying the anisotropic permittivity of materials, and to Manasa Kaniselvan for her assistance with the K-p method and scattering implementation in NEGF.

I am profoundly grateful for the financial assistance provided during my studies, which includes the Graduate Research Studentship, the Richard and Elizabeth Madter Scholarship

from the University of Waterloo, and the Nanofellowship awarded by the Waterloo Institute for Nanotechnology. Additionally, I sincerely appreciate the Digital Research Alliance of Canada for supplying the essential computing resources that significantly enhanced this study. Their collective support has been indispensable to my research and academic progress.

Dedication

To my parents, whose endless encouragement is my stronghold; To my wife, whose boundless love is my constant; To my child, whose laughter is the light of our home.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Abstract iv

Acknowledgements vi

Dedication viii

List of Figures xiii

List of Tables xviii

1 Introduction 1

1.1 Quantum Computing Platform 1

1.2 Cryogenic CMOS 4

1.2.1 Numerical Simulation for Cryogenic CMOS 5

1.3 2D-material-based MOSFETs 7

1.3.1 2D-based MOSFETs for Quantum Computing 9

1.4 Motivation 10

1.4.1 Numerical Simulation of 2D-based Cryo-MOSFETs 10

1.4.2 Multi-Level Simulation for Nanotransistors 13

2 Cryogenic CMOS Transistor **16**
2.1 Characteristics and behavior of Cryo-MOSFETs . . . 16
2.1.1 Inflection . . . 17
2.1.2 Carrier Mobility . . . 18
2.1.3 Threshold Voltage Shifts . . . 19
2.1.4 Reduced Thermal Energy on Electron Transport . . . 20
2.2 Device Physics of Cryo-MOSFETs . . . 22
2.2.1 Fermi Distribution Function . . . 23
2.2.2 Incomplete Ionization . . . 24
2.2.3 Interface Trapped States . . . 25
2.2.4 Band Tail Effect . . . 26
2.3 Summary . . . 28

3 Modeling and Simulation Method **29**
3.1 NEGF-Poisson Formalism . . . 29
3.1.1 Electrostatic Potential Calculation . . . 32
3.1.2 Non-Equilibrium Green's Function Method . . . 34
3.2 Electrostatic at Cryogenic Temperature . . . 40
3.2.1 Incomplete Ionization . . . 40
3.2.2 Interface Traps . . . 42
3.3 Scattering and Band Tail: Büttiker Probe Method . . . 43
3.3.1 NEGF with Scattering . . . 43
3.3.2 Büttiker Probe Self Energy . . . 45
3.3.3 Current Conservation . . . 45
3.4 Simulation Robustness . . . 47

4 Simulation of HfS_2 MOSFETs at Cryogenic Temperature **49**
4.1 Simulation Setup . . . 50
4.2 Scattering Self-energy . . . 51
4.2.1 Resistance vs Channel Length . . . 52
4.2.2 Mobility Calibration . . . 54
4.3 Ionization: Complete vs. Incomplete . . . 56
4.4 Subthreshold Slope Inflection Analysis . . . 57
4.4.1 Interface Trapped States . . . 58
4.4.2 Effect of Band Tail . . . 59
4.5 Channel Length Scaling . . . 60
4.6 Temperature Dependence . . . 62
4.7 Summary . . . 63

5 Multi-Level Simulation Scheme **65**
5.1 Introduction . . . 65
5.1.1 Simulation Overview . . . 67
5.2 Materials Parameterization . . . 69
5.2.1 DFT Calculations . . . 69
5.2.2 Tight-Binding (TB) Parameterization . . . 72
5.3 Device Simulation . . . 77
5.3.1 Simulation Flowchart . . . 77
5.3.2 GeH MOSFET . . . 77
5.4 MVS Compact Model . . . 81
5.4.1 DOS and Injection Velocity . . . 84
5.4.2 Quantum Capacitance . . . 85
5.4.3 Fitting Parameters . . . 86
5.5 HSPICE Simulation . . . 87
5.6 Summary . . . 90

6 Conclusion **91**

6.1 Summary 91

6.2 Outlook 92

6.3 Future Work 93

References **95**

List of Figures

1.1 The quantum-classical interface of a quantum computer. (a) illustrates the essential components of a quantum computing stack. (b) details the control and readout subsystems spanning room temperature (RT) environments to 100 mK, highlighting the integration of a readout transceiver equipped with a field-programmable gate array (FPGA), digital-to-analog converter (DAC), and analog-to-digital converter (ADC) for simultaneous multiple qubit readouts. The cryo-CMOS chip (represented in brown) is specifically designed to alleviate the input-output (IO) bottleneck associated with control signal processing [1]. 2

1.2 The device structure of a single-gate MOSFET used in the Sentaurus Simulation. 11

1.3 Drain current versus gate voltage plotted with (a) a long channel device ($L_{ch} = 1.75\,\mu m$) where the current at $T = 20\,K$ fails to converge and (b) a short channel device ($L_{ch} = 65\,nm$) where current converged to an unreasonable saturation. 12

2.1 Log I_D vs V, traces at three different temperatures [2]. 17

2.2 Influence on current at 4.2 K due to incomplete ionization ($g_A = 4$), resulting in reduced threshold voltage [3]. 20

2.3 Band diagrams depicting distinct cryogenic models with conduction band edge (E_C) and Fermi energy level (E_f) provided as reference in (a) the Fermi-Dirac distribution of room temperature and cryogenic temperature, (b) a comparison of ionization energy (Φ_A) with thermal energy (Φ_b) and visualization of interface trap states with trap density (N_{IT}), and (c) the impact on the Density of States (DOS) below the conduction band edge (E_c) resulting from band tail states. 22

3.1 (Top) Schematic structure of device coupled to the source and drain contacts. (Bottom) Self-consistent procedure for the coupled NEGF–Poisson computational simulation framework [4]. 30

3.2 (a) A structure diagram of ultra-thin body double-gate MOSFET, within a two-dimensional simulation domain outlined by a solid line. The domain utilizes uniformly spaced grids in both the x and z directions, with the spatial constants denoted as a and b, respectively. This grid arrangement facilitates detailed electrostatic analysis within the specified simulation area, ensuring comprehensive coverage and precision in modeling the device's behavior [5]. (b) The computational model adopts a 5-point star configuration, denoted as a 'computational molecule,' to effectively address the electrostatic calculations required for the device's analysis [6]. 32

3.3 Schematic representation illustrating the computational efficiency gained by substituting the extensive Hamiltonian of the reservoir, H_R, with a self-energy term, Σ. This approach simplifies the calculations without losing the physical significance of the reservoir's impact on the device [7]. 35

3.4 Overview of the NEGF-Poisson method incorporating cryogenic models in the Poisson module and the Büttiker probe within the NEGF module. . . . 41

3.5 Overview of Büttiker Probe Calculations. **Top**: Depicts a quasi-one-dimensional (Quasi-1D) semiconductor device featuring two actual contacts, represented as discretized segments for computational purposes. **Middle**: Illustrates the approach for modeling the semi-infinite contacts by incorporating self-energies at the device terminals, thereby emulating the effect of infinite contact extension. **Bottom**: Demonstrates the integration of Büttiker probe self-energies into the device model to simulate various scattering phenomena. This is achieved by strategically placing a comprehensive array of Büttiker probes along the device, facilitating the emulation of diverse scattering mechanisms within the simulation framework [8]. 44

4.1 Graphical illustration of the device structure employed in the simulations, showcasing the configuration and arrangement of its elements. 50

4.2 Supercell band structure of HfS_2. 51

4.3 Benchmark results for scattering effects with η ranging from 1×10^{-3} eV to 5×10^{-3} eV: Calculated resistance is shown in (a) for $T = 77$ K and in (b) for $T = 300$ K, while the mean free path (λ) is presented in (c) at $T = 77$ K and in (d) at $T = 300$ K. 53

4.4 Drain current (I_D) versus gate voltage (V_G) plots comparing ballistic transport (lines) with scattering-inclusive models (markers) at temperatures of $T = 300$ K (in red) and $T = 50$ K (in blue). 55

4.5 Comparison of the simulated transfer characteristic at $T = 77$ K, showcasing devices with incomplete ionization (markers), complete ionization (solid lines), and the intrinsic devices (dotted lines). 57

4.6 Comprehensive parametric study at 50 K (left) and 300 K (right) for varying interface trap density (N_T) and trap distribution width (W_T): (a) A marked inflection effect is observed as N_T is increased from 5.7×10^{12} cm^{-2} to 2.9×10^{13} cm^{-2}, which notably decreases the ON-state current and alters its curvature in moderate inversion. The onset of inflection remains consistent, attributed to the fixed Gaussian width. (b) Deterioration of SS due to inflection and tunneling effects, with SS increasing at the lower end of I_D. (e) and (f) Enhanced W_T from 20 meV to 80 meV at the inflection initiation point ($N_T = 1.7 \times 10^{13}$ cm^{-2}). (c)-(d) and (g)-(h) Display results at 300 K for comparison, showing minimal inflection in SS and limited impact of trap states on the ON-state current. 58

4.7 (a) Illustration of drain current (I_D) versus gate voltage (V_G) showcasing a range of band tail lengths from 25 meV to 100 meV (solid lines), compared to the ballistic case (markers). The data is presented for two temperature conditions: $T = 50$ K (in blue) and $T = 300$ K (in red).(b) LDOS plotted for a 18 nm-MOSFETs with $\kappa = 100$ meV at $T = 50$ K. 60

4.8 Simulated performance of a 18 nm HfS_2 MOSFET incorporating full cryogenic models: (a) Examining the effect of increasing interface trap density (N_T) with a constant trap distribution width (W_T); (b) Analyzing the impact of increasing W_T while maintaining a fixed N_T; (c) Comparing device performance for two different channel lengths, $L_{ch} = 28$ nm and $L_{ch} = 8$ nm, illustrating how the influence of cryogenic models is altered by channel scaling. 61

4.9 Temperature scaling from 50 K up to 300 K. 63

5.1 An overview of the multi-level simulation. It starts with intrinsic device simulation, including materials parameterization and numerical device simulation. A compact model bridges the gap between device and circuit simulations. . 68

5.2 Two-center Slater-Koster integrals for atomic orbitals s, p, and d, illustrating bond types like σ, π, and δ [9]. 74

5.3 The electronic band structure of germanane (GeH) as determined through Density Functional Theory (DFT) is represented by blue circles, while the solid red lines indicate the fit achieved through tight-binding (TB) parametrization. The inset presents a top view of the germanane lattice, with the unit cell demarcated by red dotted lines [10]. 75

5.4 The flowchart demonstrating the iterative NEGF and Poisson self-consistent simulation methodology for GeH FETs. This illustrative flowchart visually elucidates the intricacies of the iterative procedure employed to solve the Poisson and NEGF equations concurrently. This iterative approach culminates in precisely determining the electrostatic potential and electron density of the GeH FET device, facilitating a comprehensive understanding of its electronic behavior. 78

5.5 Drain current of GeH FETs at drain bias ($|V_{DS}|$) ranging from 0 V to 0.7 V and gate bias ($|V_{GS}|$) from 0.5 V to 1.1 V. Results from both NEGF simulations (dots) and the modified MVS compact model (solid lines) are shown. I_{DS}-V_{DS} for (a) NMOS and (d) PMOS; I_{DS}-V_{GS} for (b) NMOS and (e) PMOS where DIBL vs. V_{DS} are plotted in the insert in (b) and (e) for both NMOS and PMOS, respectively. Total channel charge as a function of gate voltage Q_{ch}-V_{GS} are also plotted for (c) NMOS and (f) PMOS at different drain voltages (V_{DS}). 79

5.6 Potential profiles of (a) the valence band for PMOS at $V_D = -0.7$ V and (b) the conduction band for NMOS at $V_D = 0.7$ V, with different gate voltages. The location of the top of the barrier for each gate voltage is indicated with red and blue dots for PMOS and NMOS, respectively. 83

5.7 (a) Charge density (ρ_{vs}) and (b) virtual-source velocity (v_{vs}) at the virtual source/top of barrier point are plotted as a function of gate voltage, respectively, at $V_D = 0.7$ V. Charge densities are plotted on both a linear scale and a logarithm scale. NEGF simulation results are shown in blue square dots, while results from the MVS model are in black solid curves. 83

5.8 (a) Energy-resolved density of state (DOS) for GeH and silicon. For comparison, the unit of DOS for GeH was converted to $eV^{-1}cm^{-3}$ by dividing 2D DOS of GeH with its thickness (0.55 nm). (b) The bandstructure-limited velocity of GeH is plotted for both electrons (blue solid line) and holes (red solid line). The red dashed line shows the velocity of light holes and that of heavy holes by the red dotted line. 84

5.9 Oxide capacitance (C_{ox}, black solid line), quantum capacitance (C_Q, colored solid line), and total gate capacitance (C_G, dash-dot line) for (a) NMOS and (b) PMOS at $V_D \approx 0V$. 85

5.10 A power supply voltage window (grey box) is plotted to illustrate the relation of the reference voltages at which circuit parameters, such as propagation delay, power dissipation, and energy-delay product, are evaluated. 88

5.11 (a) Total power dissipation (P_T), propagation delay (τ_p) and (c) energy-delay product (EDP) vs. threshold voltage (V_{TH}) at different power supply voltages of $V_{DD} = 0.7\,V$ and $1.0\,V$ and (e) Color map of 1/EDP as a function of V_{DD} and V_{TH} are plotted with a higher contact resistance ($R_c = 1\,k\Omega \cdot \mu m$). The corresponding results with a lower contact resistance $R_c = 0.1\,k\Omega \cdot \mu m$ are plotted in (b), (d), and (f). In colormap, a warmer color indicates a smaller EDP. The optimal operation point (i.e, the minimum EDP point) is located at $V_{DD} = 0.8\,V$ and $V_{TH} = 0.5\,V$ with $EDP = 9.8 \times 10^{-3}\,fJ \cdot ps$ for $1\,k\Omega \cdot \mu m$ in (e) and at $V_{DD} = 0.9\,V$ and $V_{TH} = 0.5\,V$ with $EDP = 3.1 \times 10^{-3}\,fJ \cdot ps$ for $0.1\,k\Omega \cdot \mu m$ in (f). 89

List of Tables

4.1 Scattering Benchmark Parameters 54

5.1 Fitting Parameters for NMOS and PMOS. 86

Chapter 1

Introduction

1.1 Quantum Computing Platform

Quantum computing has emerged as a transformative technology that holds the potential to revolutionize computation and problem-solving, including cryptography, optimization, simulation, and machine learning [11, 12, 13]. Unlike classical computers, which rely on bits to process information, quantum computers leverage the principles of quantum mechanics to perform computations, offering the prospect of solving complex problems exponentially faster. Quantum computers can be faster than classical computers due to their ability to leverage quantum superposition, parallelism, and entanglement. These properties allow quantum computers to perform multiple computations simultaneously and explore different solution paths, potentially leading to exponential speedups for specific problems.

While the advantages of quantum computing are promising, they are still in their early stages of development, and significant challenges lie ahead. Overcoming noise and decoherence in qubits, scaling up the number of qubits while maintaining their coherence, and developing robust error correction techniques are critical challenges to achieving fault-tolerant quantum computing [14]. Furthermore, quantum software development, quantum algorithms, and quantum control methodologies need to be refined to exploit the full potential of quantum computers in practical applications.

In the past few decades, researchers have proposed several implementations of quantum computers. Among all proposals, superconducting qubits [15, 16], iron traps [17], photonic qubits [18], neutral atoms [19] and topological qubits [20] are being actively researched and developed for their potential to achieve scalable and robust quantum computing capabilities.

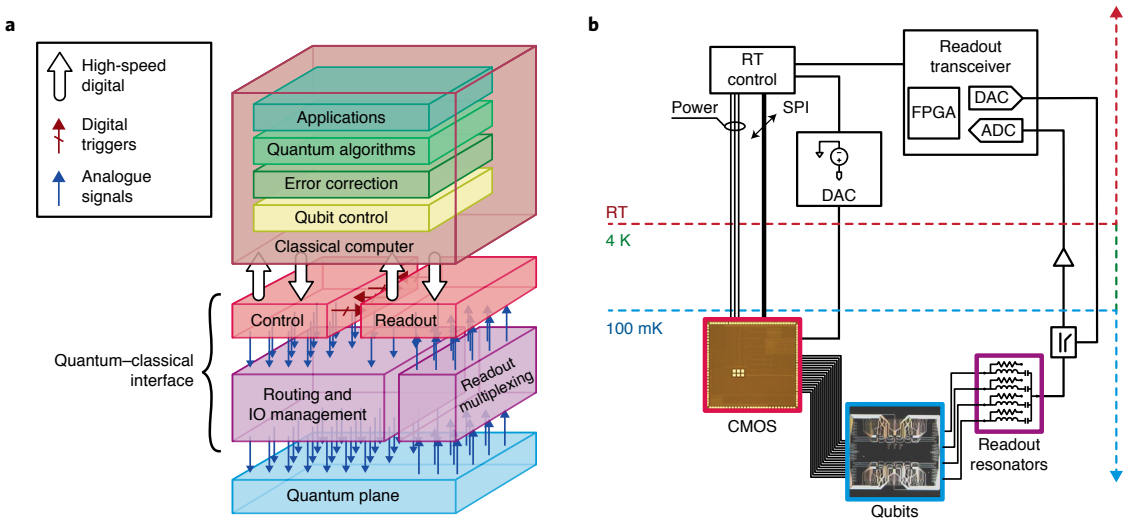


Figure 1.1: The quantum-classical interface of a quantum computer. (a) illustrates the essential components of a quantum computing stack. (b) details the control and readout subsystems spanning room temperature (RT) environments to 100 mK, highlighting the integration of a readout transceiver equipped with a field-programmable gate array (FPGA), digital-to-analog converter (DAC), and analog-to-digital converter (ADC) for simultaneous multiple qubit readouts. The cryo-CMOS chip (represented in brown) is specifically designed to alleviate the input-output (IO) bottleneck associated with control signal processing [1].

A state-of-the-art quantum computer comprises a complex hierarchy of interconnected components and subsystems that work harmoniously to execute quantum computations [1]. This hierarchical structure ensures the precise control and manipulation of qubits and the efficient communication and processing of quantum information.

In Figure 1.1, we could see an example of the diagram of a superconducting quantum computer. The system comprises three modules: quantum processor, cryogenic electronics, and CMOS devices from bottom to top with low to high temperatures. While different quantum computing approaches may have additional requirements and configurations, integrating cryogenic electronics, CMOS circuits, and the quantum processor is employed in almost all quantum computers. The quantum processor is a quantum computer's core component responsible for performing quantum computations. It consists of qubits and the necessary control systems to manipulate and measure these qubits. The quantum processor executes quantum algorithms by applying quantum gates to the qubits, allowing for the execution of parallel computations and exploiting quantum phenomena like superposition

and entanglement. Cryogenic electronics refer to the electronic components that operate at extremely low temperatures, typically near absolute zero ($T < 4\,\mathrm{K}$). Cryogenic electronics are critical for providing the conditions to manipulate and control qubits while maintaining their fragile quantum states and minimizing the effects of noise and decoherence. The cryogenic devices include low-temperature sensors, amplifiers, signal generators, and other integrated circuits designed to operate reliably in cryogenic environments. CMOS circuits are commonly used in classical computers and utilized in some aspects of quantum computing. While the primary processing in a quantum computer occurs using qubits and quantum gates, specific control systems and ancillary operations can be implemented using classical electronic components based on CMOS technology. These CMOS circuits control and interface with the quantum processor, manage qubit operations, error correction, and measurement, and communicate between classical and quantum systems.

1.2 Cryogenic CMOS

Quantum computing's success critically depends on creating a cryogenic environment, approaching the vicinity of absolute zero, which is essential for qubits' stability - the essence of quantum computing. Cryogenic CMOS (cryo-CMOS) technology, pivotal in this venture, transcends the realm of mere thermal management [21]. It embodies advanced cooling methods and the sophisticated design of integrated circuits, proving indispensable for sustaining ultra-low temperatures. This technological advancement paves the way for the monolithic integration of control and readout circuits, directly addressing the scalability challenges of the extensive qubit requirements intrinsic to quantum computing [22]. Through experimental demonstrations, including the development of noise-canceling low-noise amplifiers and class-F oscillators optimized for cryogenic operation, cryo-CMOS technology emerges as a foundational cornerstone, promising a scalable solution for future quantum computers. This comprehensive approach facilitates the practical realization of quantum computing and marks a significant leap toward harnessing the power of quantum mechanics for computational breakthroughs.

The operation of quantum processors within cryogenic temperatures is a fundamental prerequisite for ensuring the stability and coherence of qubits. This cryogenic milieu significantly diminishes decoherence mechanisms, predominantly thermal noise, thereby preserving the fragile quantum states indispensable for computation [23]. Cryogenic CMOS (cryo-CMOS) technology assumes a critical role in this context, as it enables the achievement and maintenance of ultra-low temperatures and fosters a scalable solution for the future of quantum computing.

Decoherence, primarily engendered by thermal noise, presents a formidable challenge in maintaining coherent quantum states. With its state-of-the-art cooling techniques and integrated circuit design, Cryo-CMOS technology surmounts this challenge by offering a stable cryogenic environment crucial for quantum computation. The capability of cryo-CMOS to operate efficaciously at temperatures as low as 4 K facilitates closer system integration and heralds a scalable solution to empower future quantum computers.

Confronting the multifaceted engineering challenges inherent in creating a viable cryogenic environment for quantum computing involves developing efficient cooling techniques to reach and maintain temperatures near absolute zero and the fabrication of integrated circuits capable of reliable operation under such extreme conditions. Cryo-CMOS technology, leveraging advanced cooling methods such as dilution refrigerators, adiabatic demagnetization refrigeration (ADR), and pulse tube refrigerators, stands at the vanguard of attaining these ultra-low temperatures. This sophisticated technology enables operation at temperatures

as low as 4 K, facilitating closer system integration and offering a scalable, pivotal solution for advancing future quantum computers. The amalgamation of cryo-CMOS with quantum computing signifies a profound advancement, merging the domains of low-temperature physics and semiconductor technology to meet the complex requirements of quantum computation.

Cryogenic Metal-Oxide-Semiconductor Field-Effect Transistors (cryo-MOSFETs) play a crucial role in the progression of cryo-CMOS technology for quantum computing applications. Despite significant advancements in cryo-CMOS technology, culminating in the successful experimental demonstration of essential circuit elements tailored for cryogenic operation, a discernible research gap persists in the comprehensive understanding and simulation of cryo-MOSFETs. This gap poses a considerable obstacle to fully leveraging the potential of cryo-CMOS technology within the quantum computing arena.

1.2.1 Numerical Simulation for Cryogenic CMOS

Here, we would like to briefly discuss the needs and challenges posed to numerical simulation. With the excitement in exploring cryogenic CMOS in recent years, the need for more accurate numerical simulations of cryogenic CMOS has become increasingly evident. At these temperatures, the properties of transistors and other semiconducting devices perform differently from those at room temperature. Complementing experimental exploration, numerical simulation has always been an effective tool in studying material properties and device physics. Given the significant effort and resources required for experiment and characterization at cryogenic temperature, such a tool will arguably play a more important role in cryogenic CMOS.

Cryogenic CMOS numerical simulation, as expected, involves modeling the behavior of CMOS devices at extremely low temperatures, from a few Kelvin down to below 100 mK. Besides the significant physical effects (e.g., incomplete ionization) at cryo-T to be included in cryo-TCAD tools, the very small temperature value poses numerical challenges for computation and convergence. For example, Newton's method is a numerical technique commonly used to solve the nonlinear system of equations arising from the discretization of the drift-diffusion equations. The issue of convergence at low temperatures is due to Newton's method being prone to overshooting, where the length of the correction step for a given iteration is overestimated. This can lead to slow convergence and require numerous iterations to obtain a solution. To address this issue, a damping parameter can be used to restrict the size of each correction step and improve convergence. For each bias condition, Newton's method requires an initial approximation to find the solution. However, Newton's

method will fail to converge if the initial guess does not adequately approximate the actual solution (i.e. if it is not within the solution's region of attraction).

Loss of numerical precision can also cause Newton's method to fail to converge, as subtracting nearly equal numbers using finite precision arithmetic may result in corrupted update vectors that move the solution approximation outside the solution's region of attraction. At low temperatures, double precision arithmetic (8-byte) is inaccurate enough to resolve solution updates for a given iteration. Moreover, the device equations involve both positive and negative exponential functions of temperature, making the system more ill-conditioned and sensitive to precision loss and other sources of numerical error. Consequently, convergence becomes increasingly challenging with decreasing temperature. Critical calculations are performed with extended (16-byte) precision at temperatures below 150K to overcome this issue. While extended-precision arithmetic may slow down the computation compared to normal arithmetic, it is necessary for obtaining accurate low-temperature solutions.

1.3 2D-material-based MOSFETs

In pursuing more powerful and energy-efficient electronic devices, the boundaries of traditional semiconductor technologies, encapsulated by the well-known Moore's Law, are increasingly being met. The continuous miniaturization of conventional transistors is encountering substantial hurdles, including escalated power dissipation, leakage currents, and the complexities of advanced manufacturing techniques. In response to the growing appetite for faster, more efficient electronics, the scientific community has shifted its focus to pioneering materials and innovative device architectures. These new directions promise to circumvent existing limitations and herald a transformative computing and information-processing era. This exploration prominently features 2D-based field-effect transistors (FETs), which emerge as a groundbreaking avenue with the potential to redefine the landscape of electronics and set new benchmarks for semiconductor technology. This section delves into the comprehensive study of 2D Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs), examining their fundamental principles, fabrication methodologies, and diverse applications. We will explore the exceptional characteristics of pivotal 2D materials such as graphene, transition metal dichalcogenides (TMDs), and black phosphorus, highlighting their suitability for 2D MOSFET development. Furthermore, we will contrast the unique attributes of 2D MOSFETs with traditional MOSFETs, focusing on aspects like scaling impacts, carrier confinement, and quantum effects to illuminate the complexities and opportunities within the realm of 2D MOSFET technology.

Over the past decade, the quest to harness the unique properties of 2D materials for cutting-edge electronic devices has captured the imagination of the research and engineering communities. Key players in this revolution include graphene, TMDs, and black phosphorus, each demonstrating exceptional potential in a broad spectrum of applications, such as FETs, memory devices, and optoelectronics. A series of pioneering discoveries in 2D materials underpin the emergence of 2D nanoelectronics. Graphene, a trailblazer with its extraordinary single-layer carbon structure, exhibits unprecedented electrical conductivity and mechanical strength, captivating researchers with its vast possibilities. TMDs, exemplified by molybdenum disulfide (MoS_2) [24], are known for their distinctive semiconductor properties, including a substantial band gap and excellent electrostatic integrity, making them ideal for low-power switching applications. Black phosphorus [25], recognized for its anisotropic behavior, offers enticing prospects in device engineering and performance optimization. GeH [26, 27], well-known for its high theoretical predicted mobility, is a potential channel material for next-generation high-performance devices. The unique band structures, high carrier mobility, and robust electron-electron interactions inherent in 2D materials pave the way for ultrafast electronics, flexible wearable technology, and quantum

computing applications. By leveraging these distinct properties, researchers aspire to engineer transistors that surpass the performance of conventional counterparts, potentially ushering in a new era of faster, smaller, and more energy-efficient electronic devices, which could transform industries and redefine the technological landscape.

Among the myriad applications envisioned for 2D materials, their use in field-effect transistors (FETs) is particularly promising. FETs, the cornerstone of modern electronics, govern electron flow in semiconductor channels via an external electric field. While predominantly silicon-based, traditional FETs have been the semiconductor industry's mainstay for decades, the shift towards nanoscale dimensions introduces new physical phenomena and challenges. For example, quantum mechanical effects, such as carrier confinement and tunneling, become increasingly significant as devices shrink. These issues and the inherent limitations of bulk materials have steered researchers toward exploring alternative materials and device architectures. The exceptional properties of 2D materials, including high carrier mobility and saturation velocity, have led to groundbreaking progress in FET technology, positioning them as strong contenders for next-generation electronics. Graphene, with its unmatched carrier mobility, has spurred extensive research, particularly in high-frequency applications, igniting a flurry of innovation in the field of nanoelectronics. TMDs, recognized for their wide band gap and superior electrostatic integrity, have emerged as attractive options for low-power switching applications. The significant band gap of TMDs enables efficient on-off switching, ideal for low-power devices. At the same time, their exceptional electrostatic integrity minimizes leakage current, enhancing energy efficiency and prolonging battery life in portable electronics.

In the evolving landscape of 2D nanoelectronics, developing 2D metal-oxide-semiconductor field-effect transistors (MOSFETs) has garnered significant attention. These devices, which merge the intrinsic properties of 2D materials with traditional MOSFET architecture, promise to revolutionize semiconductor technology by offering unparalleled performance and energy efficiency. The distinct electronic properties of atomically thin materials are exploited in 2D-based MOSFETs, setting them apart from conventional silicon-based FETs, where bulk material properties are predominant. Quantum confinement effects in 2D materials lead to exceptional electron mobility and reduced scattering, enhancing current capacity and switching speeds. The atomically thin nature of these materials allows for highly scalable devices with precise thickness control, creating devices with customized electrical and optical properties and paving the way for next-generation electronics.

The engineering of 2D-based FETs offers fertile ground for innovation. Researchers can tailor band structures and electronic properties by stacking different 2D materials or creating heterostructures. These heterostructures, composed of varied 2D materials,

combine multiple functionalities in a single device, expanding the horizons for device innovation. This opens doors to exciting new concepts like tunable bandgaps, valleytronics, and spintronics, which hold the potential to revolutionize modern electronics and chart the course for future devices and applications.

1.3.1 2D-based MOSFETs for Quantum Computing

The quest for leveraging 2D materials in cutting-edge electronic devices, particularly in quantum computing applications, has captivated the research and engineering communities. The distinct electronic properties of atomically thin materials, such as high carrier mobility and saturation velocity, and their quantum confinement effects position 2D-based MOSFETs as prime candidates for next-generation electronics. These properties facilitate the development of devices with unprecedented performance and energy efficiency and open new avenues for quantum computing. Quantum computing benefits immensely from integrating 2D materials in MOSFETs, primarily due to the precise control of electronic properties and the potential for reduced device dimensions without compromising performance. The exceptional electrostatic control offered by 2D materials minimizes leakage currents, a critical factor for the stable operation of qubits. Moreover, the high carrier mobility in 2D materials enhances quantum operations' speed, contributing to quantum processors' efficiency and scalability. Given these compelling attributes, 2D-based MOSFETs emerge as a promising solution for overcoming the challenges in cryogenic MOSFET applications within quantum computing. The inherent advantages of 2D materials, such as their stability at low temperatures and the potential for integrating with silicon technology, make them ideal for developing cryogenic MOSFETs. These advancements are poised to enhance the performance and scalability of quantum computers and pave the way for new computational paradigms.

1.4 Motivation

The goal of my PhD research unfolds across two distinct yet interconnected segments. On the one hand, my research is to study the cryogenic CMOS behavior of 2D-based MOSFETs by developing an advanced atomistic simulation based on Tight-binding NEGF. On the other hand, developing a multi-scale simulation tool bridges the gap between the intrinsic device performance and circuit analysis for novel nanoelectronics based on 2D channel materials. This chapter aims to lay the groundwork by shedding light on the driving forces behind each segment, thereby setting the stage for their in-depth exploration in the subsequent sections.

1.4.1 Numerical Simulation of 2D-based Cryo-MOSFETs

The first segment of my research delves into the intricate realm of cryo-CMOS simulation, which has gained significant traction in recent years. The motivation for this exploration is rooted in the burgeoning field of quantum computing, which demands a deep understanding of electronic behaviors at cryogenic temperatures. This segment aims to unravel the complexities of electronics under such extreme conditions, a quest fundamental to quantum computing technologies' evolution and success. By focusing on cryo-CMOS simulation, this research seeks to contribute meaningfully to the advancement of quantum computing, addressing the challenges posed by these low-temperature environments [28].

The Failure of Conventional Simulation Tool

Within the realm of semiconductor device simulation, commercial software packages like Synopsys Sentaurus TCAD excel at modeling devices operating under standard room temperature conditions (300 K). However, when temperatures descend into the cryogenic regime (4 K and below), these tools often reveal their limitations. Commercial TCAD (Technology Computer Aided Design) software, commonly used in device research at RT and above, fails to reach converged results at deep-cryogenic temperatures due to numerical challenges and the lack of important cryogenic effects. TCAD, designed explicitly for deep-cryogenic temperatures, is still in its early stages. Companies like Nanoacademic Technologies are actively pushing the frontier in developing commercial TCAD tools targeting cryogenic and quantum applications.

Using the device structure (see fig. 1.2), we examined the simulation results with Sentaurus TCAD from room temperature down to cryogenic regime. The results are

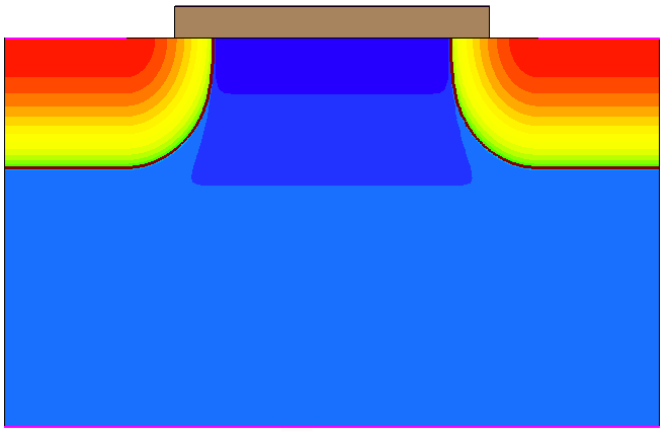


Figure 1.2: The device structure of a single-gate MOSFET used in the Sentaurus Simulation.

plotted in fig. 1.3. Simulation failures become increasingly prevalent, with convergence issues typically arising around 25 K. This breakdown has its roots in the fundamental difficulty of convergence, as the potential change is more susceptible to the fluctuation in carrier density when T becomes smaller. The failure can also be attributed to the absence of crucial cryogenic phenomena within the fundamental modeling frameworks of these simulators. The physics governing Si-MOSFET behavior undergoes a profound transformation at extremely low temperatures. Carrier freeze-out dramatically reduces the availability of free charge carriers as a significant portion of dopant atoms cease to ionize fully. Furthermore, the ionization energy of these dopants increases, further suppressing the carrier concentration. Simultaneously, lattice vibrations that typically impede carrier flow diminish, leading to a surge in carrier mobility. In the most extreme cryogenic environments, quantum mechanical effects such as carrier confinement and tunneling can no longer be ignored. Standard simulators, fundamentally calibrated for room-temperature behavior, lack the specialized models necessary to capture the nuances of these cryogenic effects. This shortcoming severely restricts their ability to accurately predict, analyze, and optimize cryogenic Si-MOSFET performance, highlighting the inherent limitations of state-of-the-art commercial software in this increasingly important domain.

2D-based MOSFETs at Cryogenic Temperature

Semiconductor devices based on two-dimensional (2D) materials, such as transition metal dichalcogenides (TMDs) and graphene, have emerged as promising candidates for next-generation electronics due to their unique electronic properties that enable superior

performance and functionality[29, 30, 31, 32, 24]. Among these devices, Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) play a pivotal role in modern integrated circuits, as they are the building blocks of logic and memory devices[33, 34]. However, their performance at cryogenic temperatures, a critical operational condition for certain applications like quantum memory and quantum computing[35, 36, 21, 37], electronics in space exploration[38], and high-performance cryogenic device[39], remains inadequately understood and poses a significant challenge for practical implementation. The issue revolves around the lack of comprehensive models that account for the intricate interplay of various physical phenomena influencing MOSFET behavior under cryogenic conditions[40, 41].

Understanding and predicting the performance of MOSFETs at cryogenic temperatures is crucial for emerging applications such as quantum computing, which require low-noise and high-speed operation[42, 1]. However, the device can behave quite differently in cryogenic temperatures compared to room-temperature conditions. Therefore, new models should be included to characterize the cryogenic devices. At low temperatures, quantum effects become more pronounced, and the behavior of carriers is governed by phenomena such as scattering[43, 44], band tailing[45, 46], incomplete ionization of dopants[3], and the presence of interface trap states[47]. These phenomena affect carrier transport, device characteristics, and device reliability, and they need to be accurately captured by device models. Conventional modeling approaches, such as drift-diffusion, struggle to capture these subtleties accurately, leading to a gap in our understanding of the cryogenic operation of 2D material-based MOSFETs.

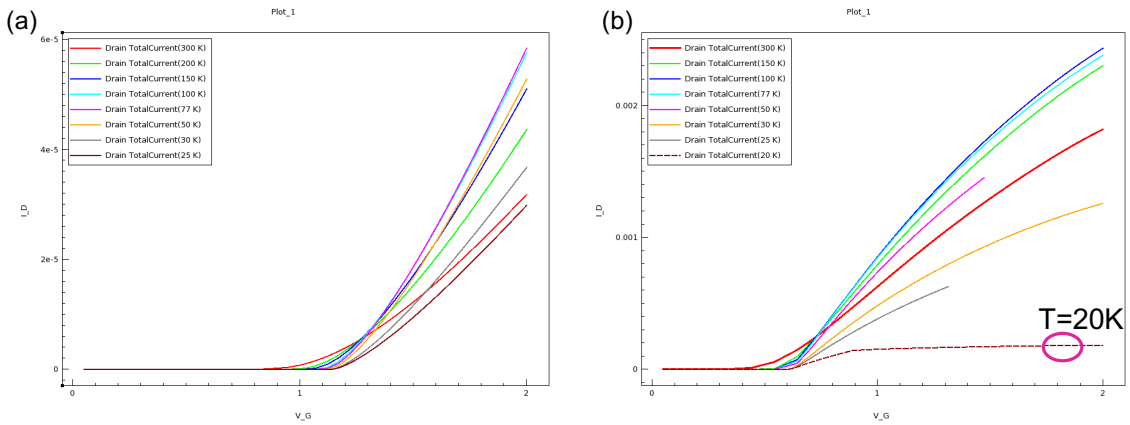


Figure 1.3: Drain current versus gate voltage plotted with (a) a long channel device ($L_{ch} = 1.75\,\mu\text{m}$) where the current at $T = 20\,\text{K}$ fails to converge and (b) a short channel device ($L_{ch} = 65\,\text{nm}$) where current converged to an unreasonable saturation.

To Achieve the Goals

We need a novel and comprehensive modeling framework to bridge this gap to capture quantum transport, inelastic scattering, and strong correlation effects at low temperatures. The current state of the art in the cryogenic simulation of 2D material-based MOSFETs has achieved notable results but is still limited by the lack of a unified and consistent modeling framework. Existing methods[3, 48, 49, 50] overlook critical factors like scattering mechanisms and interface trap states, or rely on empirical parameters and assumptions, leading to a disparity between theoretical predictions and experimental observations. Moreover, other methods[51, 52, 53] are often computationally intensive to achieve convergence and stability. We address these shortcomings by proposing an advanced cryogenic simulator for 2D material-based MOSFETs. Our NEGF-based simulator can include temperature-dependent electrostatics and scattering effects such as carrier freeze-out, interface trap, dissipation, and band tail states.

Our research aims at the numerical simulation down to 50 K serving as the first step in understanding device behavior at the ultra-cold regime of 4 K. The intermediate temperature, while not as extreme, is critical in bridging the gap between room-temperature electronics and superconducting quantum devices, offering a practical platform for studying cryogenic effects and device scalability. The insights gained at these intermediary temperatures provide a crucial foundation for future explorations aimed at reaching the lower temperature thresholds necessary for quantum computing applications. Differences in carrier mobility, quantum interference effects, and phonon scattering at 50 K versus 4 K suggest nuanced behavior as temperatures approach absolute zero. Understanding these variations at 50 K sets the stage for targeted simulations and experimental validations at 4 K, forming a comprehensive approach to device design and optimization in the quantum realm.

1.4.2 Multi-Level Simulation for Nanotransistors

The second segment of my research is geared towards developing a multi-level simulation platform for 2D-material-based MOSFETs and creating a custom compact model. The rationale behind this development is multifaceted. Primarily, this model serves as a critical bridge, linking the intrinsic device simulation with circuit analysis, thereby streamlining the process of understanding and predicting device behavior. The impetus for this endeavor is spurred by the rapid emergence of 2D-based MOSFETs as a key player in the field of advanced semiconductor technology. This segment aims to enrich our understanding and predictive capabilities regarding these devices, contributing significantly to the technological advancements in the semiconductor sector [54].

More than Moore's

The rapid and relentless advance of technology has generated an ever-growing demand for electronic devices that boast swifter speeds, heightened efficiency, and affordability. Consequently, researchers have embarked on an exploration of the untapped potential held by two-dimensional (2D) materials, including molybdenum disulfide (MoS_2), black phosphorus (BP), and germanane (GeH), as potential channel materials for the forthcoming era of complementary metal oxide semiconductor (CMOS) technology [32, 24, 25, 55, 56]. Even in their nascent stages, these materials have showcased remarkable promise due to their exceptional attributes, which encompass a high on/off current ratio, rapid switching capabilities, and resistance to short-channel effects [57, 58, 59, 60, 61]. For instance, early theoretical studies predicted that short-channel GeH Field-Effect Transistors (FETs) held great potential, boasting high field-effect mobility ($\sim 150\,\mathrm{cm^2{\cdot}V^{-1}{\cdot}s^{-1}}$) and large ON/OFF current ratio ($> 10^5$) [55, 62]. Despite strenuous endeavors in fabricating devices founded on 2D materials, a conspicuous gap persists in the realm of systematic design and simulation frameworks, which are indispensable for device engineering and circuit performance optimization. Given the often prohibitive costs associated with experimental exploration, this void assumes heightened significance, particularly in the context of novel nanoelectronics reliant on 2D materials.

Lack of Comprehensive Tools for Simulating 2D-Based MOSFETs

Hence, the imperative arises to develop a multi-level simulation framework that comprehensively encompasses materials, devices, and circuits. Central to this endeavor is the creation of a suitably tailored compact model designed explicitly for 2D-material transistors. Such a model serves as the critical bridge connecting the inherent material properties to the performance characteristics of the devices, ultimately influencing circuit performance. The existing arsenal of transistor compact models [63, 64, 65], primarily rooted in classical drift-diffusion theory [66], proves inadequate when addressing nanotransistors employing 2D materials. While a few compact models have made strides in capturing ballistic transport for silicon devices, exemplified by the MVS model [67], they falter in encapsulating the distinct physics inherent in low-dimensional materials, such as the quantum capacitance effect in 2D-material FETs. Therefore, there is a need for a compact model explicitly tailored to 2D-material Field-Effect Transistors (FETs), comprehensively encapsulating the unique effects and physics inherent in the 2D structure. Although some compact models [68, 69, 70, 71] have made headway in characterizing the features of 2D FETs, a conspicuous void persists in the literature concerning comprehensive multi-level simulations.

Such simulations must span the entire spectrum, encompassing materials, devices, compact models, and circuit analysis for 2D-material-based FETs exhibiting ballistic transport.

To Achieve the Goals

This research aims to bridge this substantial gap, advancing our understanding and application of 2D materials in the realm of cutting-edge electronics.

Chapter 2

Cryogenic CMOS Transistor

The investigation of MOSFET behavior at cryogenic temperatures reveals unique operational characteristics shaped by quantum mechanics and diminished thermal noise. This chapter delivers an in-depth analysis of Cryogenic CMOS Transistors, with a particular focus on the behavior of MOSFETs in cryogenic environments. Structured to encapsulate both the observed cryogenic behavior and the fundamental physics of MOSFETs, the chapter addresses a range of phenomena that manifest at these low temperatures. Importantly, while much of our study concentrates on the 50 K range, the insights gained provide vital clues about the changes that might be expected as temperatures approach the ultra-cold regime of 4 K. By analyzing effects at 50 K, such as shifts in carrier mobility and alterations in electronic scattering processes, we begin to piece together the broader implications of quantum behavior and reduced noise levels on device performance, laying the groundwork for a more comprehensive understanding of MOSFETs as temperatures edge closer to absolute zero.

2.1 Characteristics and behavior of Cryo-MOSFETs

Understanding the behavior of cryogenic MOSFETs for the simulation is essential and practical. One of the key features that degrade the transfer characteristics of MOSFETs at cryogenic temperatures is the subthreshold swing (SS). At room temperature, the SS is limited by the thermal energy of the carriers, which is proportional to the product of the Boltzmann constant (k_B) and the temperature (T). This limit can be written as $(k_B T/q \ln 10)$, where q is the elementary charge. However, the SS deviates from this limit at cryogenic temperatures and reaches a higher saturation value. This is due to the variation

of the bandgap, which results in an exponential band tail decay in the density of states. This means some states are available for the carriers below the nominal band edge. The band tail decay is described by a parameter W_t, which is larger than the thermal energy at cryogenic temperatures. Hence, the SS limit at low temperatures is $(W_t/q)\ln 10$, which is higher than the room-temperature limit. The band tail decay originates from intrinsic mechanisms, such as electron-electron interactions and electron-phonon scattering, which are independent of the device structure and material [53, 72].

2.1.1 Inflection

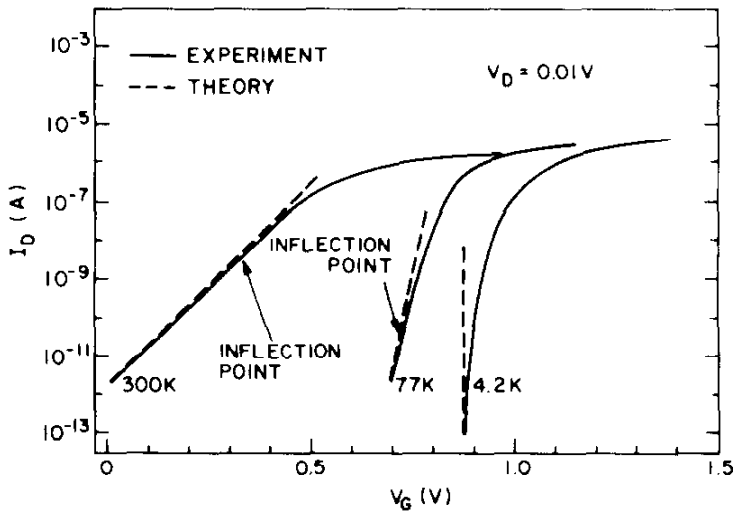


Figure 2.1: Log I_D vs V, traces at three different temperatures [2].

Another feature that affects the transfer characteristics of MOSFETs at cryogenic temperatures is the inward bending (inflection) of the current. This phenomenon occurs when the drain voltage (V_D) is high enough to induce impact ionization in the channel. Impact ionization is a process where a high-energy carrier collides with an atom in the lattice and creates an electron-hole pair. The newly generated carriers can increase the drain current, enhancing the effective transconductance of the device. However, they can also initiate a positive feedback mechanism, where the additional carriers amplify the electric field in the channel, which in turn boosts the impact ionization rate. This can lead to a runaway situation, where the drain current becomes unstable and shows a negative differential resistance. This phenomenon is called the kink effect or snapback, and it can

harm the device or cause circuit malfunction. To prevent the kink effect, the drain voltage must be maintained below a certain threshold, which depends on the temperature and the device parameters. At room temperature, the impact ionization threshold is relatively high, and the current inflection is negligible in the transfer characteristics. However, at cryogenic temperatures, the impact ionization threshold is considerably lower, and the current inflection is more evident. This is because the carrier distribution in the channel is narrower at low temperatures due to the reduced thermal broadening. Consequently, the carriers have higher energies and are more prone to impact ionization [73]. The current inflection constrains the maximum drain voltage that can be applied to cryogenic MOSFETs, and it also impacts the output resistance and the gain of the device.

2.1.2 Carrier Mobility

At cryogenic temperatures, carrier mobility in MOSFETs typically experiences significant alterations due to the pronounced reduction in phonon scattering. This phenomenon generally leads to an increase in mobility, as the thermal vibrations of the lattice, which hinder the free movement of charge carriers, are greatly reduced. However, this improvement in carrier mobility is not uniform across all temperature ranges and may exhibit a complex behavior as temperatures approach the deep cryogenic regime [36].

In the deep cryogenic temperature range, other scattering mechanisms begin to assert their influence more strongly. Impurity scattering and interface roughness scattering [74], in particular, become predominant due to the reduced thermal energy. These mechanisms can limit the mobility increase achieved through reduced phonon scattering. Impurity scattering is related to the interaction of charge carriers with dopant atoms within the semiconductor, while interface roughness scattering arises from the imperfections at the semiconductor-oxide interface. Both types of scattering are less temperature-dependent than phonon scattering but become more relevant as the latter is minimized in cryogenic conditions. The complex interplay between these scattering mechanisms necessitates a comprehensive investigation that extends beyond the capabilities of standard modeling approaches, which are typically calibrated for higher-temperature operations. Traditional compact models [67, 65] often fail to accurately predict the behavior of MOSFETs at cryogenic temperatures, as they do not incorporate the detailed physics of low-temperature phenomena. For instance, the industry-standard compact models are not readily available for deep cryogenic temperatures ($< 50\,\mathrm{K}$), indicating a gap in the current understanding and modeling of MOSFET behavior in this regime.

Moreover, the transition to cryogenic temperatures also highlights the need to revisit the assumptions made in these models. For example, the Maxwell-Boltzmann approximation

of Fermi-Dirac statistics, commonly used in the modeling of semiconductor devices [75], remains valid for non-degenerate doping concentrations at deep-cryogenic temperatures. This is crucial for the development of accurate electrostatic models of MOSFETs operating in these conditions. As such, developing a physical model that accurately captures the essence of carrier mobility at cryogenic temperatures is paramount. This model would need to consider the reduced phonon scattering, the increasing significance of impurity and interface roughness scattering, and the validity of the Maxwell-Boltzmann approximation in the cryogenic regime. Such a model would provide a more reliable basis for understanding the operation of MOSFETs at cryogenic temperatures and for the development of compact models for circuit simulation and design purposes.

2.1.3 Threshold Voltage Shifts

Threshold voltage (V_{TH}) shifts at cryogenic temperatures are a complex interplay of several factors that significantly influence the performance and operational characteristics of MOSFET devices. These shifts can be primarily attributed to changes in the semiconductor's dielectric properties, dopant freeze-out effects, and alterations in the surface potential. At cryogenic temperatures, the semiconductor's dielectric constant may change, affecting the capacitance of the MOSFET and hence its threshold voltage. The reduction in temperature leads to a decrease in the dielectric's thermal vibrations, which can alter its permittivity. This change in permittivity directly impacts the threshold voltage, as V_{TH} is inversely proportional to the capacitance between the gate and the channel[41].

Furthermore, the dopant freeze-out effect becomes significant at these low temperatures [3]. As the temperature decreases, the ionization energy of the dopants exceeds the thermal energy available, leading to a decrease in the number of free charge carriers. This reduction in carrier concentration effectively increases the device's threshold voltage, as a higher gate voltage is now required to invert the semiconductor's surface and induce a conducting channel. The dopant freeze-out phenomenon is particularly noticeable in non-degenerately doped regions of the MOSFET, where the concentration of dopants is insufficient to maintain a significant ionization level at cryogenic temperatures.

In addition to these factors, alterations in the surface potential due to changes in the semiconductor's work function at low temperatures also contribute to the threshold voltage shift. The work function, which is the energy required to remove an electron from the solid to a point just outside the solid surface, can vary with temperature, leading to adjustments in the surface potential and, consequently, the threshold voltage. Despite these complexities, the Maxwell-Boltzmann approximation of Fermi-Dirac statistics remains

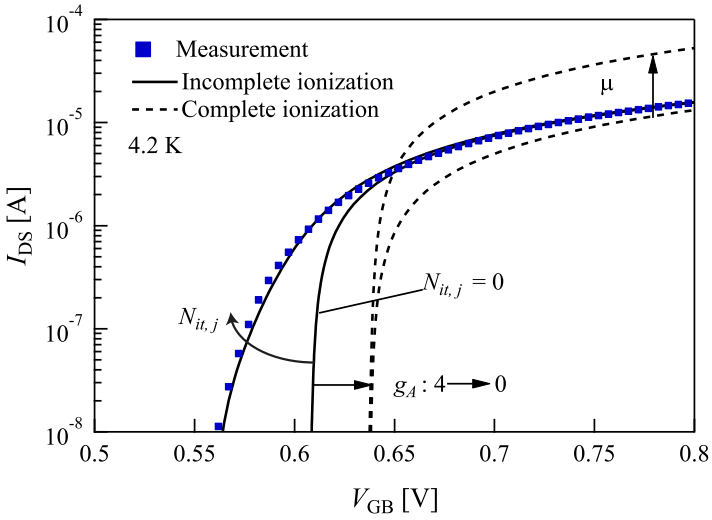


Figure 2.2: Influence on current at 4.2 K due to incomplete ionization ($g_A = 4$), resulting in reduced threshold voltage [3].

valid for non-degenerate doping concentrations at cryogenic temperatures, simplifying the analysis of these effects. This approximation allows for the use of simpler expressions to describe carrier concentrations and electrostatics within the MOSFET, facilitating the derivation of models that can accurately predict threshold voltage behavior under cryogenic conditions. A comprehensive understanding of these factors is crucial for accurately modeling and designing MOSFETs intended for cryogenic applications. It highlights the need for specialized models that can account for the unique physical phenomena occurring at these low temperatures, ensuring the reliable operation of devices in cryogenic environments.

2.1.4 Reduced Thermal Energy on Electron Transport

The significant reduction in thermal energy at cryogenic temperatures profoundly influences electron transport mechanisms within MOSFET devices. Traditional drift-diffusion models, which adequately describe electron transport at higher temperatures, become less effective as quantum transport mechanisms emerge prominently. One such mechanism is band-tail tunneling, which involves electron movement through localized states in the semiconductor bandgap. This phenomenon becomes increasingly dominant in the weak inversion regime as temperatures drop to deep-cryogenic levels. At temperatures below 50 K, the subthreshold slope (SS) experiences saturation, failing to achieve its theoretically predicted minimum

value based on classical diffusion current theories. This deviation from the expected thermal limit highlights a critical shortfall in conventional semiconductor device models, which do not account for the quantum mechanical effects that become significant at these low temperatures. The presence of a hopping current, characterized by electrons tunneling through band-tail states within the bandgap, further complicates the electron transport behavior. This quantum effect leads to interface-state densities that remain within reasonable limits even at temperatures approaching absolute zero, challenging the conventional understanding and modeling of MOSFET operation under such extreme conditions.

The inadequacy of standard compact models to accurately predict these phenomena necessitates a paradigm shift in the modeling approach for MOSFETs designed to operate at cryogenic temperatures. New models must embrace the quantum mechanical realities of electron transport in this regime, incorporating mechanisms like band-tail tunneling and the effects of reduced thermal energy on carrier mobility and threshold voltage. Developing such models requires a deep understanding of the interplay between quantum effects and traditional semiconductor physics, ensuring that the unique characteristics of MOSFETs at cryogenic temperatures are captured with high fidelity. This complex interplay between reduced thermal energy and electron transport mechanisms at cryogenic temperatures underscores the need for an enhanced theoretical framework. Such a framework would not only facilitate a more accurate prediction of device performance under these conditions but also drive the innovation of cryogenic MOSFET technologies for applications in quantum computing and ultra-low-power electronics, where precise control and understanding of quantum effects are paramount.

2.2 Device Physics of Cryo-MOSFETs

The second section shifts focus to the underlying physical principles governing MOSFET behavior at cryogenic temperatures. It begins with a discussion on the Fermi Distribution Function, highlighting how its characteristics are altered in a cryogenic environment and the implications of these changes on MOSFET operation. The section then explores the concept of Incomplete Ionization, examining how reduced temperatures affect the ionization of dopants in the semiconductor material. This is followed by a study of Interface Trapping, where the traps at the semiconductor-oxide interface play a crucial role in defining the transistor's characteristics under cryogenic conditions. Finally, the section addresses the Band Tail phenomenon, explaining how the tail states in the energy band structure of semiconductors influence MOSFET behavior at low temperatures.

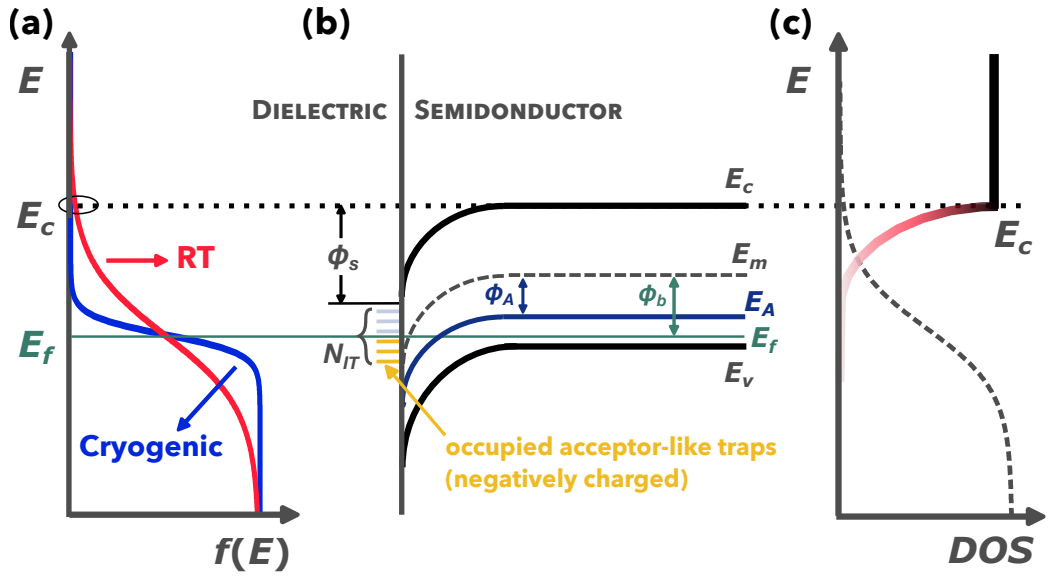


Figure 2.3: Band diagrams depicting distinct cryogenic models with conduction band edge (E_C) and Fermi energy level (E_f) provided as reference in (a) the Fermi-Dirac distribution of room temperature and cryogenic temperature, (b) a comparison of ionization energy (Φ_A) with thermal energy (Φ_b) and visualization of interface trap states with trap density (N_{IT}), and (c) the impact on the Density of States (DOS) below the conduction band edge (E_c) resulting from band tail states.

2.2.1 Fermi Distribution Function

The influence of cryogenic temperatures on semiconductor physics, particularly through the lens of the Fermi distribution function, heralds significant implications for MOSFET device characteristics. The Fermi distribution function, given by

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}}, \tag{2.1}$$

sharply delineates occupied and unoccupied energy states in a semiconductor. This delineation becomes more pronounced at cryogenic temperatures due to the significant reduction in thermal energy, kT, leading to a step-like behavior of the function. This section explores the ramifications of this behavior on carrier concentration, mobility, and the MOSFET's subthreshold characteristics.

The sharper Fermi function at cryogenic temperatures drastically influences carrier concentration within the semiconductor. Electrons below the Fermi level (E_F) are nearly completely in occupied states, while those above are largely in unoccupied states. This clear demarcation reduces the thermal generation of electron-hole pairs, thus affecting intrinsic carrier concentration. In doped semiconductors, the dominant mechanism is dopant ionization, which may be influenced by incomplete ionization at such low temperatures. Moreover, the reduced thermal agitation at cryogenic temperatures leads to diminished phonon scattering, thereby enhancing carrier mobility. This enhancement directly contributes to an increase in the MOSFET's on-current (I_D) for a given gate voltage (V_G), improving device performance.

The impact of cryogenic temperatures extends to the MOSFET's sensitivity to gate voltage, particularly in the subthreshold region where the device transitions from OFF to on states. The sharpened Fermi distribution facilitates a more efficient modulation of carrier concentration in the channel by the gate voltage, potentially leading to improved subthreshold swing (SS) characteristics. A steeper SS represents a more efficient switching behavior, requiring smaller voltage changes to increase the current by an order of magnitude. However, it is crucial to consider the effects of band tails and localized states at the semiconductor-oxide interface, which can alter these idealized behaviors. The inflection phenomenon in MOSFET transfer characteristics at cryogenic temperatures, attributed to localized states with a Gaussian distribution, reflects the complex interplay between the Fermi distribution and the semiconductor's energy landscape, affecting electrical characteristics.

Cryogenic temperatures offer enhanced carrier mobility and gate voltage sensitivity, which benefit high-performance MOSFET applications. Nonetheless, localized states and the specifics of cryogenic device physics, including incomplete ionization and interface

traps, necessitate a nuanced understanding of the accurate prediction and optimization of MOSFET performance in cryogenic conditions. This area remains an active field of research, focusing on leveraging the advantages while addressing the challenges presented by operating at cryogenic temperatures.

2.2.2 Incomplete Ionization

Incomplete ionization, a critical phenomenon in semiconductor physics, particularly impacts device behavior at cryogenic temperatures. This phenomenon arises when the thermal energy available is insufficient to ionize all dopant atoms within the semiconductor lattice, resulting in a lower effective doping concentration. The ionization energy required for dopants to contribute free carriers becomes a pivotal factor at reduced temperatures, decreasing the number of ionized dopant atoms. This reduction significantly influences the threshold voltage (V_T) of MOSFETs due to changes in the built-in potential of the semiconductor.

At room temperature, the thermal energy is generally sufficient to ionize most dopant atoms. However, as the temperature decreases, the thermal energy (kT) falls below the ionization energy of the dopants, leading to incomplete ionization. The expression for the carrier concentration considering incomplete ionization can be modified as follows:

$$p = N_A - n_i \exp\left(\frac{E_F - E_A}{kT}\right), \tag{2.2}$$

where N_A is the acceptor concentration, n_i is the intrinsic carrier concentration, E_F is the Fermi level, E_A is the acceptor energy level, and kT is the thermal energy.

The effect of incomplete ionization is more pronounced at cryogenic temperatures, where the dopant atoms remain largely un-ionized. This scenario leads to a significant deviation from the expected behavior based on complete ionization models. The threshold voltage of a MOSFET, which is intricately linked to the doping concentration and the built-in potential, is therefore affected. A lower effective doping concentration results in an altered threshold voltage, necessitating adjustments in device design and operation parameters to maintain performance.

Engineering strategies to mitigate the impact of incomplete ionization include selecting dopants with lower ionization energies or increasing the overall doping concentration. These approaches aim to ensure a higher proportion of dopant atoms contribute free carriers even at reduced temperatures, thus maintaining the device's operational integrity. The

relationship between the threshold voltage and the doping concentration, considering incomplete ionization, is given by:

$$V_T = V_{T0} + \frac{kT}{q} \ln \left(\frac{N_A - n_i \exp\left(\frac{E_F - E_A}{kT}\right)}{n_i} \right), \tag{2.3}$$

where V_{T0} is the threshold voltage at a reference temperature, and q is the charge of an electron.

Incorporating the effects of incomplete ionization into device models is crucial for accurate prediction and optimization of MOSFET performance at cryogenic temperatures. It allows for a deeper understanding of the fundamental limitations and operational capabilities of semiconductors under such extreme conditions, guiding the development of cryogenic electronics for applications in quantum computing, deep space exploration, and low-noise amplifiers.

2.2.3 Interface Trapped States

Interface trapping, particularly at the semiconductor-oxide boundary, plays a crucial role in defining the operational characteristics of MOSFETs, more so under cryogenic conditions. These traps, capable of capturing or emitting carriers, directly influence the surface potential and, thus, the threshold voltage (V_T) of the device. The dynamics of these traps, including their capture and emission rates, undergo significant changes at low temperatures, which can markedly affect the device's electrical response. The altered trap dynamics can lead to slower carrier release rates, potentially resulting in a more pronounced impact on device characteristics due to the prolonged presence of carriers within these traps.

The charge density per unit area trapped in interface states, Q_{it}, is described by the integral

$$Q_{it} = -q \int_{E_i}^{+\infty} D_{it}(E) f(E) dE, \tag{2.4}$$

where $D_{it}(E)$ represents the density of interface traps across the energy spectrum, and $f(E)$ is the occupancy probability of these traps, typically modeled by the Fermi-Dirac distribution at equilibrium. The interface trap charge density, Q_{it}, can be decomposed into two parts: $Q_{it} = Q_u + Q_0$, where Q_u is associated with a uniform distribution of traps across the bandgap, and Q_0 models the exponentially increasing trap density near the band edges, which becomes significant at cryogenic temperatures.

The energy-dependent trap density $D_{it}(E)$ can exhibit a Gaussian distribution around the band edge, characterized by an exponential decay into the bandgap. This distribution is mathematically expressed as

$$D_0(E) = \frac{N_0}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(E - E_c)^2}{2\sigma^2}\right), \tag{2.5}$$

where N_0 is the peak density of traps, E_c is the conduction band edge, and σ represents the standard deviation of the distribution.

The presence of interface traps affects the threshold voltage of MOSFETs through the modification of the surface potential. The alteration in carrier dynamics due to trapping and detrapment processes at the interface leads to a variation in the effective channel charge, thereby influencing V_T. The threshold voltage, incorporating the effects of interface traps, can be adjusted to reflect these changes as follows:

$$V_T = V_{T0} + \frac{q(Q_u + Q_0)}{C_{ox}}, \tag{2.6}$$

where V_{T0} is the threshold voltage without considering the interface traps, C_{ox} is the oxide capacitance, and q is the elementary charge.

At cryogenic temperatures, the altered trap dynamics significantly impact MOSFET performance, necessitating high-quality interface engineering during device fabrication. The slower dynamics of trap capture and emission processes can lead to enhanced sensitivity of V_T to these traps, affecting device stability and performance. Optimizing the interface quality and selecting materials with minimized trap densities become paramount in ensuring the reliability of cryo-MOSFETs.

The profound impact of interface trapping on MOSFET behavior, especially at cryogenic temperatures, underscores the importance of understanding and controlling these traps. The quantitative treatment of interface traps, through models incorporating their energy-dependent distributions, provides insights into their effects on device parameters such as the threshold voltage. As the field of cryogenic electronics continues to grow, particularly for applications in quantum computing and space exploration, the role of interface traps and strategies for their mitigation will remain areas of active research and development.

2.2.4 Band Tail Effect

The band tail phenomenon significantly affects the operational characteristics of MOSFETs, particularly in cryogenic environments. This phenomenon is characterized by an extended

distribution of electronic states into the bandgap, arising from various forms of disorder within the semiconductor lattice, such as crystalline disorder, strain, residual impurities, and surface roughness. These extended states, or tail states, can trap carriers, thus impacting mobility and recombination rates, which in turn influence the device's subthreshold behavior and OFF-state current.

The band tail phenomenon can be attributed to the potential fluctuations within the semiconductor material due to the aforementioned forms of the disorder. These fluctuations lead to localized states within the bandgap, which are energetically closer to the conduction or valence bands than the idealized sharp band edges. Mathematically, the density of these states can be described by an exponential or Gaussian distribution, extending from the band edges into the bandgap. The exponential band tail model is often expressed as:

$$D(E) = D_0 \exp\left(-\frac{|E - E_c|}{W_t}\right), \tag{2.7}$$

where $D(E)$ is the density of states at energy E, D_0 is a constant, E_c is the conduction band edge energy, and W_t is the characteristic decay length of the tail states into the bandgap.

At cryogenic temperatures, the occupancy of band tail states changes due to the reduced thermal energy available to carriers. This altered occupancy affects carrier mobility by increasing the probability of carriers being trapped in these states, thereby reducing effective mobility. Similarly, recombination rates can be affected due to the increased likelihood of carriers being captured by states within the band tail, altering device performance parameters such as subthreshold swing (SS) and OFF-state current (I_{OFF}).

The presence of band tail states at cryogenic temperatures leads to a degradation in the subthreshold swing (SS) of MOSFETs. The SS is a critical parameter that describes how efficiently the device transitions from OFF to on state with a change in gate voltage. An ideal SS is limited by the thermal voltage (kT/q), but the trapping and slow release of carriers in band tail states increase the required gate voltage change to achieve the same current increase, thus degrading the SS. The OFF-state current (I_{OFF}) can also increase due to the enhanced carrier trapping in the tail states, reducing the device's power efficiency.

The band tail phenomenon plays a significant role in determining the behavior of MOSFETs at cryogenic temperatures. Understanding and accurately modeling the effect of band tail states are crucial for predicting device performance in cryogenic environments. This understanding informs the design and optimization of low-power electronic systems intended for operation in such conditions, highlighting the importance of material quality and fabrication techniques in minimizing the impact of band tails on device performance.

2.3 Summary

This chapter has provided an in-depth exploration of the significant alterations in MOSFET behavior and underlying physical mechanisms when operated in cryogenic environments. Through a detailed examination of phenomena such as the Fermi Distribution Function, Incomplete Ionization, Interface Trapping, and the Band Tail effect, we've shed light on the complex dynamics that characterize MOSFET performance at extremely low temperatures. The insights garnered from this analysis are invaluable for both researchers and practitioners in the burgeoning field of cryogenic electronics, offering guidance on navigating the challenges inherent to these conditions.

By elucidating the intricate interplay between decreased temperature and semiconductor physics, this discussion underscores the critical aspects that must be considered for device optimization in cryogenic applications. The nuanced understanding of each phenomenon—ranging from the sharpened Fermi distribution leading to enhanced mobility and device sensitivity, the complications posed by incomplete ionization on doping efficiency, to the pivotal roles of interface trapping and band tail states in carrier dynamics—provides a comprehensive framework for addressing the operational challenges faced by MOSFETs under cryogenic conditions.

Moreover, this chapter emphasizes the importance of innovative engineering strategies and advanced material science in mitigating adverse effects and enhancing device reliability and performance. As the demand for cryogenic electronics continues to grow, driven by applications in quantum computing, space exploration, and ultra-low noise amplifiers, the foundational knowledge presented here serves as a cornerstone for future advancements. It propels the development of more efficient, robust, and reliable systems that can withstand the rigors of cryogenic operation, marking a significant step forward in the quest to unlock the full potential of electronics at the frontiers of temperature.

Chapter 3

Modeling and Simulation Method

In this chapter, we conduct a detailed exploration of sophisticated modeling and simulation strategies, specifically targeting electronic transport properties. The focus is primarily on the Non-Equilibrium Green's Function (NEGF) formalism enhanced with the Büttiker probe method. The structure of this chapter is meticulously crafted to impart a holistic understanding of these advanced simulation tools.

3.1 NEGF-Poisson Formalism

The Non-Equilibrium Green's Function (NEGF) coupled with the Poisson equation is referred to as the NEGF-Poisson method. In the first section of the chapter, we introduce the NEGF-Poisson solver, a powerful tool for simulating quantum transport in nano-electronic devices, and how it is adeptly utilized to characterize the non-equilibrium behavior of electrons and holes in cryogenic simulations. This method synergistically combines the NEGF, primarily for transport computations, with the Poisson equation to manage electrostatic aspects. Recognized for its high efficiency in modeling nanoscale electronic devices, the NEGF-Poisson approach provides a schematic overview in Figure 3.1. The NEGF component employs Green's functions to account for quantum mechanical effects meticulously. This inclusion is paramount for a granular analysis of charge transport phenomena at the nanoscale. The approach rigorously considers the quantum mechanical nature of carrier transport, offering a detailed perspective on the behavior of charge carriers. Complementing the NEGF, the Poisson equation is seamlessly integrated to handle the electrostatics within the device. This amalgamation forms a robust framework for simulating

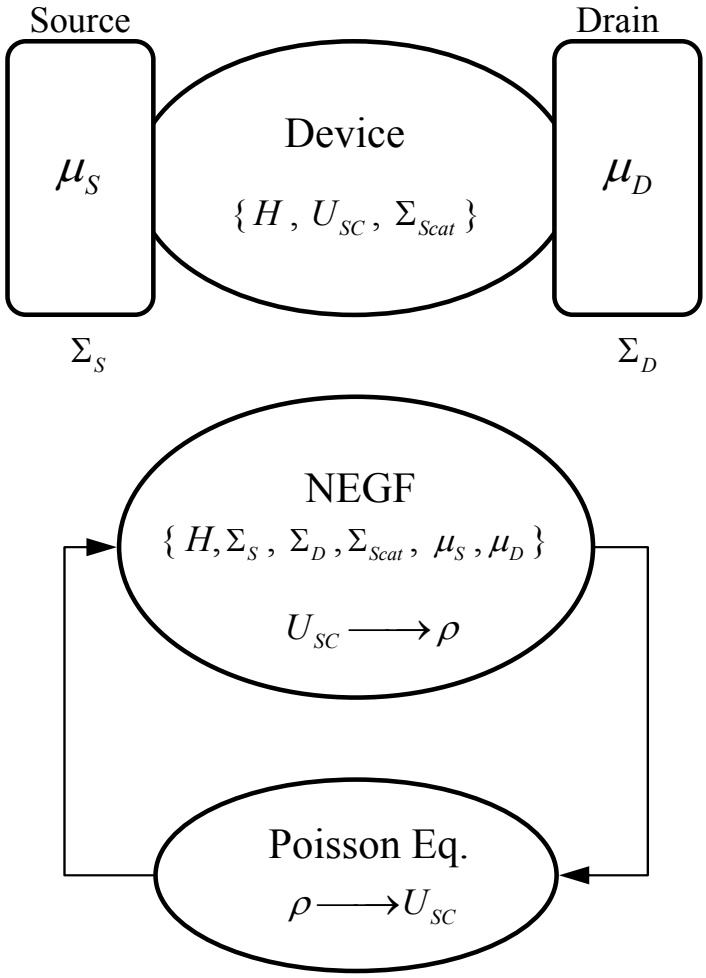


Figure 3.1: (Top) Schematic structure of device coupled to the source and drain contacts. (Bottom) Self-consistent procedure for the coupled NEGF–Poisson computational simulation framework [4].

the complex interplay between quantum mechanics and electrostatics, yielding an intricate representation of device behavior, particularly pivotal for nanoscale electronics.

Within the framework of the Non-Equilibrium Green's Function (NEGF) formalism, a device is typically segmented into three main regions: the source, drain, and channel. The source and drain are considered as infinite reservoirs, interacting with the channel through self-energy matrices, Σ_S and Σ_D, respectively. The Fermi levels of the source (μ_S) and drain (μ_D) are modulated by the applied bias voltage, influencing carrier distribution across the device. The channel region itself is characterized by its Hamiltonian, H. Incoherent transport phenomena, such as scattering processes, are encapsulated by the self-energy term Σ_{incoh}. To facilitate numerical treatment, operators are discretized using either the effective mass (EM) or the tight-binding (TB) approximation. A mode-space representation is often preferred over a full three-dimensional real-space simulation to diminish the computational burden by reducing the dimensionality of the involved operator matrices.

The self-consistent loop for calculating the transmission coefficient at a specific energy ($T(E)$), the electron density (ρ), and the current (I) alongside Poisson's equation involves several steps:

1. Employ the Finite Difference Method (FDM) to discretize all operators for the device in question.
2. Initiate the self-consistent loop with an initial guess for the potential, $U_{\text{SC}}^{(1)}(\vec{r})$.
3. Utilizing $U_{\text{SC}}^{(1)}(\vec{r})$, the device Hamiltonian, H, is defined in line with the chosen approximation method (EM or TB). The self-energies Σ_S and Σ_D are also computed at this stage.
4. Calculate the density matrix through the retarded Green's function.
5. Derive the charge density (ρ) of the device from the density matrix, which is then used to solve for a new self-consistent potential $U_{\text{SC}}^{(1)}(\vec{r})$ via Poisson's equation.
6. Repeat steps (3) to (5) iteratively until convergence is achieved based on predefined criteria.
7. Finally, calculate the current using the converged self-consistent potential and density matrix.

This iterative process ensures the self-consistency of the potential within the device, enabling accurate modeling of carrier transport phenomena.

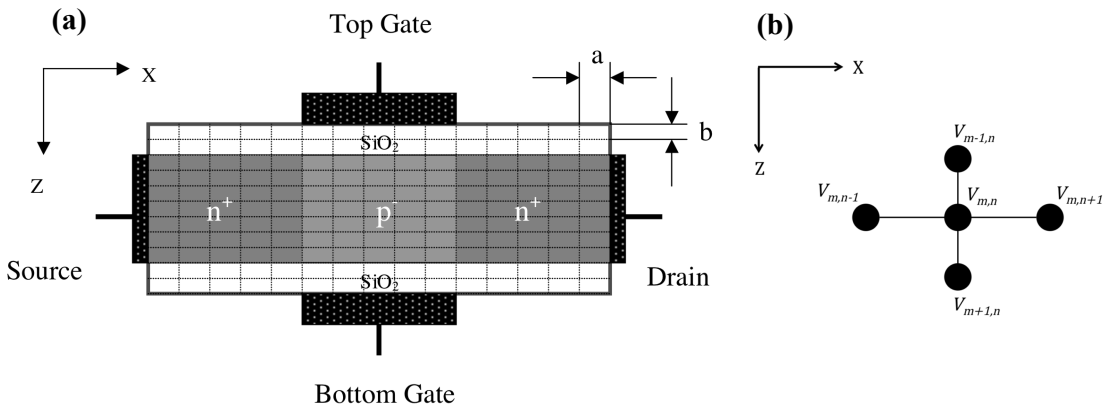


Figure 3.2: (a) A structure diagram of ultra-thin body double-gate MOSFET, within a two-dimensional simulation domain outlined by a solid line. The domain utilizes uniformly spaced grids in both the x and z directions, with the spatial constants denoted as a and b, respectively. This grid arrangement facilitates detailed electrostatic analysis within the specified simulation area, ensuring comprehensive coverage and precision in modeling the device's behavior [5]. (b) The computational model adopts a 5-point star configuration, denoted as a 'computational molecule,' to effectively address the electrostatic calculations required for the device's analysis [6].

3.1.1 Electrostatic Potential Calculation

This section succinctly outlines the methodology for obtaining numerical solutions to Poisson's Equation, employing the Finite Difference Method (FDM) tailored to analyze double-gate MOSFET architectures. Figure 3.2(a) illustrates the device structure, featuring schematically represented, uniformly spaced grids for simulation. The designation of these grids is detailed in Figure 3.2(b), providing a foundation for the ensuing computational analysis. This approach facilitates a structured exploration of the electrostatic properties inherent in MOSFET devices, leveraging grid-based numerical techniques for precision and clarity.

Initially, we posit that the semiconductor and the insulating layer share an identical dielectric constant. The numerical solution for Poisson's equation is obtained using Gauss's law:

$$\nabla \cdot \mathbf{D}(\mathbf{r}) = \rho(\mathbf{r}) \tag{3.1}$$

where $\mathbf{D}$, $\mathbf{r}$, and ρ represent the electric flux density, position vector, and charge density,

respectively. Under the assumption of a uniformly distributed dielectric material and employing the constitutive relation $\mathbf{D}(\mathbf{r}) = \epsilon_0 \epsilon_r \mathbf{E}(\mathbf{r})$, where ϵ_0 is the vacuum permittivity and ϵ_r is the dielectric constant of the semiconductor and insulating layer, we arrive at:

$$\nabla^2 V(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0 \epsilon_r} \tag{3.2}$$

The mesh spacings along X and Z directions are denoted by a and b, respectively, with $V_{m,n}$ representing the potential at grid $[m, n]$. The solution domain encompasses N_X nodes along the X direction and N_Z nodes along the Z direction, requiring an $N_X \times N_Z$ grid for the 2D numerical solution to Poisson's equation. At each node, we rewrite the equation Equation (3.2) as:

$$\frac{\partial^2 V_{m,n}}{\partial x^2} + \frac{\partial^2 V_{m,n}}{\partial z^2} = -\frac{\rho_{m,n}}{\epsilon_0 \epsilon_r} \tag{3.3}$$

where V is the electrostatic potential, x and z are the spatial coordinates, $\rho_{q,j}$ is the charge density at the node $[m, n]$, and ε denotes the permittivity of the medium.

The Finite Difference Method (FDM) is employed to solve this equation numerically, which approximates the second derivatives using discrete differences. Consequently, Equation (3.3) is transformed into Equation (3.4), enabling numerical solution:

$$\frac{V_{m+1,n} - 2V_{m,n} + V_{m-1,n}}{\Delta x^2} + \frac{V_{m,n+1} - 2V_{m,n} + V_{m,n-1}}{\Delta z^2} = -\frac{\rho_{q,j}}{\varepsilon} \tag{3.4}$$

Here, Δx and Δz represent the distances between adjacent nodes in the x and z directions, respectively. This equation is central to the simulation as it allows for the computation of potential values at each node within the mesh by considering the local charge distribution and the boundary conditions. The potential distribution across the device is obtained through iterative solution techniques, facilitating further analysis of device characteristics such as field distributions and capacitance.

Boundary Condition

In simulations of semiconductor devices, especially in the context of ballistic transport models, accurately specifying boundary conditions is crucial for determining potential distributions. At the gate contacts, the electrostatic potential is predetermined by the applied gate bias, which is expressed through the equation $V_{m,n} = V_G$, where V_G is the gate voltage. This scenario requires the application of Dirichlet boundary conditions, which

directly assign a value to the potential at the boundary based on the gate bias and the metalwork function.

Conversely, a nuanced approach is adopted for the remaining boundaries, particularly at the extent of the simulation domain adjacent to the source and drain. Floating boundary conditions, characterized by the condition $\mathbf{n} \cdot \nabla V = 0$, where $\mathbf{n}$ denotes the surface unit normal vector, are implemented. This condition implies that the electric field normal to the boundary is zero, effectively allowing the potential to adjust dynamically. Such an approach is in stark contrast to the fixed potential conditions often used in scattering-dominated transport simulations, where equilibrium charge neutrality is assumed.

This distinction becomes particularly relevant in ballistic transport regimes, especially under high bias conditions (V_{DS}), where the traditional equilibrium assumptions fail to hold. In these scenarios, the imposition of a zero-field boundary condition at the source and drain ends allows for a more realistic representation of carrier behavior, acknowledging the partial carrier depletion at the source end due to suppressed carrier injection from the drain. The floating potential at these boundaries dynamically aligns with the conditions dictated by the Non-Equilibrium Green's Function (NEGF) formalism, thus ensuring macroscopic charge neutrality without predetermining the potential values, providing a more accurate depiction of ballistic transport phenomena.

3.1.2 Non-Equilibrium Green's Function Method

The NEGF framework is pivotal in quantum transport computations, allowing for the modeling of open quantum systems and their interaction with semi-infinite contacts. Notation from reference [76] is employed, and the methodology aligns with that presented in reference [77]. Consider a quantum device in a non-scattering regime, where the Green's function, indicative of the system's response, is defined as:

$$G = [EI - H]^{-1} \tag{3.5}$$

where EI represents the product of energy and the identity matrix, and H signifies the Hamiltonian of the system. Defining:

$$A' = EI - H \tag{3.6}$$

The relationship between A' and G across various regions, namely the left terminal (L), device central region (D), and right terminal (R), can be articulated as:

$$A'G = I = \begin{bmatrix} A'_{LL} & A'_{LD} & 0 \\ A'_{DL} & A'_{DD} & A'_{DR} \\ 0 & A'_{RD} & A'_{RR} \end{bmatrix} \begin{bmatrix} G_{LL} & G_{LD} & G_{LR} \\ G_{DL} & G_{DD} & G_{DR} \\ G_{RL} & G_{RD} & G_{RR} \end{bmatrix} \tag{3.7}$$

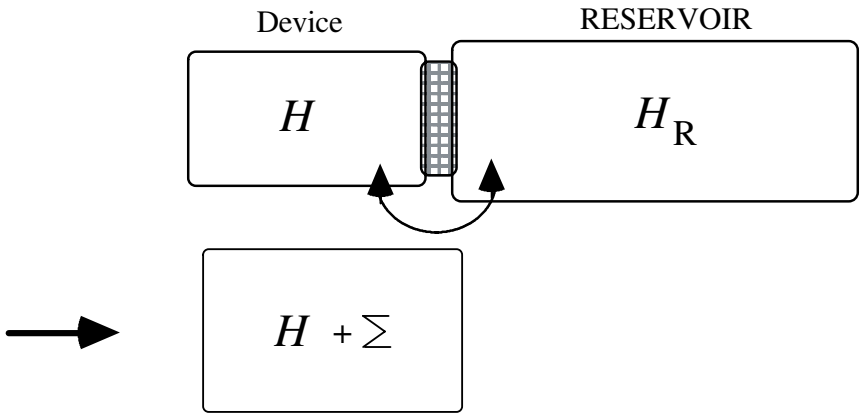


Figure 3.3: Schematic representation illustrating the computational efficiency gained by substituting the extensive Hamiltonian of the reservoir, H_R, with a self-energy term, Σ. This approach simplifies the calculations without losing the physical significance of the reservoir's impact on the device [7].

It follows that:

$$[A'_{DD} - A'_{DL}A'^{-1}_{LL}A'_{LD} - A'_{DR}A'^{-1}_{RR}A'_{RD}]G_{DD} = I \tag{3.8}$$

The matrices A' exhibit a sparse configuration, with non-zero entries predominantly along the diagonal, reflecting the localized nature of the interactions within the system. Subsequently, the Green's function for isolated semi-infinite terminals is defined as:

$$g_L = A'^{-1}_{LL} \quad \text{and} \quad g_R = A'^{-1}_{RR} \tag{3.9}$$

The surface Green's functions $g_{L,11}$ and $g_{R,r1}$, corresponding to the edge layers of the left and right terminals, respectively, are thus:

$$g_{L,11} = (A'^{-1}_{LL})_{11} \quad \text{and} \quad g_{R,r1} = (A'^{-1}_{RR})_{11} \tag{3.10}$$

Self-energy: Ohmic Contact

Incorporating self-energy terms, which capture the essence of semi-infinite contacts, we can reformulate the equation as follows:

$$G(E) = \left[(E + i0^+)I - H - U - \Sigma_S - \Sigma_D\right]^{-1}. \tag{3.11}$$

In this expression, E represents the energy level for which $G(E)$ is evaluated, $i0^+$ denotes an infinitesimally small positive imaginary number, I is the identity matrix, and H symbolizes the Hamiltonian matrix, which was elaborately discussed in section 2.2. U stands for the potential profile across the device, while Σ_S and Σ_D are the self-energy terms pertinent to the source and drain, respectively. The necessity for Σ_S and Σ_D arises from the considerable dimensional disparity between the channel and the electrodes. Therefore, it is computationally more expedient to substitute the extensive Hamiltonian of the reservoir (H_R) with a self-energy term Σ as depicted in Figure 3.3 [77, 76, 7]. For simplification, Σ_S and Σ_D will henceforth be collectively referred to as Σ_{Lead}, computed via:

$$\Sigma_{\text{Lead}}(E) = \tau G_S \tau^\dagger. \tag{3.12}$$

Due to the focus of this thesis being on ohmic contact-based devices, discussions regarding Schottky-type contacts will not be included. For ohmic contacts, a semi-infinite contact approximation is assumed, wherein the contact material is heavily doped yet identical to the channel material. Several methods exist to approximate the surface Green's function g_S for a semi-infinite contact. Among these, the recursive surface Green's function (SGF) approach and the Sancho-Rubio methods are most prevalent.

The recursive SGF method computes g_S iteratively, beginning with an initial guess ($g_S = H_{00}^{-1}$), and utilizes the following update equation:

$$g_S^{new} = \left[(E + i0^+)I - H_{00} - U - H_{01} g_S^{old} H_{01}^\dagger\right]^{-1}, \tag{3.13}$$

where H_{00} represents the Hamiltonian's intra-cell interaction component, and H_{01} accounts for inter-cell interactions. This iterative process continues until convergence is reached, that is, $g_S^{new} \approx g_S^{old}$. The Sancho-Rubio method, outlined in detail in prior works, initiates by defining $\omega = (E + i0^+)I$, $\epsilon_0 = H_{00}$, $\alpha_0 = H_{01}$, and $\beta_0 = \alpha_0^\dagger$. Subsequently, the following recursive relations are solved:

$$\alpha_i = \alpha_{i-1}(\omega - \epsilon_{i-1})^{-1}\alpha_{i-1}, \tag{3.14}$$
$$\beta_i = \beta_{i-1}(\omega - \epsilon_{i-1})^{-1}\beta_{i-1}, \tag{3.15}$$
$$\epsilon_i' = \epsilon_{i-1} + \alpha_{i-1}(\omega - \epsilon_{i-1})^{-1}\beta_{i-1} + \beta_{i-1}(\omega - \epsilon_{i-1})^{-1}\alpha_{i-1}, \tag{3.16}$$

which may be resolved iteratively until α_i and β_i diminish sufficiently. Consequently, one can obtain the surface Green's function as:

$$g_S = (\omega - \epsilon^S)^{-1}. \tag{3.17}$$

The Sancho-Rubio methods typically outperform the recursive SGF approach because they consider a wider scope of layers within the device, as the recursive SGF accounts for only N successive layers, whereas the Sancho-Rubio methods incorporate 2^N layers, thereby enhancing computational efficiency.

Upon determining the self-energy terms, the broadening function $\Gamma(E)$ can be ascertained, encapsulating the rate of charge injection from the leads into the channel:

$$I_{\text{Lead}}(E) = i\left[\Sigma_{\text{Lead}}(E) - \Sigma_{\text{Lead}}(E)^\dagger\right] = -2\,\text{Im}\left[\Sigma_{\text{Lead}}(E)\right]. \tag{3.18}$$

The in-scattering ($\Sigma_\alpha^{\text{in}}$) and out-scattering ($\Sigma_\alpha^{\text{out}}$) self-energy components for electrode α (where $\alpha = \text{S}, \text{D}$) can be computed utilizing Γ. These components are significant as they yield insights into the occupancy states of the contacts:

$$\Sigma_\alpha^{\text{in}}(E) = \Gamma_\alpha(E) f_\alpha, \tag{3.19}$$

$$\Sigma_\alpha^{\text{out}}(E) = \Gamma_\alpha(E)(1 - f_\alpha), \tag{3.20}$$

where f_α denotes the Fermi-Dirac distribution given by:

$$f_\alpha = \frac{1}{1 + e^{(E-E_f)/k_B T}}. \tag{3.21}$$

Here, E_f represents the Fermi level, k_B is the Boltzmann constant, and T is the absolute temperature.

Recursive Green's Function Method

The Recursive Green's Function Method is pivotal for computing the Green's functions in nanoscale devices, where traditional direct methods become computationally prohibitive. This section elaborates on the method's application to calculate specific elements of Green's function matrix, focusing primarily on the left terminal, although the methodology is symmetrically applicable to the right terminal. To initiate, we calculate the left terminal's surface Green's function, $g_{L,11}$, by solving the recursive equation:

$$A_l - T_l^\dagger g_{L,11} T_l g_{L,11} = I, \tag{3.22}$$

where A_l represents the Hamiltonian matrix for the left lead, modified to include the self-energy terms, T_l is the coupling matrix between the device and the left lead, and I is the identity matrix. This equation iteratively computes $g_{L,11}$, starting from the surface

of the terminal and proceeding towards the device region. Once $g_{L,11}$ is obtained, we use Dyson's equation to iteratively calculate the main diagonal elements of the Green's functions associated with the left side:

$$g_{q+1,q+1} = (A_{q+1,q+1} - A_{q+1,q} g_{q,q} A_{q,q+1})^{-1}, \tag{3.23}$$

where the site index $q+1$ ranges from the first lattice point in the device, l_n, to the last, N. For indices from l_n to $l1$, an energy-dependent imaginary potential $i\eta$ is introduced to the diagonal elements of the Hamiltonian within A, accounting for the scattering-induced broadening in the density of states. The density of states, $D(E)$, around the emitter's quasi-bound state in the equilibrium reservoir, influenced by the broadening factor η, is expressed as:

$$D(E) = \frac{\eta(E)/2\pi}{(E-\epsilon)^2 + \eta(E)^2}, \tag{3.24}$$

reflecting the imaginary part of the self-energy and the intrinsic broadening due to reservoir coupling.

For the quantum region, starting from the terminal element $g^{LN}_{N,N}$, which is equivalent to the last diagonal element of $G_{N,N}$, we recursively compute the remaining elements of G using:

$$G_{q+1,q} = -G_{q+1,q+1} A_{q+1,q} g_{q,q}, \tag{3.25}$$
$$G_{q,q+1} = -g_{q,q} A_{q,q+1} G_{q+1,q+1}, \tag{3.26}$$
$$G_{q,q} = g_{q,q} - g_{q,q} A_{q,q+1} G_{q+1,q}. \tag{3.27}$$

For computing the charge in the equilibrium reservoir region, these relations enable the deduction of G''s diagonal elements. The non-equilibrium region's correlation function, G^n, is determined by:

$$G^n = G(\Gamma_1 f_1 + \Gamma_2 f_2) G^\dagger, \tag{3.28}$$

indicating that to calculate G^n, only the first and N-th column of G are necessary. This approach efficiently captures the quantum mechanical effects essential for nanoscale device modeling.

Transmission and Current

By resolving Green's function $G(E)$, one can compute the electron G^n and hole G^p spectral functions, which are pivotal for determining charge occupancy within the energy-space

of the device at every point. These spectral functions are derived from $G(E)$ and the in-scattering ($\Sigma_\alpha^{\text{in}}$) and out-scattering ($\Sigma_\alpha^{\text{out}}$) self-energy terms:

$$G^n(E) = G(E)\Sigma_\alpha^{\text{in}}(E)G^\dagger(E), \tag{3.29}$$

$$G^p(E) = G(E)\Sigma_\alpha^{\text{out}}(E)G^\dagger(E). \tag{3.30}$$

The local density of states (LDOS), denoted as $LDOS(E)$, and spectral function $A(E)$ can be approximated by:

$$LDOS(E) = A(E) = i\left[G(E) - G^\dagger(E)\right] = -2\,\text{Im}\left[G(E)\right] = G^n(E) + G^p(E). \tag{3.31}$$

The charge densities for electrons n and holes p are integral to electronic devices and are obtained by integrating G^n and G^p over all energies:

$$n(\mathbf{r}) = \int \frac{dE}{2\pi} G^n(E), \tag{3.32}$$

$$p(\mathbf{r}) = \int \frac{dE}{2\pi} G^p(E). \tag{3.33}$$

Subsequently, the current density J_c can be calculated as:

$$J_c = \frac{2e}{h}\int \frac{dE}{2\pi}\left[\Sigma_\alpha^{\text{in}}(E)A(E) - \Sigma_\alpha^{\text{out}}(E)A(E)\right]. \tag{3.34}$$

Notably, the Non-Equilibrium Green's Function (NEGF) method described herein constitutes only a portion of the complete simulation. For a comprehensive device simulation, NEGF must be integrated self-consistently with Poisson's equation, which determines the charge density ρ and the resulting electrostatic potential profile U.

3.2 Electrostatic at Cryogenic Temperature

In the cryogenic regime, the standard NEGF-Poisson simulation framework proves inadequate for capturing the complete physical behavior of nanoelectronic devices. To address this limitation, we have developed an advanced simulation framework that extends the original NEGF-Poisson methodology, specifically tailored to analyze nanoelectronics under cryogenic temperatures. A schematic representation of this enhanced framework is depicted in Figure 3.4, illustrating its comprehensive approach to model device physics at low temperatures accurately.

3.2.1 Incomplete Ionization

Incomplete ionization refers to the phenomenon where a fraction of dopant impurities in the semiconductor material fails to donate or accept charge carriers due to various physical and temperature-dependent factors. This phenomenon becomes particularly pronounced in cryogenic conditions, where thermal energy is significantly reduced. Under thermal equilibrium, the majority carrier concentration may deviate from the implanted doping value N_0, owing to the incomplete ionization of dopants. In cryogenic equilibrium, the effect of incomplete ionization is pronounced and referred to as freeze-out, as the thermal dopant-ionization remains minimal[3]. However, during MOS operation, field-assisted ionization becomes a significant factor. Incorporating the model of incomplete ionization into the Poisson module allows us to more realistically account for the distribution of charged and uncharged dopant states, thus providing a more accurate representation of the electrostatics within the device.

Fermi-Dirac statistics offer a fundamental framework for modeling incomplete ionization, including dopant-ionization mechanisms. The concentration of ionized dopants denoted as N_A, is then expressed as the product of the total concentration of implanted dopants and the Fermi-Dirac occupation probability of the acceptor energy E_A (i.e., $N_A = N_0 \times f(E_A)$), specifically,

$$N_A = \frac{N_0}{1 + g_A e^{[(E_A - E_f)/k_B T]}} \tag{3.35}$$

where E_f is the quasi-fermi level, modulated by V_G and V_D, throughout the device as a function of position x. The degeneracy factor (g_A) is set based on the degeneration arising from heavy-light holes and spin-up/down configurations. Assigning g_A a value of zero is equivalent to assuming complete ionization.

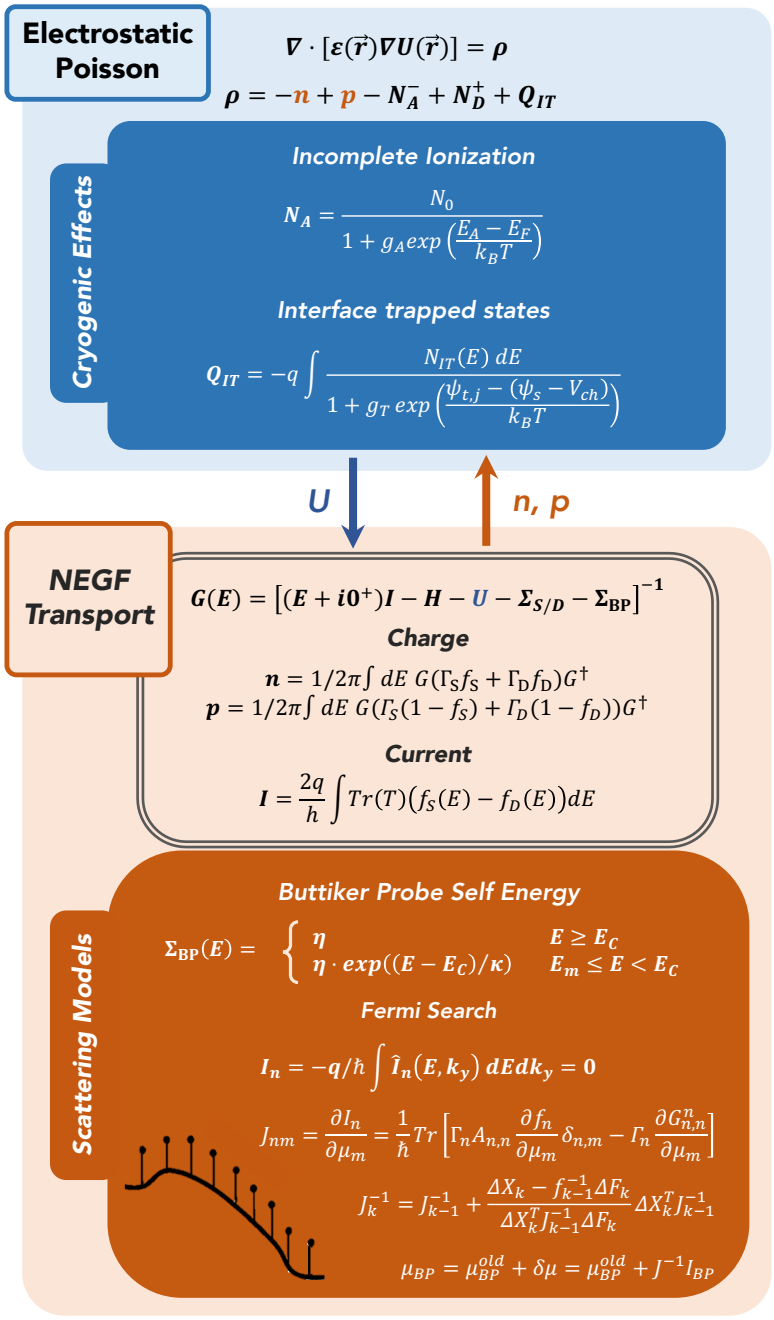


Figure 3.4: Overview of the NEGF-Poisson method incorporating cryogenic models in the Poisson module and the Büttiker probe within the NEGF module.

3.2.2 Interface Traps

On the other hand, interface trap states in semiconductors represent localized energy levels at the interface between the semiconductor and its adjacent dielectric material. These states arise from imperfections, impurities, or defects in the interface, creating discrete energy levels that can capture and release charge carriers. The significance of interface trap states is particularly pronounced in cryogenic temperatures. The reduced thermal energy (k_BT) amplifies the impact of defects and impurities on charge transport dynamics. At lower temperatures, the thermal energy available to carriers diminishes, rendering them more susceptible to being trapped or influenced by these localized states. Understanding and accurately modeling interface trap states are crucial in cryogenic semiconductor device simulations, such as those involving HfS_2 MOSFETs, as they play a pivotal role in determining carrier behavior, device performance, and overall reliability under extreme low-temperature conditions.

$$Q_{IT} = -q \int_{-\infty}^{+\infty} \frac{N(E)}{1 + g_T e^{[(E_T - E_F)/k_B T]}} \mathrm{d}E, \tag{3.36}$$

where the localized states density $N(E)$ is assumed to be the Gaussian distribution function,

$$N(E) = \sqrt{\frac{2}{\pi}} \frac{N_T}{W_T} e^{-2(E-E_c)^2/W_T^2}. \tag{3.37}$$

This Gaussian model is introduced to provide a more accurate physical representation of trap states at the interface. The distribution of localized states is centered around the sharp conduction band edge at the channel/oxide interface (E_c). The numbers of trapped states N_T and the distribution width W_T are two variables for characterizing the interface traps. This characteristic ensures a more accurate depiction of the localized states, accounting for various sources of disorder such as surface roughness, residual impurities, strain, vacancies, and other contributing factors[47, 78].

3.3 Scattering and Band Tail: Büttiker Probe Method

This section transitions to the Büttiker probe approach, pivotal for elucidating scattering mechanisms within quantum transport. We commence with an examination of the Büttiker Probe Self Energy, illustrating its deployment within the NEGF formalism to simulate scattering processes' influence on electronic transport. This discussion extends to the integration of self-energy terms into Green's function to model inelastic scattering events effectively. Subsequently, we delve into the Fermi Search subsection, which is dedicated to the meticulous determination of the Fermi energy level amidst scattering phenomena. This subsection emphasizes the necessity of accurately identifying the Fermi level, a prerequisite for the authentic simulation of transport properties in nano-electronic devices.

Further, we introduce the band tail simulation model within the NEGF framework by incorporating an innovative self-energy term designated to capture electron-environment interactions. The Büttiker probe self-energy, routinely integrated into NEGF for scattering analysis, provides an optimal equilibrium between simulation precision and computational manageability. Within NEGF, electron and hole characteristics are elucidated through recursive solutions of the Green's function.

3.3.1 NEGF with Scattering

In the ballistic transport regime of the Non-Equilibrium Green's Function (NEGF) model, charge carriers are restricted from redistributing their energy and momentum due to the absence of scattering events. To simulate the impact of scattering in the channel, the Büttiker probe method introduces virtual contacts, which are conceptually attached to each atom within the channel. These contacts act as theoretical sites for both elastic and inelastic scattering processes. The scattering events are thereby phenomenologically accounted for within the simulation framework. This approach is graphically represented in Figure 3.5, and both types of scattering events are effectively captured by the Büttiker probe model [79]. The Büttiker probes are virtual contacts introduced to model elastic and inelastic scattering mechanisms [80] efficiently. The retarded Green's function of electrons can be calculated by:

$$G = \left[(E + i0^+)I - H - \Sigma_S - \Sigma_D - \mathbf{\Sigma}_{BP}\right]^{-1} \tag{3.38}$$

Within each partition of the device, all Green's functions and self-energies are represented as matrices with the dependence on in-plane momentum k_y and energy E, which is

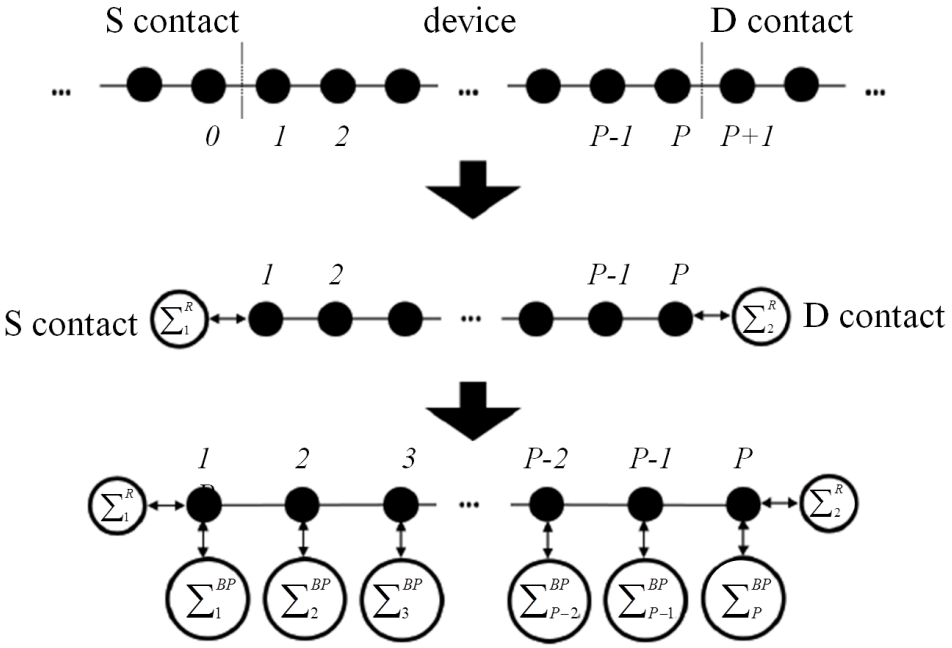


Figure 3.5: Overview of Büttiker Probe Calculations. **Top**: Depicts a quasi-one-dimensional (Quasi-1D) semiconductor device featuring two actual contacts, represented as discretized segments for computational purposes. **Middle**: Illustrates the approach for modeling the semi-infinite contacts by incorporating self-energies at the device terminals, thereby emulating the effect of infinite contact extension. **Bottom**: Demonstrates the integration of Büttiker probe self-energies into the device model to simulate various scattering phenomena. This is achieved by strategically placing a comprehensive array of Büttiker probes along the device, facilitating the emulation of diverse scattering mechanisms within the simulation framework [8].

intentionally excluded for clarity. The Hamiltonian is treated with a tight-binding approach. The self-energies for source and drain contacts are denoted as Σ_S and Σ_D. The inclusion of electron and hole scattering involves Büttiker probe self-energies (Σ_{BP})[81, 79, 82]. Various quantities, such as the density of states, particle density, state occupancy, and current density, can be derived from Green's functions in the NEGF method. The iterative solution for the retarded source and drain self-energies ($\Sigma_{S,D}$) follows the procedure outlined by Datta[76]. It should be noted that Σ_{BP} is purposely incorporated into the calculation of $\Sigma_{S,D}$, to ensure the seamless transitions for electrons and holes from the device to the semi-infinite contact introduced by $\Sigma_{S,D}$.

3.3.2 Büttiker Probe Self Energy

The advanced Büttiker probe consolidates intra-band scattering processes, including interactions with phonons, impurities, and electron-electron scattering, into the empirical scattering parameter η. To consider the band tail of HfS_2 in the simulation, η exponentially decays into the bandgap. The self-energy for electrons in the conduction band associated with the advanced Büttiker probe is expressed as

$$\Sigma_{BP}(x) = \begin{cases} i\eta & \text{if } E \geq E_c(x) \\ i\eta \cdot e^{\frac{E-E_c(x)}{\kappa}} & \text{if } E_c(x) > E > E_m(x) \end{cases} \tag{3.39}$$

where $E_c(x)$ and $E_m(x)$ are the energies of the conduction band bottom and the middle of the band gap at position x. It should be noted that the $\Sigma_{BP}(x)$ is calculated independently for each k_y and E.The nominal band tail parameter κ is chosen to be 100 meV for electrons[83, 84], and is applied throughout the work.

3.3.3 Current Conservation

With Green's functions specified, the pivotal step in the Büttiker probe methodology is to ensure charge conservation within each virtual contact. Unlike tangible contacts where carriers can leave and return, virtual probes influence carrier dynamics by redistributing their energy, momentum, spin, and other properties. It is posited that these probes are sufficiently large to achieve thermal equilibrium hence the Fermi-Dirac statistics dictate the carrier distribution. Benefiting from a defined temperature at each probe, this approach simplifies the model to a single governing parameter—the chemical potential of the probe.

Büttiker probes are integrated into the simulation of devices analogously to how source and drain contacts are applied. The implementation of the Büttiker probe approach

mandates the determination of the probes' chemical potentials such that the total current, described in Eq. 3.40, through each probe is nullified, thereby achieving equilibrium. A study of Büttiker Probe elucidates two principal methodologies for current conservation: the momentum relaxing model and the energy relaxing model. The momentum relaxing model ensures the conservation of current at a particular energy for a specified probe, potentially at the expense of momentum conservation for that probe. Consequently, the formulation of charge current through each virtual contact takes into account these relaxation processes, which are amenable to quantitative analysis by the ensuing equation:

$$I_n = -\frac{q}{\hbar}\int_{-\infty}^{\infty} I_n(E)\mathrm{d}E \tag{3.40}$$

where $I_n(E) = \sum_{m\neq n} T_{m,n}\left[f_n(E) - f_m(E)\right]$ is the current density function of k_y and E at n-th probe. Therefore, each probe's chemical potential μ_n should be adjusted accordingly to satisfy the condition.

The momentum relaxing model stipulates that the current at each probe be nullified, expressed mathematically as

$$I_n(E) = 0 \tag{3.41}$$

This model necessitates the establishment of a probe at each physical and energy coordinate, which is its primary limitation due to the substantial computational demands it imposes, particularly when employing finer energy meshes within the Newton-Raphson algorithm. The scattering phenomena accounted for by this model arise due to the non-conservation of momentum, allowing for the exchange of carriers with disparate momenta, a feature also described as phase-breaking. This constraint does not impede the electrons' energy states, yet it leads to a diminished current due to backscattering.

Conversely, the more inclusive energy relaxation model dictates that the cumulative current be zero across all degrees of freedom except spatial coordinates. This is denoted by the integral

$$I_n = \int_{-\infty}^{\infty} I_n(E)dE = 0 \tag{3.42}$$

The continuity of charge current mandates that the aggregate charge current at each probe I_e^m be zero. This approach diverges from the momentum relaxing model as it permits a singular probe to address all scattering phenomena at a given location by adjusting the chemical potentials to ensure the conservation of energy-integrated current, thus encapsulating the effects of incoherent scattering losses.

Thus far, two distinct paradigms for conserving current have been delineated: the energy relaxing model, which will be the focus for the remainder of our discussions, is acclaimed for its universality.

Newton-Raphson Method vs Broyden's Method

To achieve zero net current flow at each probe, an inner loop (Fermi Search in Figure 3.4) is added in the NEGF module to search for the target of the Fermi levels. We utilize the Newton-Raphson approach[85] for root-finding (Fermi searching). The Jacobian, representing partial first derivatives, is crucial for Newton-Raphson-based methods. The element of Jacobian is defined as follows,

$$J_{nm} = \frac{\partial I_n}{\partial \mu_m} = \frac{1}{\hbar} \mathrm{Tr} \left[\Gamma_n A_{n,n} \frac{\partial f_n}{\partial \mu_m} \delta_{n,m} - \Gamma_n \frac{\partial G^n_{n,n}}{\partial \mu_m} \right] \tag{3.43}$$

Newton's method calculates the Jacobian precisely at each step, eliminating the need for fitting parameters and reducing computational load. The secant method[85] is further applied to speed up the convergence, where the exact Jacobian is computed only in the first iteration (J_0). Subsequent iterations employ Broyden's update formula [41] to update J_k ($k \geq 1$) approximately as shown below,

$$J_k^{-1} = J_{k-1}^{-1} + \frac{\Delta X_k - f_{k-1}^{-1} \Delta F_k}{\Delta X_k^T J_{k-1}^T \Delta F_k} \Delta X_k^T J_{k-1}^{-1} \tag{3.44}$$

where the traditional RGF function is sufficient for updating J. In this paper, most results utilize the secant method, where the exact Jacobian is computed in the first iteration, and Broyden's update formula is applied for subsequent iterations.

3.4 Simulation Robustness

The robustness of numerical simulations is crucial for accurate and reliable modeling of semiconductor devices, especially at cryogenic temperatures where the simulation becomes inherently more unstable, as briefly discussed in Section 1.4.2. This section delves into the reasons behind this instability and outlines mitigation strategies, ensuring more stable and convergent simulations.

First, thermal energy significantly influences the relationship between charge density and Fermi level, represented by the equation $k_B T$, where k_B is the Boltzmann constant and

T is the temperature. As the temperature decreases, k_BT becomes smaller, leading to an increased sensitivity of charge density to fluctuations in the Fermi level. This sensitivity can cause simulation instability. To address this issue and enhance simulation convergence, several techniques are employed:

1. Utilize the initial guess from previously converged results at higher temperatures, where simulations are more stable. This approach involves gradually approaching the target cryogenic temperature, allowing for easier convergence.

2. Refine the energy grid, ensuring a finer resolution that can better capture the subtle changes in the system with temperature variations. This refinement aids in maintaining the stability of the charge density against fluctuations in the Fermi level.

Second, simulations become increasingly sensitive to small energy differences at cryogenic temperatures, making the precise modeling of potential variations along the channel and across the oxide layer critical. To accommodate this sensitivity, the simulation grid mesh is meticulously refined, especially in regions where potential fluctuations are pronounced. For the work presented in this thesis, the grid mesh density was intentionally increased relative to standard room temperature simulations to capture these critical details.

Additionally, integrating cryogenic physics models, particularly scattering models using Büttiker probes and the iterative solution of chemical potentials, adds complexity to numerical simulations. Broyden's method, an approximation of Newton's method, is typically employed to find the distribution of chemical potentials that satisfy the condition of zero net currents. However, Broyden's method may exhibit unstable behavior at extremely low temperatures. In such cases, reverting to the more robust Newton's Method is recommended to ensure stability.

It is important to recognize that the strategies employed to enhance simulation stability—such as increased grid mesh density and finer energy resolution—significantly augment computational intensity. Therefore, a careful trade-off analysis is essential to balance the need for simulation accuracy and computational feasibility. Strategies like adaptive mesh refinement and selective energy grid densification can be explored to optimize computational resources while maintaining the integrity of the simulation results. Addressing the challenges of simulation robustness at cryogenic temperatures requires a multifaceted approach, combining advanced numerical techniques with strategic computational resource management. By carefully implementing these strategies, the reliability and accuracy of simulations for 2D material-based MOSFETs at cryogenic temperatures can be significantly improved, paving the way for deeper insights into device physics and performance optimization in low-temperature environments.

Chapter 4

Simulation of HfS_2 MOSFETs at Cryogenic Temperature

Two-dimensional (2D) materials, notably transition metal dichalcogenides (TMDs) and graphene hold immense promise for revolutionizing the electronics industry, particularly in the context of quantum computing. However, the performance of 2D Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) under cryogenic temperatures poses significant challenges. Addressing this, our research introduces an advanced simulation framework to accurately characterize and predict the behavior of 2D MOSFETs in cryogenic environments. Leveraging the non-equilibrium Green's function (NEGF) formalism, the framework intricately integrates together temperature-dependent electrostatic, scattering, and dissipation effects, along with carrier freeze-out and band tail phenomena, incorporating a Büttiker probe for scattering. Our simulations, rigorously aligned with experimental measurements through detailed mobility calibrations, prove effective in elucidating the behavior of MOSFETs at low temperatures. This alignment not only validates our model but also provides deep insights into the physical processes governing 2D materials in cryogenic settings. By offering a sophisticated tool for designing and optimizing cryogenic devices based on 2D materials, our work significantly impacts the development of electronics for quantum computing, marking a substantial advancement in the understanding and application of 2D materials in extreme temperature conditions.

In this chapter, we explore the cryogenic Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs), a class of two-dimensional (2D) materials that garner significant interest in quantum computing. Our research on HfS_2 MOSFETs under cryogenic conditions spans four scenarios thoroughly discussed in: band tail (Section 4.2), incomplete ionization (Section 4.3), interface traps (Section 4.4), channel length scaling (Section 4.5)

and temperature scaling (Section 4.6). Utilizing the Büttiker probe model (Section 4.2), we align our simulations with experimental data to derive meaningful physical parameters. Each cryogenic model is initially implemented in isolation within our simulation framework, enabling an in-depth analysis of its individual impact. These findings are detailed in the respective sections. Devices with varying channel lengths are simulated in Section 4.5. A holistic simulation incorporating all models is conducted to understand their combined effects over a broad temperature range, from standard conditions to the cryogenic extreme in Section 4.6.

4.1 Simulation Setup

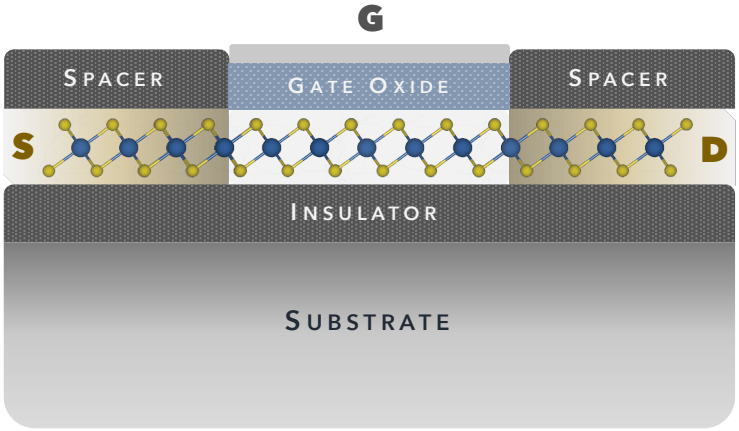


Figure 4.1: Graphical illustration of the device structure employed in the simulations, showcasing the configuration and arrangement of its elements.

We use HfS_2 (Hafnium Disulfide) MOSFETs as an example throughout the paper to illustrate the effects of cryogenic temperatures on semiconductor device behavior. HfS_2, a member of the transition metal dichalcogenides (TMDCs) family, exhibits intriguing electronic properties due to its layered structure and relatively large band gap ($E_g = 1.54\,\text{eV}$). The device structure of the HfS_2 MOSFET is depicted in Figure 4.1. The nominal device consists of an 18-nm channel region and 15-nm extensions for both source and drain, each with a doping concentration of $5 \times 10^{13}\,\text{cm}^{-2}$ in S/D regions. Periodicity is observed in the transverse directions. The channel length scaling incorporates a quantum well layer to enhance gate control and potentially reduce short-channel effects.

The band structure of HfS_2 in Figure 4.2 reveals its indirect band gap. The band structure notably includes multiple valleys at the conduction band minimum (CBM), which are critical in determining the transport properties. This complex electronic structure necessitates a comprehensive simulation approach to predict device performance accurately.

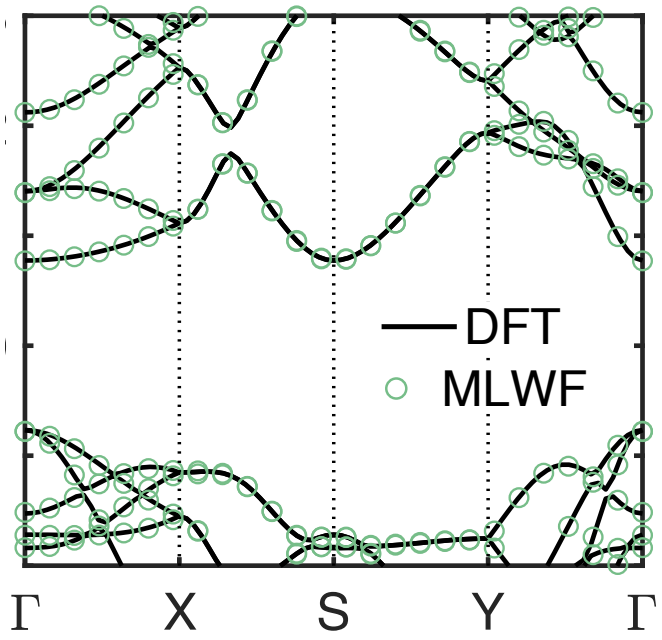


Figure 4.2: Supercell band structure of HfS_2.

4.2 Scattering Self-energy

This section presents a comprehensive analysis of the scattering mechanisms and mobility benchmarking in Hafnium Disulfide (HfS2) Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) operating under cryogenic conditions. The primary objective of this study is to systematically examine the influence of scattering processes on the resistivity and carrier mobility within these devices using a rigorous and well-structured simulation setup. With the scattering parameter η, we consider the influence of carrier scattering processes within the device, which can lead to energy loss as carriers traverse through a uniformly doped semiconductor material. The value of η is validated by aligning it with the experimentally observed carrier mobility. In Section 4.2.1, the focus shifts to

analyzing the resistive behaviors caused by scattering effects. This analysis is fundamental to the overall performance of the devices, as it underscores the importance of these effects. The detailed exploration in this part provides a deeper understanding of how scattering influences device functionality under cryogenic conditions. The Section 4.2.2 is dedicated to ensuring that the mobility results derived from simulations closely align with experimental data. This alignment is critical for validating the accuracy and dependability of the simulation framework. It guarantees that the simulations are not only theoretically sound but also practically applicable in predicting the behavior of HfS2 MOSFETs in cryogenic environments. Overall, this study provides a comprehensive understanding of the scattering mechanisms and mobility benchmarking in HfS_2 MOSFETs under cryogenic conditions.

The transport calculations are proposed to be conducted with minimal bias across the device. This ensures the presence of free carriers in quantities corresponding to the doping level by establishing the appropriate potential through the solution of the Poisson equation. The relationship $R(L) = V_D/I(L)$ allows for the determination of resistance, where V_D is the source-drain voltage and $I(L)$ the current for a given device length L. Resistance can thus be evaluated for various device lengths. By analyzing the slope of $R(L)$, the resistivity ρ can be extrapolated as delineated in Equation (4.1). The mobility, a critical parameter for characterizing carrier dynamics, is computed using Equation (4.2), where μ_n and μ_p represent the electron and hole mobilities, respectively:

$$\rho = \frac{dR}{dL} \tag{4.1}$$

$$\rho = \frac{1}{q(\mu_n n + \mu_p p)} \tag{4.2}$$

Here, q is the elementary charge, n is the electron concentration, and p is the hole concentration.

4.2.1 Resistance vs Channel Length

The parameter η can be validated further by aligning it with experimentally observed mobility. In general, to conduct alignment calculations, a small bias of $V_{DS} = 10\,\mathrm{mV}$ is applied across the device. Using $R(L) = V_{SD}/I(L)$, the resistance is computed for various device lengths, as depicted in Figure 4.3(a) and (b) for $T = 77\,\mathrm{K}$ and $300\,\mathrm{K}$, respectively. The slope of $R(L)$ allows the deduction of the resistivity ($\rho = \mathrm{d}R/\mathrm{d}L$). Considering the $n \gg p$ in n-type transistor, mobility can subsequently be approximated with $\mu_n = 1/qn\rho$

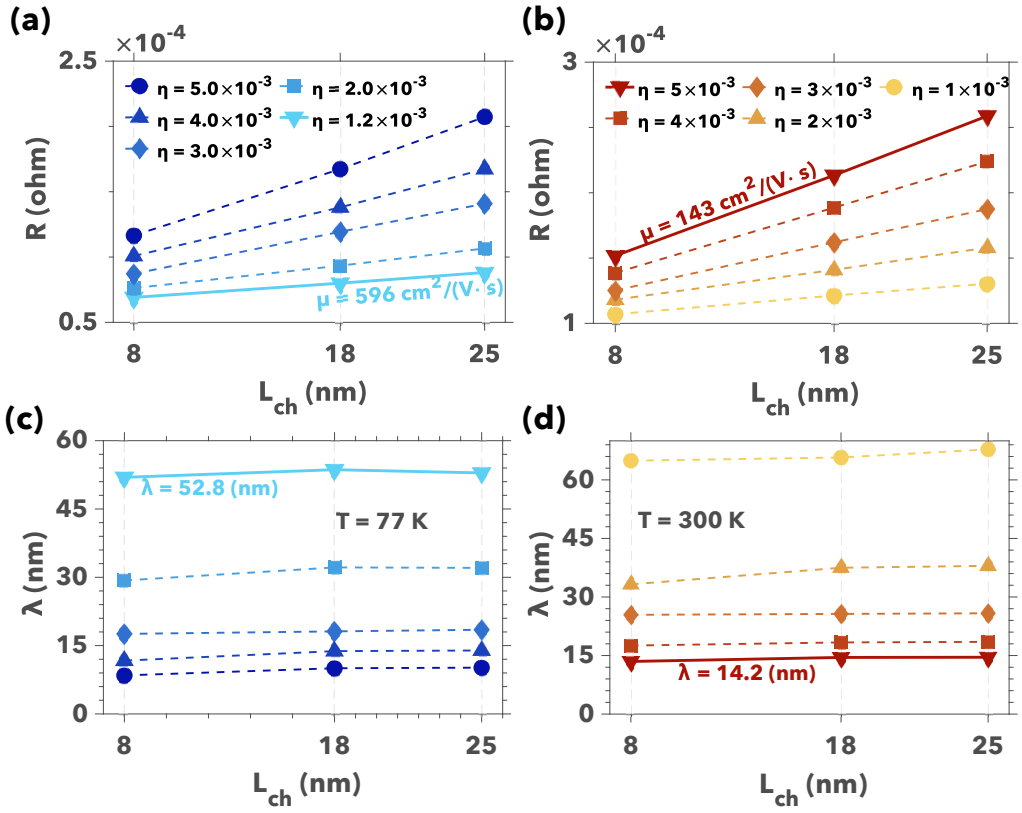


Figure 4.3: Benchmark results for scattering effects with η ranging from 1×10^{-3} eV to 5×10^{-3} eV: Calculated resistance is shown in (a) for $T = 77$ K and in (b) for $T = 300$ K, while the mean free path (λ) is presented in (c) at $T = 77$ K and in (d) at $T = 300$ K.

using Equation (4.2). The mean free path (λ), on the other hand, can be approximated by Equation (4.3) below:

$$\lambda \approx \frac{I}{I_B - I} L_{ch} \tag{4.3}$$

where the I is the drain current measured with band tail and I_B is the current with Ballistic assumption.

4.2.2 Mobility Calibration

Our newly developed Büttiker probes undergo benchmarking against state-of-the-art experimental measurements over a broad temperature range, spanning from 300 K down to 50 K. While it may seem convenient to extract parameters by simulating one temperature and then applying the parameters universally, we observed a slight deviation in fitted results for two distinct temperatures with identical scattering parameters. Consequently, our approach involves conducting simulations tailored to each specific temperature. In Figure 4.3(a) and (b), we present a series of scattering parameters for $T = 77\,\mathrm{K}$ and 300 K, respectively, serving as exemplary cases. Notably, when the scattering constant η is set to 1.2×10^{-3}, the extracted mobility aligns with the experimental value of $\mu = 596\,\mathrm{cm}^2/(\mathrm{V}\cdot\mathrm{s})$ at 77 K, whereas setting η to 5×10^{-3} achieves $\mu = 143\,\mathrm{cm}^2/(\mathrm{V}\cdot\mathrm{s})$ at 300 K. Table 4.1 presents the scattering parameters for all other temperatures using the same approach, with the mobility extracted from simulations listed against experimental data. Notably, the band tail length remains constant throughout the benchmark run.

Furthermore, we calculate the mean free path (λ) associated with each η using the formula $\lambda = L_{ch} \cdot I/(I_B - I)$, where I denotes the drain current at a specific η, and I_B is the ballistic current ($\eta = 0$). The results are visually presented in Figure 4.3(c) and (d)

Table 4.1: Scattering Benchmark Parameters

			μ ($\mathrm{cm}^2\cdot\mathrm{V}^{-1}\cdot\mathrm{s}^{-1}$)	
T (K)	κ (meV)	η (meV)	Simulation	Experiment[58]
300	100	5.0	143	140
200	100	3.5	200	201
100	100	1.5	472	469
77	100	1.2	596	597
50	100	1.0	675	680

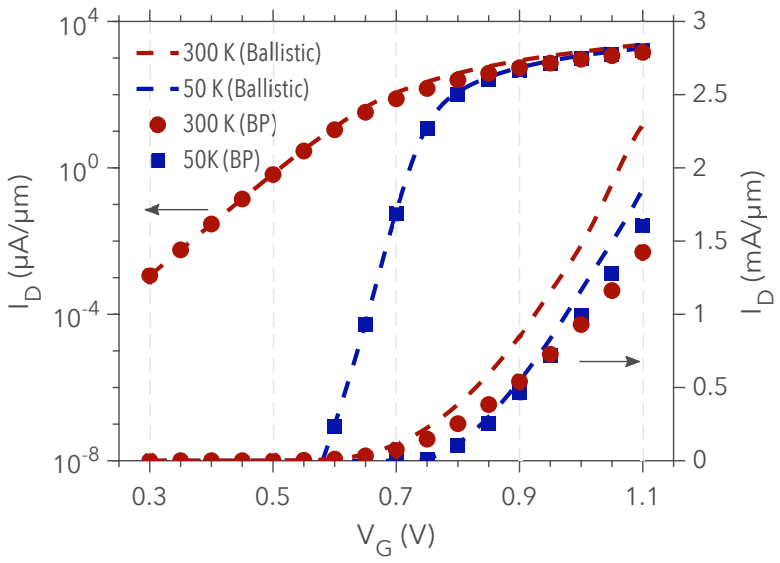


Figure 4.4: Drain current (I_D) versus gate voltage (V_G) plots comparing ballistic transport (lines) with scattering-inclusive models (markers) at temperatures of $T = 300\,\mathrm{K}$ (in red) and $T = 50\,\mathrm{K}$ (in blue).

for 77 K and 300 K, respectively, highlighting a consistent mean free path across various device channel lengths. This uniformity in mean free path values indicates the accuracy of the underlying physics in our simulation. The thorough alignment process between simulations and experimental measurements ensures the derivation of robust parameters with clear physical significance, establishing a solid foundation for subsequent simulations under diverse scenarios.

4.3 Ionization: Complete vs. Incomplete

This section is devoted to examining the phenomena of complete and incomplete ionization in HfS_2 MOSFETs at cryogenic temperatures. It explores how these ionization processes affect the electrical properties of the device, particularly focusing on the implications for quantum computing applications. The study provides a comparative analysis between complete and incomplete ionization states, elucidating their roles in determining the operational efficiency and stability of HfS_2 MOSFETs in low-temperature environments.

As a part of our comprehensive exploration of cryogenic device simulations, we meticulously delved into the numerical intricacies of HfS_2 MOSFETs, employing a model that accounts for incomplete ionization (Equation (3.35)). Our primary goal was to identify the impact of incomplete ionization on the device's behavior under cryogenic conditions. A channel dopant with dopant density of $1.5 \times 10^{13}\,\mathrm{cm}^{-2}$ and shallow dopant level at 0.1 eV below E_c is introduced in our simulation. We found that incomplete ionization has a very limited effect on our simulation configuration. This finding is observable in the I-V results at 77 K, plotted in Figure 4.5. The absence of discernible effects on the transfer characteristics associated with incomplete ionization was unsurprising, as Beckers has reported similar behavior [86].

A crucial factor contributing to this alignment with experimental results lies in the distinctive structural attributes of our simulation device, designed as a Silicon-On-Insulator (SOI) with a notably thin body. Incomplete ionization tends to manifest at deeper positions within the device structure. However, the inherently thin body of our SOI configuration limits the presence of these deeper positions, effectively diminishing the significance of incomplete ionization effects. This insight underscores the pivotal role played by the device's structural composition in shaping the observed phenomena during cryogenic simulations. Furthermore, the gate voltage electric field ionization played a crucial role in influencing the impact of incomplete ionization within the channel region of the device. The interplay between the gate voltage and ionization dynamics is the key to the device behavior, as shown in cryogenic device simulations. Considering the limited impact of incomplete ionization

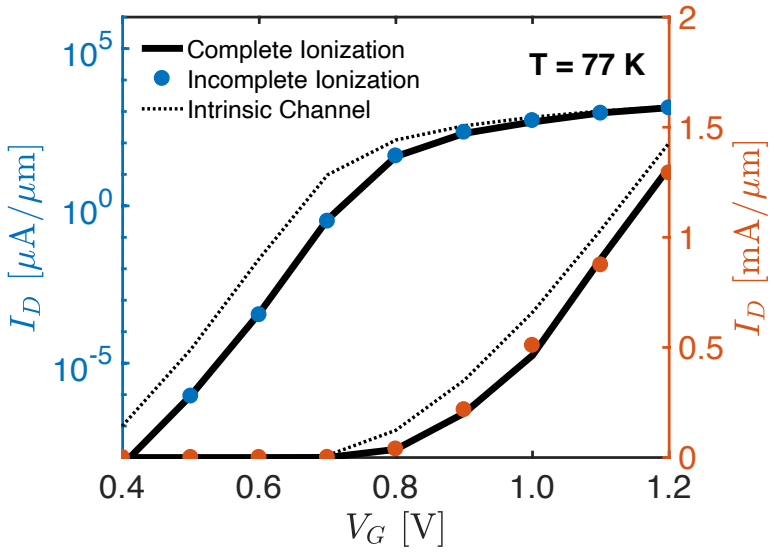


Figure 4.5: Comparison of the simulated transfer characteristic at $T = 77\,\mathrm{K}$, showcasing devices with incomplete ionization (markers), complete ionization (solid lines), and the intrinsic devices (dotted lines).

for our device, we purposely neglect the impact of substrate freeze-out for the following simulations.

4.4 Subthreshold Slope Inflection Analysis

This section offers an in-depth analysis of the Subthreshold Swing (SS) inflection in HfS_2 MOSFETs, a critical aspect of device performance at cryogenic temperatures. The subsection on Interface Trapped States investigates the impact of traps at the semiconductor-oxide interface on the SS behavior, highlighting the intricate relationship between interface states and device characteristics. Following this, the Band Tail subsection discusses the role of tail states in the energy band structure, analyzing how they contribute to the SS inflection and overall device behavior under cryogenic conditions.

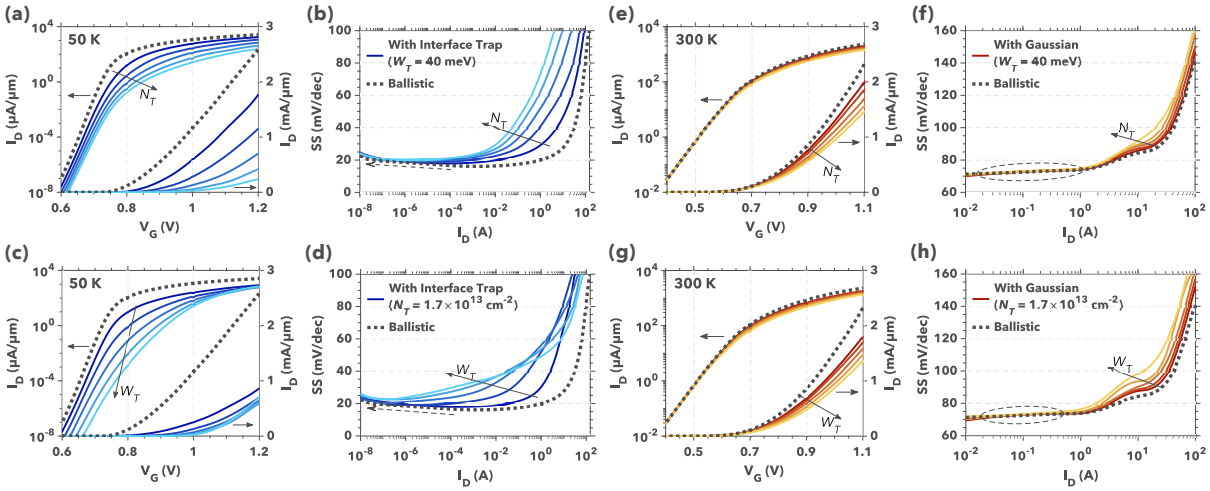


Figure 4.6: Comprehensive parametric study at 50 K (left) and 300 K (right) for varying interface trap density (N_T) and trap distribution width (W_T): (a) A marked inflection effect is observed as N_T is increased from $5.7 \times 10^{12}\,\mathrm{cm}^{-2}$ to $2.9 \times 10^{13}\,\mathrm{cm}^{-2}$, which notably decreases the ON-state current and alters its curvature in moderate inversion. The onset of inflection remains consistent, attributed to the fixed Gaussian width. (b) Deterioration of SS due to inflection and tunneling effects, with SS increasing at the lower end of I_D. (e) and (f) Enhanced W_T from 20 meV to 80 meV at the inflection initiation point ($N_T = 1.7 \times 10^{13}\,\mathrm{cm}^{-2}$). (c)-(d) and (g)-(h) Display results at 300 K for comparison, showing minimal inflection in SS and limited impact of trap states on the ON-state current.

4.4.1 Interface Trapped States

This section focuses on individually examining the interface trap's impact, facilitated by enabling an electrostatic model within the Poisson module. Characterized by a Gaussian distribution (refer to Section 3.2.2), these traps are assumed to be primarily situated around the conduction band edge, as shown in the inset of Figure 3.4. The density and width can be varied throughout the simulation. The simulations were carried out at temperatures of 50 K and 300 K, examining the effects of varying trap density (N_T) and the width of the Gaussian distribution (W_T). Two distinct simulation sets were performed. The first set adjusted N_T while maintaining a constant W_T, whereas the second set varied W_T, keeping N_T constant. The outcomes are presented in Figure 4.6, including I_D-V_G and SS-I_D graphs at both temperatures for various configurations. In Figure 4.6(a)-(b) and (e)-(f), a standard ballistic current scenario is set as a baseline (with $\eta = 0$ and $N_T = 0$), depicted by black dotted lines at $T = 50$ K. The inclusion of localized interface trapped

states, indicated by arrows in Figure 4.6(a) and (e), leads to a discernible deterioration in the subthreshold slope (SS). For a specific width W_T, the current exhibits an inflection point around $I_D = 2 \times 10^{-3}$ A, as observed in Figure 4.6(a) and (b). An increment in the trap density N_T, from $5.7 \times 10^{12}\,\mathrm{cm}^{-2}$ to $2.9 \times 10^{13}\,\mathrm{cm}^{-2}$, increases surface disorder or trap effects, noticeably degrading I_D and its slope, especially during moderate inversion phases. The variation in gate voltage causes the Fermi level (E_F) to traverse through the interface trap energy levels within the Gaussian distribution. Distinct responses in the I-V curves are associated with the Gaussian distribution's inflection points and peak. For instance, the inflection of SS happens when the Fermi energy sweeps through the first inflection point as V_G increases, as shown in Figure 4.6(a). An increase in W_0 from 20 to 80 meV, as illustrated in Figure 4.6(e), results in a convergence of the current curves along the dashed arrow, due to the invariance of trapped charge at the Gaussian maximum ($E_{F,n} = E_c$).

Additionally, these simulations were replicated at a room temperature of 300 K, showcased in fig. 4.6(c)-(d) and (g)-(h). Device behavior significantly deviates from the cryogenic conditions at this higher temperature, even under the same parameter settings. The current and switching characteristics, which show noticeable degradation at 50 K, exhibit a less pronounced decline at 300 K (see the dashed circles in fig. 4.6(d) and (h)). The influence of interface traps on the device becomes less dominant, and the transfer characteristics vary over a more restricted range at this elevated temperature. Despite the static nature of the trap distribution, there is a marked shift in the electron distribution near the Fermi level, primarily influenced by the Fermi-Dirac statistics, as elaborated in fig. 2.3. It should noted that a small kink is observed for S-I_D at 300 K in fig. 4.6(d) and (h), which could be attributed to numerical fluctuations. The influence of band tail and interface traps, more pronounced at lower temperatures, will be explored comprehensively in the following section.

4.4.2 Effect of Band Tail

The significance of band tail states is magnified as the performance of MOSFETs is increasingly influenced by the transition to cryogenic temperatures. Assessing the impact of the extended tail on the band edge requires an in-depth understanding of the band tail's characteristic length. In our study, we go beyond traditional experimental methods to investigate the band tail's implications by simulating various characteristic lengths of the Urbach tail. All configurations remain unchanged except for the band tail length κ, which is varied from 0 to 100 meV at both standard room temperature (300 K) and cryogenic temperature (50 K).

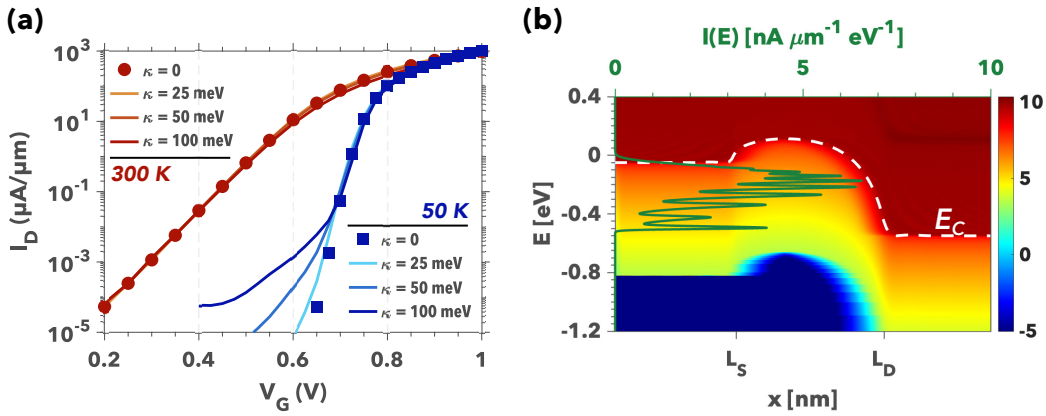


Figure 4.7: (a) Illustration of drain current (I_D) versus gate voltage (V_G) showcasing a range of band tail lengths from 25 meV to 100 meV (solid lines), compared to the ballistic case (markers). The data is presented for two temperature conditions: $T = 50$ K (in blue) and $T = 300$ K (in red).(b) LDOS plotted for a 18 nm-MOSFETs with $\kappa = 100$ meV at $T = 50$ K.

The I_D-V_G transfer characteristics at these temperatures are illustrated in Figure 4.7(a). At both 50 K and 300 K, the current in the ON-state shows negligible variation. However, the current behavior in the subthreshold region is markedly different. The rise in drain current is primarily noticeable when V_G falls below 0.7 V, a phenomenon largely attributed to tunneling current. The extended band tail introduces additional states within the band gap, facilitating a tunneling pathway. This results in a more effective channel narrowing below E_C, enhancing the probability of electron tunneling. An upward trend in current curves is observed with increasing band tail length, especially below 0.7 V. This trend confirms our analysis that attributes this behavior to tunneling current via the in-gap states created by the extended band tail. A Local density of state (LDOS) profile is plotted in Figure 4.7(b) as an example. The longer the characteristic length of the band tail, the greater the accumulation of states beneath the conduction band, increasing the likelihood of electron tunneling through the potential barrier in the channel region.

4.5 Channel Length Scaling

In this section, the focus shifts to the effects of channel length scaling on the performance of HfS_2 MOSFETs at low temperatures. The study explores how reducing the channel length influences the device's electrical characteristics, including threshold voltage, on-off

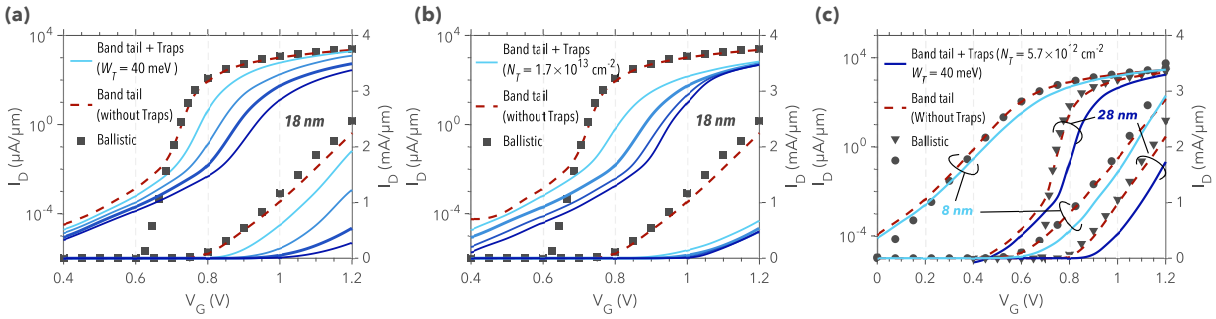


Figure 4.8: Simulated performance of a 18 nm HfS_2 MOSFET incorporating full cryogenic models: (a) Examining the effect of increasing interface trap density (N_T) with a constant trap distribution width (W_T); (b) Analyzing the impact of increasing W_T while maintaining a fixed N_T; (c) Comparing device performance for two different channel lengths, $L_{ch} = 28\,\text{nm}$ and $L_{ch} = 8\,\text{nm}$, illustrating how the influence of cryogenic models is altered by channel scaling.

ratio, and current flow. This analysis is crucial for understanding the scaling limits of HfS_2 MOSFETs in cryogenic environments, providing insights into the design and optimization of miniaturized devices for advanced quantum computing applications. Our analysis zeroes in on the effectiveness of various cryogenic devices, seamlessly integrating the cryogenic models previously discussed. We incorporate both band tail effects and interface traps into our comprehensive simulation framework, ensuring a holistic approach to the analysis. The simulation results for a channel length of $L_{\text{ch}} = 18\,\text{nm}$ are prominently displayed in Figures fig. 4.8(a) and (b), evaluated at a brisk cryogenic temperature of $T = 50\,\text{K}$.

A critical component of our analysis is the comparative assessment of transfer characteristics. This comparison juxtaposes the Ballistic model (indicated by squares in the figures) against scenarios with a band tail. Notably, the ON-state current demonstrates remarkable stability in these settings, showing only a slight decrease due to the influence of the band tail. However, the contrast becomes particularly pronounced in the subthreshold region. Here, the disparity in current characteristics grows more acute as the gate voltage decreases. The notable increase in drain current when V_{G} falls below 0.7 V, as detailed in section 4.4.2, can be traced back to the band tail's extension into the band gap. This extension results in additional states that facilitate electron tunneling, a phenomenon critical to understanding device behavior under these conditions.

Moreover, the introduction of interface traps, as depicted in fig. 4.8(a) and (b), exerts a distinct influence on the current behavior. This mirrors the effects we explored in section 4.4.1, where interface trap states were examined in isolation. However, a key difference

emerges in the altered dynamics of the subthreshold current, now heavily influenced by the simultaneous presence of the band tail effect.

To delve deeper into the significance and impact of each cryogenic model and its operational domains, we investigated the effect of channel length scaling on device performance under cryogenic conditions. This aspect of the study, illustrated in Fig. 4.8(c), is vital for discerning the dominance of each model in different scenarios. Our findings reveal that short channel effects (SCEs) become more pronounced as the channel length decreases. This intensification is primarily attributed to an increase in drain-induced barrier lowering (DIBL) and a decrease in channel modulation. This trend remains consistent across various channel lengths. Intriguingly, longer channels, such as those measuring 28 nm, exhibit behavioral patterns similar to the ones observed at 18 nm, underscoring the pervasive nature of these effects in cryogenic environments.

Overall, this section presents a nuanced, multi-faceted examination of cryogenic devices, offering valuable insights into the interplay of different mechanisms and their implications on device performance in extreme conditions.

4.6 Temperature Dependence

The final section of the chapter addresses the impact of temperature scaling on the behavior of HfS_2 MOSFETs. It examines how varying temperature conditions, particularly in the cryogenic range, alter the electrical properties of these devices. This analysis is pivotal in understanding the temperature-dependent performance of HfS_2 MOSFETs, guiding the development of reliable and efficient 2D material-based devices for use in extreme temperature environments, such as those encountered in quantum computing applications.

Our study incorporates a comprehensive range of cryogenic models into simulations to evaluate the transfer characteristics of HfS_2 MOSFETs from 300 K to 50 K. As shown in Fig. 4.9, the results are analyzed under two distinct scenarios: 1) ballistic transport and 2) the integration of all cryogenic models. A salient observation is the enhancement in switching performance (SS) with decreasing temperature, aligning with the Boltzmann limit in both scenarios. The activation of cryogenic models leads to a decrease in the ON-state current (I_{on}) across all temperatures, attributable to phonon scattering effects, as represented by the Büttiker probe. Intriguingly, lower temperatures coupled with active cryogenic models exhibit higher I_{on} values at elevated gate voltages, a trend absent in the ballistic model. Examining SS under ballistic transport reveals consistent values across temperatures, but a reduction from 69 mV/dec to 17 mV/dec is noted as temperature decreases from

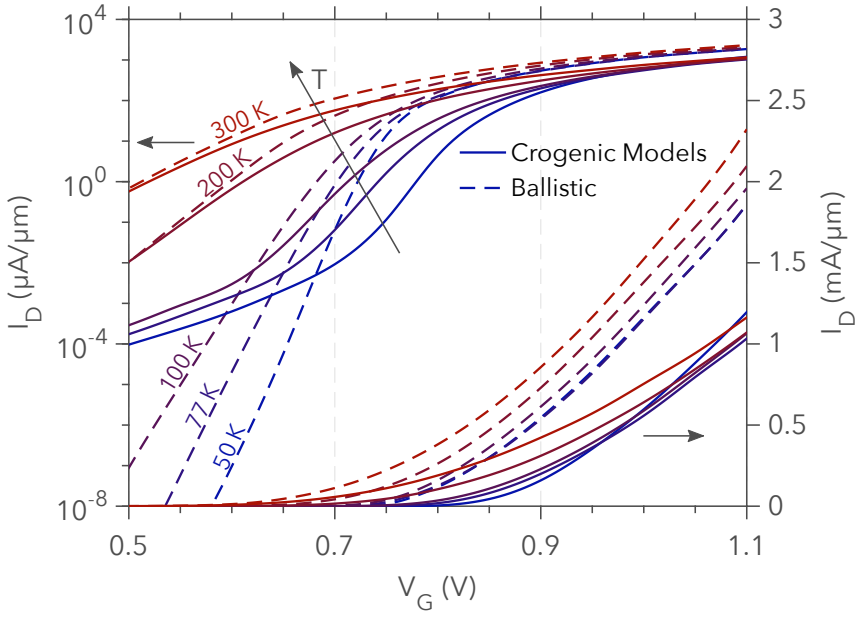


Figure 4.9: Temperature scaling from 50 K up to 300 K.

300 K to 50 K. Conversely, with cryogenic models, SS demonstrates variability with gate voltage, manifesting an inflection of SS phenomenon, consistent with the observation by Kamgar[2]. This inflection is particularly noticeable below 100 K. Given the extremely low off-state current ($< 1 \times 10^{-4}$ µA/µm), the pronounced electron distribution near the Fermi level markedly influences the in-gap density of states below E_c, resulting in a significant degradation of SS at lower temperatures. Our results corroborate experimental data, thereby validating our simulation model. Additionally, the implementation of cryogenic models unveils substantial shifts in current and SS behavior, particularly at lower temperatures.

4.7 Summary

This chapter provides an in-depth exploration of the cryogenic behavior of Hafnium Disulfide (HfS_2) Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs), a crucial class of two-dimensional (2D) materials pivotal for quantum computing advancements. Structured into five essential sections, it meticulously examines various aspects of HfS_2 MOSFETs in low-temperature environments.

The chapter commences with "Scattering and Mobility Benchmarking," where a simulation setup is established to investigate the effects of scattering on resistivity and carrier mobility in HfS_2 MOSFETs. This section's detailed approach to mobility calibration ensures that the simulations closely mirror experimental measurements, providing a dependable depiction of device behavior at cryogenic temperatures. Following this, attention is turned to "Ionization: Complete vs. Incomplete," which analyzes the impact of these ionization states on the electrical properties of HfS_2 MOSFETs. This comparative study is instrumental in comprehending operational efficiencies and stability within cryogenic settings. The "SS Inflection Analysis" section delves into the Subthreshold Swing (SS) inflection, focusing on interface-trapped states and band tail phenomena. This analysis is key in elucidating the intricate relationship between these factors and overall device performance in low-temperature conditions. "Channel Length Scaling," the subsequent section, investigates the effects of reducing channel length on the electrical characteristics of HfS_2 MOSFETs. This exploration is critical for the design and optimization of small-scale devices intended for cutting-edge quantum computing applications. The chapter concludes with "Temperature Scaling," which examines how different temperature conditions, particularly in the cryogenic range, affect the electrical properties of these MOSFETs. This section highlights the significance of understanding temperature-dependent performance in developing reliable 2D material-based devices for extreme environments.

In summary, we discuss a thorough and systematic investigation of HfS_2 MOSFETs under cryogenic conditions, delivering essential insights into the physical processes and challenges related to these 2D materials. It significantly contributes to the field of electronics for quantum computing and marks a step forward in the advancement of 2D material applications.

Chapter 5

Multi-Level Simulation Scheme

In the end, we present an improved multi-level simulation framework (in section 5.1.1) for 2D material-based nanoelectronics, which expands from device simulation, physics-based compact modeling, and circuit benchmarking using the GeH transistor as an example. In section 5.3, the device simulation employs the non-equilibrium Green's function (NEGF) method to obtain the characteristics of 2D GeH field-effect transistors for both NMOS and PMOS. In section 5.4, a compact model based on the MIT virtual source (MVS) model is then revised to capture the unique behaviors of 2D material FETs, including voltage-dependent virtual source velocity and DIBL, as well as the effect of quantum capacitance. In section 5.5, HSPICE circuit simulations are performed to analyze and optimize CMOS digital benchmark circuits. The case study demonstrates that 2D material-based transistors favor a different range of supply voltage and threshold voltage than their silicon counterpart to achieve the optimal energy-delay product. The impact of contact resistance is also analyzed using the proposed framework. This multi-level simulation approach bridges the gap between nanoelectronics and circuit behavior, thereby comprehensively advancing the understanding of materials, devices, and circuits. The framework tailored for GeH MOSFET provides accurate device-circuit co-optimization, which can be easily extended to devices based on other 2D materials.

5.1 Introduction

The rapid evolution of technology has led to an increasing demand for electronic devices that are faster, more efficient, and cost-effective. As a result, researchers have been exploring the potential of two-dimensional (2D) materials, such as molybdenum disulfide (MoS_2),

black phosphorus (BP), and germanane (GeH), as channel materials for the next-generation complementary metal oxide semiconductor (CMOS) technology[32, 24, 25, 55, 56]. Even at the initial stage, these materials have demonstrated tremendous potential owing to their exceptional characteristics, including high ON/OFF current ratio, fast switching speed, and immunity to short-channel effects[57, 58, 59, 60, 61]. For instance, previous theoretical studies predicted that short-channel GeH field-effect transistors (FETs) would be promising due to their high field-effect mobility ($\sim 150\,\mathrm{cm^2 \cdot V^{-1} \cdot s^{-1}}$) and large ON/OFF current ratio ($> 10^5$)[55, 62].

Despite significant efforts to fabricate devices based on 2D materials, a systematic design and simulation framework is still lacking for device engineering and circuit performance optimization. This is especially important for novel nanoelectronics based on 2D materials, as experiments can be expensive. Thus, a multi-level simulation framework that covers materials, devices, and circuits should be developed. Creating a proper compact model for 2D FETs is crucial in connecting the intrinsic material properties and device characteristics to circuit performance. Existing transistor compact models[63, 64, 65], primarily based on the classical drift-diffusion theory[66], are inadequate for nanotransistors using 2D materials. While there are a few compact models that capture ballistic transport for silicon devices, such as the MVS model[67], they fail to include the distinct physics of low-dimensional materials, such as quantum capacitance in 2D FETs. Therefore, a compact model tailored explicitly for 2D FETs is required, incorporating the effects and physics unique to the 2D structure. While several models have been developed to evaluate 2D FETs beyond their device-level performance either using compact modeling[68, 69, 71] or directly based on atomistic-level simulation[70], there remains a gap in the literature regarding comprehensive multi-level simulation with a full scope from materials, device, compact model to circuit analysis for 2D-material based FET exhibiting ballistic transport.

In recent developments, a modified MVS compact model has emerged within a simulation framework [87] aimed at assessing the performance of monolayer GeH FETs. This represents a significant stride towards establishing a comprehensive multi-level simulation paradigm. This framework seamlessly marries the Non-Equilibrium Green's Function (NEGF) simulation, responsible for extracting crucial physical parameters, with the compact model, tailored to capture specific phenomena inherent in 2D MOSFETs [88, 89]. Nevertheless, the earlier iteration encountered challenges related to accurately fitting current in the saturated region of the transistor, and it lacked the essential PMOS device characteristics necessary for circuit analysis. In the pursuit of advancing this research, we present substantial enhancements to the previous work [87] on multi-level 2D simulation, resulting in a far-reaching simulation framework that encompasses NEGF simulations, a 2D MOSFET model, and a circuit-level simulator designed for energy-delay optimization.

A key innovation involves the integration of an advanced compact model, which substantially improves the accuracy of the current fitting for the device. This enhancement incorporates critical features such as voltage-dependent carrier velocity, drain-induced barrier lowering (DIBL), and the quantum capacitance effect. These modifications are specifically tailored to refine the MVS compact model for 2D-material FETs, considering the unique phenomena stemming from the low density of states (DOS) observed in 2D materials. Moreover, we employ HSPICE simulation to comprehensively assess the circuit-level performance of GeH FETs, adopting the compact model and parameters developed for both NMOS and PMOS devices. Furthermore, we delve into the influence of contact resistance using HSPICE simulation, thereby providing a realistic evaluation of circuit performance. By leveraging the inherent synergy between NEGF and MVS, which mutually cross-validate each other's physics, we significantly bolster the reliability of our simulation results. This integrated approach empowers us to rigorously compare and validate the outputs of both models, ensuring a high degree of consistency and accuracy in predicting device behavior. Consequently, our circuit-level energy-delay analysis is invaluable for gaining profound insights into technology-device-circuit co-design.

5.1.1 Simulation Overview

The hierarchical framework of the multi-level simulation (Fig. 5.1) formed with device simulation, compact modeling, and circuit analysis was introduced in our previous work. In the device simulation, we use GeH, a hydrogenated germanium monolayer, as the channel material. Tight-binding (TB) parameters [90] of monolayer GeH are extracted by fitting the DFT bands [91] for the material input of the device simulation [10]. At the first stage, the NEGF equation [76] is adopted along with Poisson's equation to calculate the intrinsic device performance of short channel GeH FETs where quantum mechanical effects such as carrier tunneling and quantum capacitance are inherently included. In the second step, a new compact model based on the MVS model is developed to reproduce the device characteristics obtained from the NEGF simulation. Finally, circuit analysis is conducted using HSPICE with the developed compact model. Integration of the Non-Equilibrium Green's Function (NEGF) equation and the MVS compact modeling approach, which cross-validate each other's results. Notably, the velocity, DIBL, and quantum capacitance (C_Q) are examples of such parameters. This combination provides a comprehensive understanding of the device behavior by considering both quantum effects and circuit-level performance.

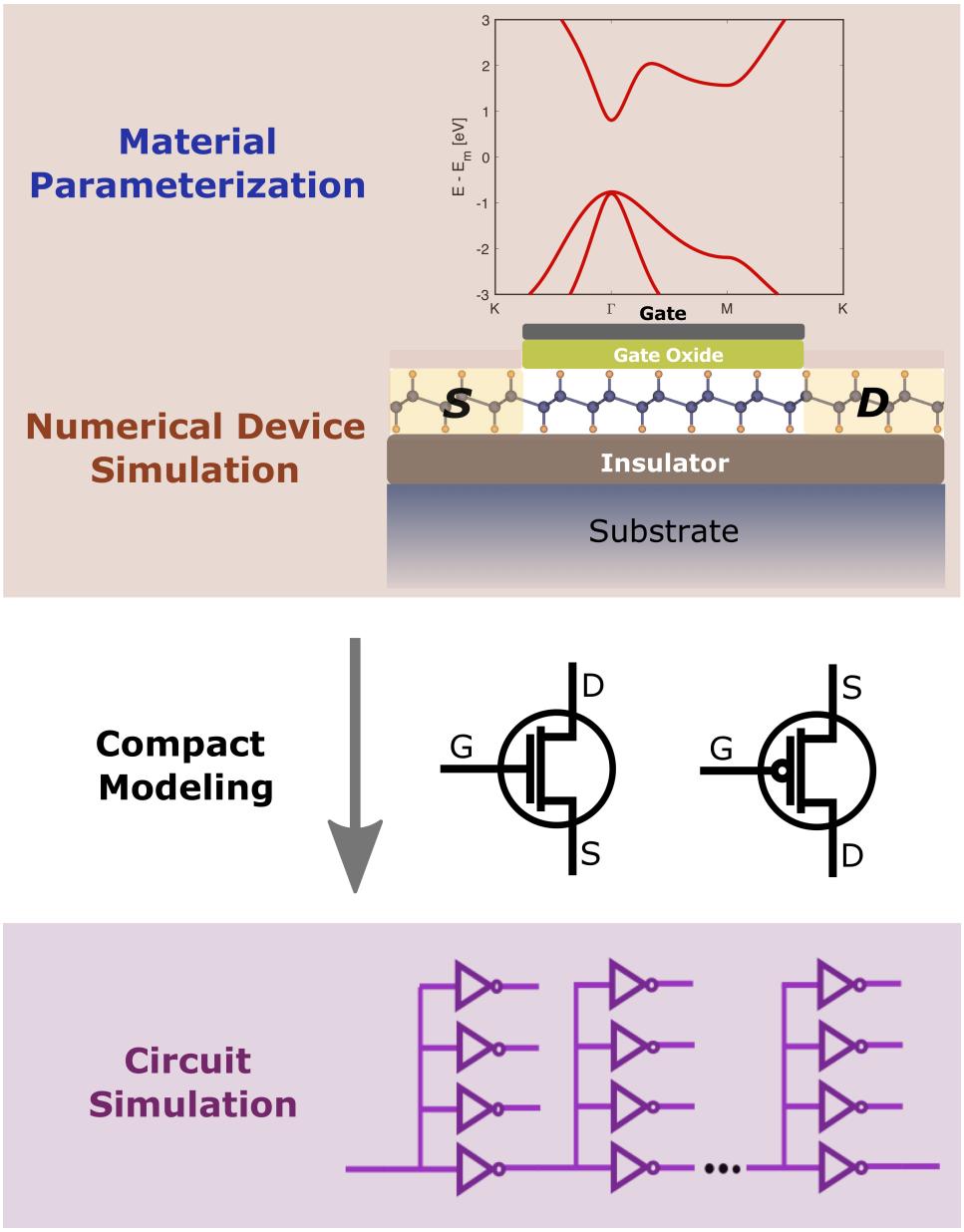


Figure 5.1: An overview of the multi-level simulation. It starts with intrinsic device simulation, including materials parameterization and numerical device simulation. A compact model bridges the gap between device and circuit simulations.

5.2 Materials Parameterization

The scaling of conventional silicon metal-oxide-semiconductor field-effect transistors (MOSFETs) is nearing fundamental limits, necessitating advanced simulation tools for novel materials and nanoscale device applications. These tools are critical for reducing the time and cost associated with experimental explorations of new materials and devices. Computer-aided design has become increasingly important in this regard for its efficiency. A fundamental aspect of device simulation is the characterization of material properties.

5.2.1 DFT Calculations

Density Functional Theory (DFT) is a quantum mechanical method used to investigate the electronic structure of many-body systems. It is based on the principle that a function of the electron density alone can determine the ground state properties of a quantum system. This approach circumvents the complexity of solving the many-body Schrödinger equation directly.

The many-body Schrödinger equation for a system of N particles is:

$$\left[\sum_{i}^{N} -\frac{\hbar^2}{2m_i}\nabla_i^2 + V(\mathbf{r}_1, ..., \mathbf{r}_N)\right] \Psi(\mathbf{r}_1, ..., \mathbf{r}_N) = E\Psi(\mathbf{r}_1, ..., \mathbf{r}_N) \tag{5.1}$$

Due to the complexity of solving this equation for systems with many electrons, approximations such as the Born-Oppenheimer approximation are utilized. This approximation decouples the nuclear and electronic motion, allowing the electrons to be treated as moving in a static potential created by the nuclei. To further simplify the problem, DFT introduces the electron density $n(\mathbf{r})$, which is a function of only three spatial coordinates, regardless of the number of electrons. This density is given by:

$$n(\mathbf{r}) = N \int |\Psi(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N)|^2 \, d\mathbf{r}_2 ... d\mathbf{r}_N \tag{5.2}$$

The problem's reformulation in terms of electron density leads to the Kohn-Sham equations, which are solvable and provide a practical method for calculating the electronic structure in many electron systems.

Kohn-Sham Equations

The Kohn-Sham approach to Density Functional Theory (DFT) represents the interacting electrons within a material by a non-interacting reference system while accounting for electron correlation and exchange effects through the use of an effective potential. The Kohn-Sham equations, which are derived from this approach, facilitate the practical calculation of the ground-state density and energy of a many-electron system. These equations are given by:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{5.3}$$

where $V_{\text{eff}}(\mathbf{r})$ includes the external potential, the average electron-electron interactions, and the exchange-correlation potential, and $\phi_i(\mathbf{r})$ are the Kohn-Sham orbitals. Solving these equations self-consistently yields the electron density $n(\mathbf{r})$, from which one can deduce material properties critical for simulating novel semiconductor devices.

The Hamiltonian in the many-electron Schrödinger equation encompasses the kinetic energy $\hat{T}$, the external potential $\hat{V}_{\text{ext}}$, and the electron-electron interaction $\hat{V}_{\text{ee}}$:

$$\hat{H} = \sum_i^N -\frac{\hbar^2}{2m}\nabla_i^2 + \sum_i^N V_{\text{ext}}(\mathbf{r}_i) + \sum_{i<j} U(\mathbf{r}_i, \mathbf{r}_j) \tag{5.4}$$

Two central theorems underpin DFT:

- The ground-state energy is a unique function of the electron density.
- The electron density that minimizes the energy functional corresponds to the true ground-state electron density.

For non-degenerate ground states, the wave function is uniquely determined by the external potential $\hat{V}$, leading to a unique potential for a given electron density. The energy functional $E[\rho]$ is expressed as:

$$E[\rho] = F[\rho] + \int \rho(\mathbf{r})V_{\text{ext}}(\mathbf{r})d\mathbf{r} \tag{5.5}$$

where $F[\rho]$ is a universal functional comprising kinetic and electron-electron interaction energies, $F[\rho] = T[\rho] + E_{\text{ee}}[\rho]$. The ground state energy E_0 for a given external potential V_{ext} is:

$$E_0 = E[\rho_0] \tag{5.6}$$

Kohn-Sham DFT effectively reduces the complexity of the many-electron problem by transforming it into a set of single-electron equations. The effective potential V_{eff} in the Kohn-Sham equations is a sum of the external potential, the Hartree potential representing electron-electron repulsions, and the exchange-correlation potential accounting for quantum mechanical exchange and correlation:

$$\hat{T} = -\frac{\hbar^2}{2m}\nabla^2 \tag{5.7}$$

$$V_{\text{ne}} = -\sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \tag{5.8}$$

$$V_{\text{H}} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \tag{5.9}$$

$$V_{\text{XC}} = \frac{\delta E_{\text{XC}}[\rho]}{\delta \rho(\mathbf{r})} \tag{5.10}$$

Thus, the Kohn-Sham equations are a pivotal tool in material science, providing a practical route to understanding the properties of complex systems.

One can determine the electronic structure by employing a self-consistent computational scheme:

1. Begin by initializing an electron density guess $n(\mathbf{r})$.
2. Solve the Kohn-Sham equations with the initial $n(\mathbf{r})$ to obtain the single-electron wave functions Ψ_i.
3. Update the electron density $n_{\text{new}}(\mathbf{r})$ using the wave functions according to $n_{\text{new}}(\mathbf{r}) = \sum_i |\Psi_i(\mathbf{r})|^2$ where the sum runs over occupied states.
4. Compare $n_{\text{new}}(\mathbf{r})$ with $n(\mathbf{r})$. If they are not equal within a set tolerance, adjust $n(\mathbf{r})$ using a mixing scheme and return to step 2. If they are equal, proceed to calculate the ground-state energy E_0.

LDA and GGA

To approximate the exchange-correlation energy E_{xc}, one may use the Local Density Approximation (LDA) or the Generalized Gradient Approximation (GGA). LDA assumes a uniform electron gas model, which is adequate for many quantum chemistry applications

but may underestimate the band gap in solid-state physics. GGA improves upon LDA by considering the gradient of the electron density, yielding more accurate results for the band structure of materials.

Several computational packages are available for DFT simulations, such as VASP, Quantum Espresso, Abinit, and Siesta. Quantum Espresso, in particular, is a widely utilized open-source software for electronic-structure calculations and materials modeling on the nanoscale, which employs plane waves and pseudopotentials.

For device simulations that extend beyond the capabilities of traditional DFT, such as when considering larger systems or requiring more nuanced models, methods like Non-Equilibrium Green's Function (NEGF), effective mass, tight-binding (TB), and maximally localized Wannier functions (MLWF) can be more suitable. These methods provide a practical means to study material properties and device performance on sub-micrometer scales.

5.2.2 Tight-Binding (TB) Parameterization

While Density Functional Theory (DFT) can produce more accurate results, generating Hamiltonians for other calculations like the Non-Equilibrium Green's Function (NEGF) method is computationally expensive. Multiple approaches have been suggested, such as the effective mass approximation and the k.p method. These methods capture only the minimal features needed for NEGF, such as the effective mass of the conduction and valence bands or the band gap. On the other hand, the Tight-Binding (TB) method can provide a closer approximation of the DFT results. The TB Hamiltonian is based on atomic-like basis sets. Instead of using the multi-body Hamiltonian operator shown in eqn. (2.19), a parameterized Hamiltonian is employed. These parameters preserve features like the effective symmetry and distance without significantly increasing computational time. This is achieved by limiting the parameters to a minimal number within the energy range of interest. Two main approaches to constructing the TB Hamiltonian are the Slater-Koster TB parametrization and the maximally localized Wannier's Function (MLWF).

The Slater-Koster TB method populates the Hamiltonian using hopping integral-based parameters. These integrals stem from the electronic Hamiltonian in Equation (5.3). If one neglects V_H and V_{XC}, the new Hamiltonian takes the form of

$$H_{elec}(r) = T + V_{NE} = -\frac{1}{2}\nabla^2 + \sum_{i,\alpha} v(\vec{r} - \vec{T}_i - \vec{R}_\alpha) = -\frac{1}{2}\nabla^2 + v(r) \tag{5.11}$$

where $\vec{T}_i$ is the lattice vector and $\vec{R}_\alpha$ is the position of the basis atom. The Bloch wavefunction of each atomic orbital with quantum numbers lm will be

$$\psi^{\alpha}_{lm}(\vec{r},\vec{k}) = \frac{1}{\sqrt{N}} \sum_i e^{i\vec{k}\cdot\vec{T}_i} \psi_{lm}(\vec{r} - \vec{T}_i - \vec{R}_\alpha) \tag{5.12}$$

The Hamiltonian and the overlap matrix for the Bloch wavefunction are calculated using

$$H^{\alpha,\alpha'}_{lm,l'm'}(\vec{k}) = \langle \psi^{\alpha}_{lm}(\vec{k}) | H_{elec} | \psi^{\alpha'}_{l'm'}(\vec{k}) \rangle \tag{5.13}$$

$$S^{\alpha,\alpha'}_{lm,l'm'}(\vec{k}) = \langle \psi^{\alpha}_{lm}(\vec{k}) | \psi^{\alpha'}_{l'm'}(\vec{k}) \rangle \tag{5.14}$$

Equations (5.13) and (5.14) define a generalized eigenvalue problem from which the eigenvalues of the wavefunction can be obtained. The new Hamiltonian $H^{\alpha,\alpha'}_{lm,l'm'}$ can be written in a tight-binding context to approximate the DFT results.

The Hamiltonian $H^{\alpha,\alpha'}_{lm,l'm'}(k)$ utilizes the atomic energy level $\varepsilon^{\alpha,0}_{l}$, the crystal-field matrix $\Delta\varepsilon^{\alpha}_{lm,l'm'}$, and the hopping integral $t^{\alpha,\alpha'}_{lm,l'm'}$ as follows:

$$H^{\alpha,\alpha'}_{lm,l'm'}(k) = \varepsilon^{\alpha,0}_{l} S^{\alpha,\alpha'}_{lm,l'm'}(k) + \Delta\varepsilon^{\alpha}_{lm,l'm'}\delta_{\alpha,\alpha'} - \frac{1}{N}\sum_{i,\alpha} e^{i(k\cdot(T_i - T_j))} t^{\alpha,\alpha'}_{lm,l'm'} \tag{5.15}$$

where $t^{\alpha,\alpha'}_{lm,l'm'}$ is defined by

$$t^{\alpha,\alpha'}_{lm,l'm'} = \int d\vec{r} \psi^{\alpha}_{lm}(\vec{r} - \vec{R}_\alpha)[v(\vec{r} - \vec{R}_\alpha)]\psi^{\alpha}_{l'm'}(\vec{r} - \vec{R}_\alpha) \tag{5.16}$$

The two-center hopping integral $t^{\alpha,\alpha'}_{lm,l'm'}$, shown in eqn. (2.50), comprises both two-center and three-center terms. The three-center term can often be neglected if the basis is sufficiently localized.

$$v^{\alpha,\alpha'}_{lm,l'm'} = \int d\vec{r} \psi^{\alpha}_{lm}(\vec{r} - \vec{R}_\alpha)[v(\vec{r} - \vec{T}_i - \vec{R}_\alpha) + v(\vec{r} - \vec{T}_j - \vec{R}_\alpha)]\psi^{\alpha}_{l'm'}(\vec{r} - \vec{R}_\alpha) \tag{5.17}$$

Slater-Koster provided a simplification using cubic harmonic orbitals. A linear combination of bond integrals can approximate the interaction between two atomic bases. Listed below are some examples of these energy integrals from the two-center approximation:

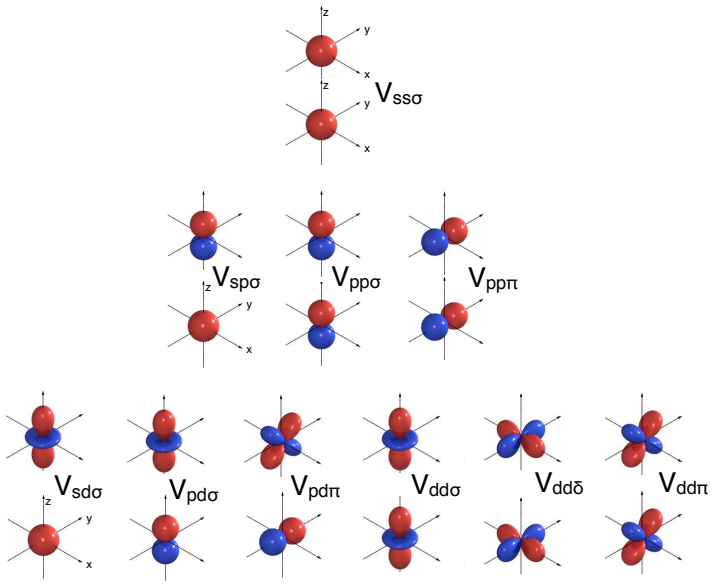


Figure 5.2: Two-center Slater-Koster integrals for atomic orbitals s, p, and d, illustrating bond types like σ, π, and δ [9].

$$
\begin{aligned}
E_{ss} &= V_{ss\sigma} \\
E_{sx} &= lV_{sp\sigma} \\
E_{xx} &= l^2 V_{pp\sigma} + (1 - l^2) V_{pp\pi} \\
&\vdots \\
E_{xyxy} &= 3l^2 m^2 V_{dd\sigma} + (l^2 + m^2 - 4l^2 m^2) V_{dd\pi} + (n^2 + l^2 m^2) V_{dd\delta}
\end{aligned}
$$

Direction cosines l, m, n are defined with respect to the distance d between the two basis atoms as:

$$
l = \frac{(R_j - R_i) \cdot \vec{x}}{d}, \quad m = \frac{(R_j - R_i) \cdot \vec{y}}{d}, \quad n = \frac{(R_j - R_i) \cdot \vec{z}}{d} \tag{5.18}
$$

The terms mentioned above account for all types of bonds, from the strongest σ bond to the π and δ bonds, which are plotted in Figure 5.2 [9].

To determine the Slater-Koster integrals, empirical fitting of DFT eigenvalues is commonly employed, following these steps:

1. Guess initial values for all integrals $v^{\alpha,\alpha'}_{lm,l'm'}$, atomic energy levels $\varepsilon^{\alpha,0}_{l}$, and crystal-field matrix $\Delta\varepsilon^{\alpha}_{lm,l'm'}$.

2. Solve the eigenvalue problem using $H^{\alpha,\alpha'}_{lm,l'm'}(k)$.

3. Compare TB eigenvalues E_{TB} with DFT ones E_{DFT}. If the Mean Square Error (MSE) is below the desired threshold, finalize; otherwise, update $v^{\alpha,\alpha'}_{lm,l'm'}$, $\varepsilon^{\alpha,0}_{l}$, and $\Delta\varepsilon^{\alpha}_{lm,l'm'}$.

TB parametrization for Germanene (GeH)

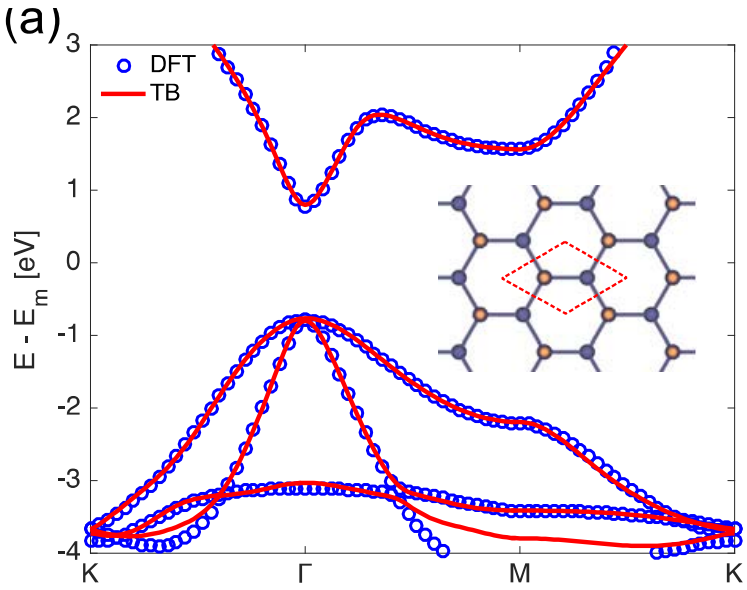


Figure 5.3: The electronic band structure of germanane (GeH) as determined through Density Functional Theory (DFT) is represented by blue circles, while the solid red lines indicate the fit achieved through tight-binding (TB) parametrization. The inset presents a top view of the germanane lattice, with the unit cell demarcated by red dotted lines [10].

Density Functional Theory (DFT) calculations were utilized to obtain the band structure of germanane (GeH). The Generalized Gradient Approximation (GGA) with the Perdew-Burke-Ernzerhof (PBE) parameterization was employed to capture the exchange-correlation energy accurately. For geometry optimization, the conjugate gradient method was utilized, setting the force tolerance to 10^{-3}eV/Å and the energy tolerance to 10^{-4}eV. The primitive cell of germanane, containing two Ge atoms and two H atoms, is depicted in the inset of

Figure 5.3. Brillouin zone sampling was performed using the Monkhorst-Pack scheme with a $50 \times 50 \times 1$ mesh.

The resulting band structure is displayed in Figure 5.3, showing isotropic effective masses near the Γ point. The electron effective mass is quantified at $0.07m_0$ and the heavy and light hole effective masses at $0.50m_0$ and $0.07m_0$, respectively (m_0 is the free electron mass). These values are in good agreement with those reported in the literature. An underestimation in the DFT-calculated band gap (E_g) was corrected to 1.56eV, using the scissors correction scheme, to align with the literature-suggested range of $1.48 - 1.60$eV. This adjustment to E_g does not influence the analysis of n-type transport behaviors in GeH FETs, which is the focus of our study.

Tight-binding (TB) parameters were extracted from the DFT band structure using a two-center Slater-Koster approximation with consideration for up to second-nearest neighbors. Both s and p atomic orbitals were included for Ge, while only the s orbital was considered for H. An orthonormal basis set was utilized for the calculations. The TB parameters were refined by fitting to the DFT band structure using a Mean Squared Error (MSE) approach. The refined TB parameters accurately replicated the DFT band structure, especially near the band edges, as evidenced in Figure 5.3.

5.3 Device Simulation

5.3.1 Simulation Flowchart

For nanotransistors based on 2D materials, conventional drift-diffusion transport models fail to capture the complex quantum mechanical phenomena at play, particularly when the device dimensions approach the mean free path of the carriers. This is especially pertinent in the context of 2D-material Field-Effect Transistors (FETs), where the typical transport models give way to ballistic or quasi-ballistic transport models. These models are essential for accurately simulating device characteristics when the carrier mean free paths are in the same order as the channel length.

The Non-Equilibrium Green's Function (NEGF) method is widely used for the simulation of these 2D-material FETs, as it incorporates the quantum transport equation and self-consistently solves Poisson's equation to determine the electrostatic potential across the device. An iterative process is necessary to achieve a converged solution, which involves initializing the device with an electrostatic potential, computing the self-energy terms based on the contacts, and solving the NEGF equation to obtain the electron density and current distribution. After this, Poisson's equation is solved to update the electrostatic potential, which is repeated until convergence is reached.

The iteration process, illustrated in Figure 5.4, ensures that the electrostatic potential and the electron density are consistently updated, leading to a solution that accurately reflects the electronic properties of the device. This is particularly important in devices where electrostatic effects are dominant, such as quantum dots and field-effect transistors, where the self-consistent treatment of the electronic states and the electrostatic environment is crucial for accurate simulation results.

5.3.2 GeH MOSFET

In this work, monolayer GeH FET with an ideal ohmic contact is simulated (see the device structure in Fig. 1). The material property is described by TB Hamiltonian with a second-nearest neighbor approximation. The contribution of s and p orbitals for Ge and s orbital for H are considered. Based on this set of TB parameters, the band structure is well-matched with DFT bands in the previous work [22]. The nominal device is set up as follows: 15-nm channel length (L_{ch}), 3.85 nm-thick Al_2O_3 ($k = 9$) gate dielectric (equivalent oxide thickness of EOT = 1.67 nm), and 20-nm extended source and drain. For the NMOS, a doping concentration of $8.25 \times 10^{12}\,\mathrm{cm}^{-2}$ is used for the source/drain, whereas

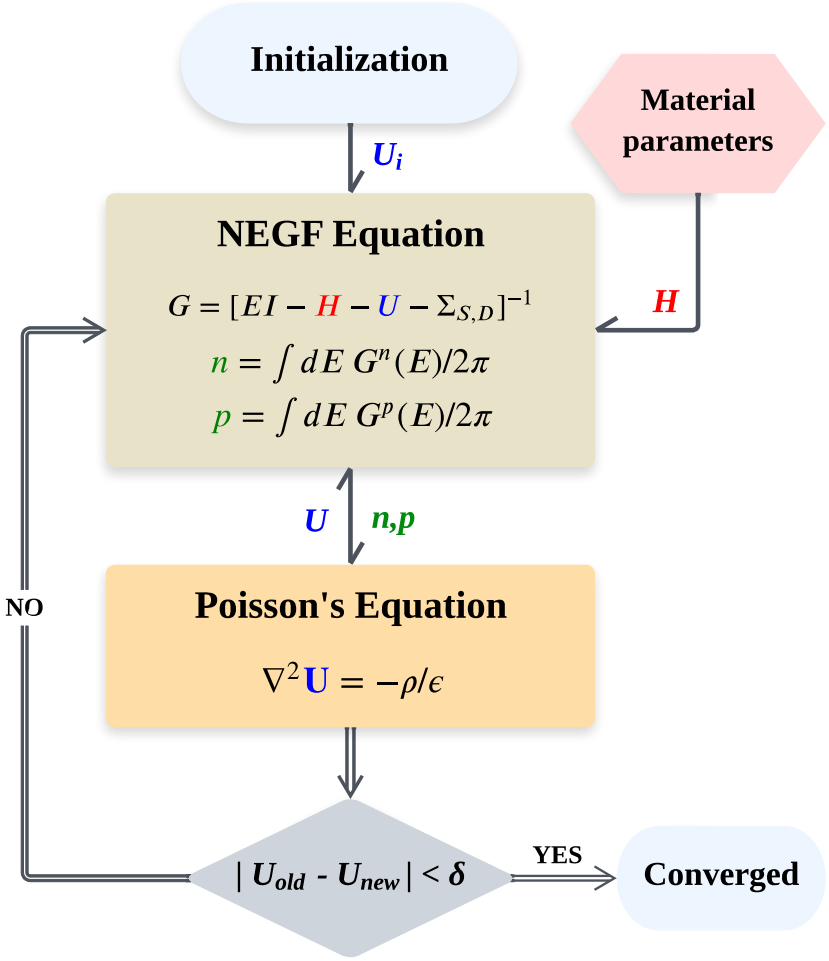


Figure 5.4: The flowchart demonstrating the iterative NEGF and Poisson self-consistent simulation methodology for GeH FETs. This illustrative flowchart visually elucidates the intricacies of the iterative procedure employed to solve the Poisson and NEGF equations concurrently. This iterative approach culminates in precisely determining the electrostatic potential and electron density of the GeH FET device, facilitating a comprehensive understanding of its electronic behavior.

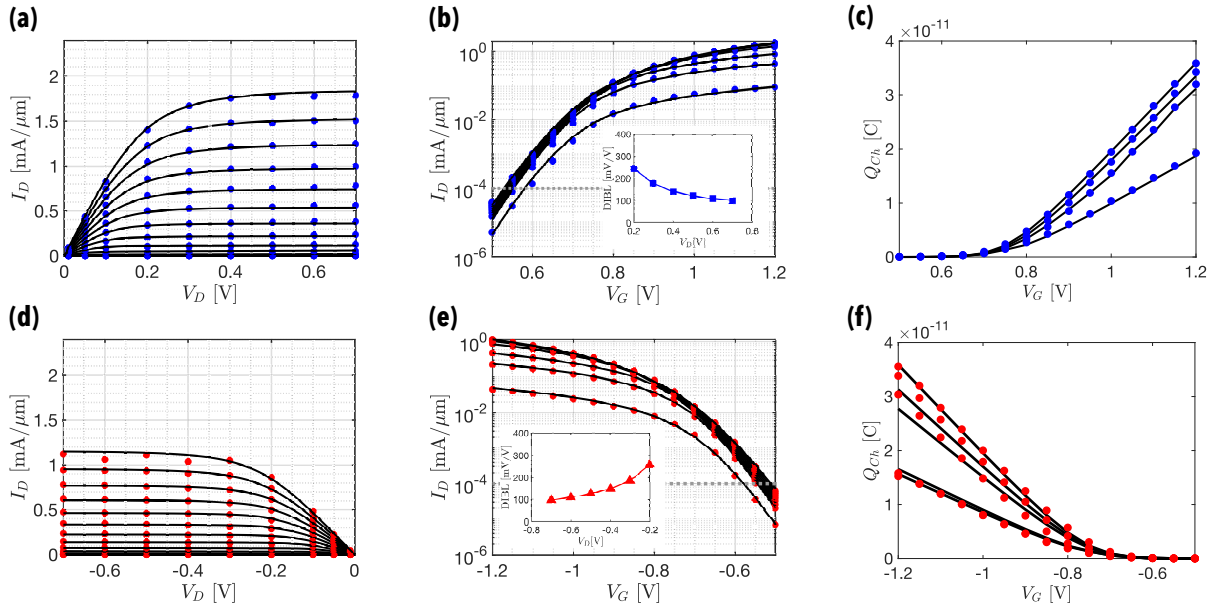


Figure 5.5: Drain current of GeH FETs at drain bias ($|V_{DS}|$) ranging from 0 V to 0.7 V and gate bias ($|V_{GS}|$) from 0.5 V to 1.1 V. Results from both NEGF simulations (dots) and the modified MVS compact model (solid lines) are shown. I_{DS}-V_{DS} for (a) NMOS and (d) PMOS; I_{DS}-V_{GS} for (b) NMOS and (e) PMOS where DIBL vs. V_{DS} are plotted in the insert in (b) and (e) for both NMOS and PMOS, respectively. Total channel charge as a function of gate voltage Q_{ch}-V_{GS} are also plotted for (c) NMOS and (f) PMOS at different drain voltages (V_{DS}).

a higher doping concentration of $4.95 \times 10^{13}\,\mathrm{cm}^{-2}$ is adopted for the PMOS considering the larger DOS in the valence band. Ballistic transport is assumed due to the relatively short channel length considered in this study. Contact resistance is assumed to be zero in NEGF simulation.

We assess the intrinsic device performance of GeH NMOS and PMOS with NEGF device simulation and present the device characteristics in Fig. 5.5, including current characteristics (Fig. 5.5(a) and (b) for NMOS with blue dots, and Fig. 5.5(d) and (e) for PMOS with red dots) and total channel charge curves as a function of gate voltage (Fig. 5.5(c) and (f) for NMOS and PMOS, respectively). Our results indicate that while some similarity can be observed between GeH n-type and p-type MOSFET (NMOS and PMOS), NMOS exhibits a higher saturated current than PMOS at all gate voltages due to the smaller effective mass of electrons in GeH [10]. Both devices show excellent gate control and demonstrate

subthreshold swing just above the theoretical Boltzmann limit (66.0 mV/dec for PMOS and 64.8 mV/dec for NMOS). Moreover, the total charge for PMOS is slightly higher than that of NMOS for the same bias point, which can be attributed to the higher density of states of holes in GeH. The details of those discussed are provided in later sections 5.4.1 and 5.4.2. NEGF simulation results serve as input for the compact model in our multi-level scheme, and the intrinsic characteristic extracted from NEGF is a good starting point for calibrating the parameters in the MVS model.

5.4 MVS Compact Model

To study the performance of circuits based on GeH FETs, a compact model is introduced to enable computationally efficient circuit simulations. A semi-empirical virtual source model (MVS) [64] that describes the device characteristics of short-channel MOSFETs is chosen as the starting point, given its ballistic nature and simplicity in its model formulation. One of the advantages of the MVS model is the usage of only a limited number of model parameters. Most of these parameters have straightforward physical meanings, allowing relatively simple verification from numerical simulation results. However, the MVS model was developed for the quasi-ballistic transport of three-dimensional (3D) silicon devices where virtual source velocity is assumed to be constant. Therefore, it is incompetent to capture some physics of 2D-material FETs. Considering the unique underlying physics of 2D materials and their device, we proposed an updated MVS model to match the transport characteristics obtained from the NEGF quantum transport simulation. The model is improved for 2D MOSFET mainly in three parts: 1) considering the effect of quantum capacitance 2) modeling the virtual source velocity with a voltage-dependent equation and 3) introducing a voltage-dependent DIBL. First, we evaluate the gate capacitance using Eq(5.19). In 2D FET, gate capacitance (C_G) is not always dominated by the oxide capacitance C_{ox} and quantum capacitance (C_Q) plays an important role in determining C_G along with C_{ox}.

$$1/C_G = 1/C_{ox} + 1/C_Q \tag{5.19}$$

Here, C_Q is considered by default in NEGF simulation. Second, the gate voltage-dependent virtual source velocity is newly introduced as

$$v_x = v_{x0} \cdot [(V_G - V_{TH})/V_0] \tag{5.20}$$

where v_{x0} is virtual source velocity, V_{GS} is gate voltage, V_{TH} is threshold voltage, and V_0 is set to 1 V (for the unitless). The virtual source velocity changes linearly at the ON state with the new empirical equation. Third, the voltage-dependent DIBL is redefined by the equation as

$$DIBL = \delta \left\{1 + \exp\left[(V_G - V_{TH})/V_1\right]\right\}^{-1} \tag{5.21}$$

where δ is a fitting parameter of nominal DIBL, and V_1 is set to 1×10^{-3} V for unitless. The modifications in virtual source velocity, DIBL, and capacitance can accurately describe 2D-material FETs, thereby improving the overall fitting of the device characteristics data.

With a set of optimized parameters, the results from MVS model fit NEGF data points for both current and charge. In Fig. 5.5, the excellent agreement between the simulated (dots) and modeled results (lines) demonstrate an accurate fit for the drain currents (I_D). Notably, in Fig. 5.5(c) and (f), the total charges in the channel are well matched between the simulated and modeled results. This is a critical aspect for the precise prediction of dynamic circuit performance. It is important to have virtual source velocity and charge density well matched separately so that a current fitting with physical meaning requires a good match in both virtual source charge density and velocity between device simulation and compact model. This section will highlight the importance of a good match in charge density and velocity at the virtual source point for successful current curve fitting.

In the previous work [87], it was found that the MVS model could not accurately fit the current-voltage characteristics of GeH FETs from NEGF simulation, particularly at the ON state, even with the best-optimized set of parameters. This discrepancy arises from the lack of physics unique to nanotransistors based on 2D materials. To investigate the correlation between NEGF and the MVS model, it is essential to analyze the electrostatic potential profile of the device. In Fig. 5.6, the channel potentials are plotted along the device length at different gate voltages (from $|V_G| = 0\,\mathrm{V}$ to $|V_G| = 1.2\,\mathrm{V}$ at a fixed drain voltage $|V_D| = 0.7\,\mathrm{V}$). From the plot, we can observe as the potential barriers are gradually modulated downward for NMOS (upward for PMOS) with increasing (decreasing) gate voltage, the top of the barrier moves closer to the source region. The top of each potential (indicated by dots in Fig. 5.6) corresponds to the virtual source point in the MVS model.

In Fig. 5.7, we present charge density and carrier velocity at the virtual source point. The NEGF simulation velocity (Fig. 5.7(b)) was calculated by dividing the drain current with the local charge density (Fig. 5.7(a)) at the virtual-source point, located at the top of the barrier. The charge density and carrier velocity obtained from the MVS model are also depicted in the same plot with solid lines. Our findings reveal a consistent match between NEGF and the modified MVS model, indicating that our new model captures the drain current and generates parameters such as gate capacitance and virtual source velocity with a solid physical interpretation. While the NMOS results are presented as an example, we applied the same approach to PMOS.

The agreement in charge and velocity is a result of the updated equation of virtual source velocity and modulated gate capacitance in the modified MVS model, as discussed in the following sections. These modifications are derived from two unique phenomena arising from the properties of 2D materials that were uncovered through NEGF simulation. To gain a deeper understanding of the physics underlying charge and velocity in 2D FETs, we discuss the density of states (DOS) and injection velocity given by the GeH band structure, as well as the effect of quantum capacitance, in sections 5.4.1 and 5.4.2, respectively.

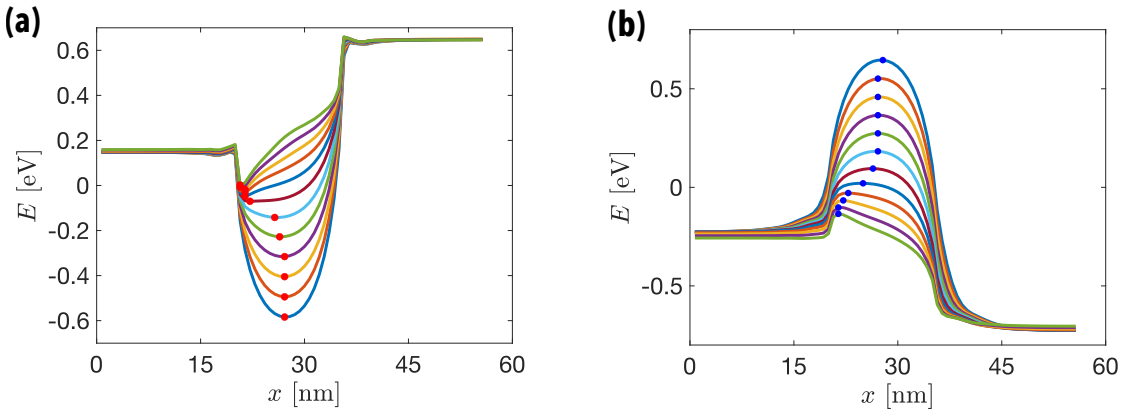


Figure 5.6: Potential profiles of (a) the valence band for PMOS at $V_D = -0.7\,\mathrm{V}$ and (b) the conduction band for NMOS at $V_D = 0.7\,\mathrm{V}$, with different gate voltages. The location of the top of the barrier for each gate voltage is indicated with red and blue dots for PMOS and NMOS, respectively.

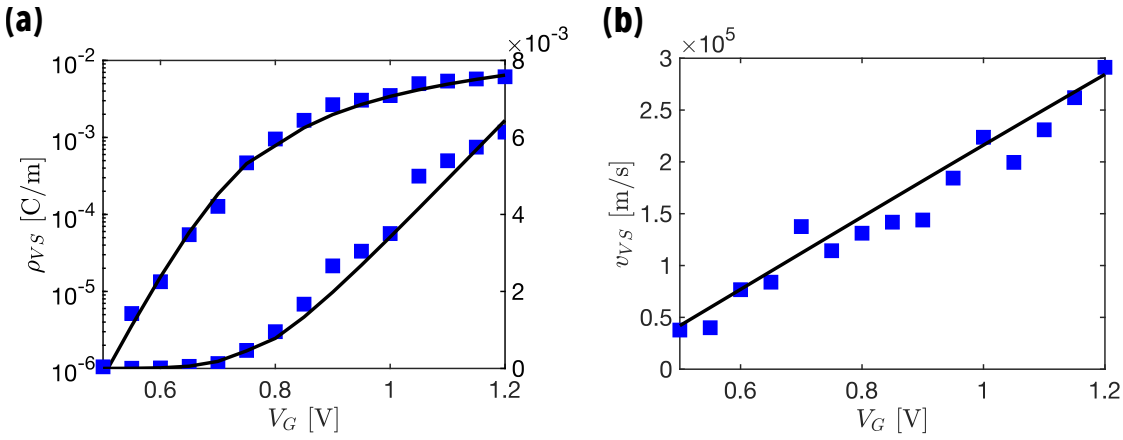


Figure 5.7: (a) Charge density (ρ_{vs}) and (b) virtual-source velocity (v_{vs}) at the virtual source/top of barrier point are plotted as a function of gate voltage, respectively, at $V_D = 0.7\,\mathrm{V}$. Charge densities are plotted on both a linear scale and a logarithm scale. NEGF simulation results are shown in blue square dots, while results from the MVS model are in black solid curves.

5.4.1 DOS and Injection Velocity

One of the unique behaviors of GeH MOSFETs is that the carrier velocity at the virtual source point increases with gate voltage in the ON state, as shown in Fig. 5.7(b). This contrasts with 3D silicon, for which the carrier velocity at the virtual source point is assumed to be relatively independent of gate voltage in the MVS model. The reason for this behavior is that 2D GeH has a significantly lower density of states (DOS) than 3D silicon. Figure 5.8(a) displays the calculated DOS for GeH and silicon, confirming that GeH has a significantly lower DOS than silicon. For example, the DOS at an energy level of 0.1 eV above the conduction band bottom (E_C) is estimated to be $3.04 \times 10^{20}\,\mathrm{cm^{-3}eV^{-1}}$ for GeH and $2.41 \times 10^{21}\,\mathrm{cm^{-3}eV^{-1}}$ for silicon.

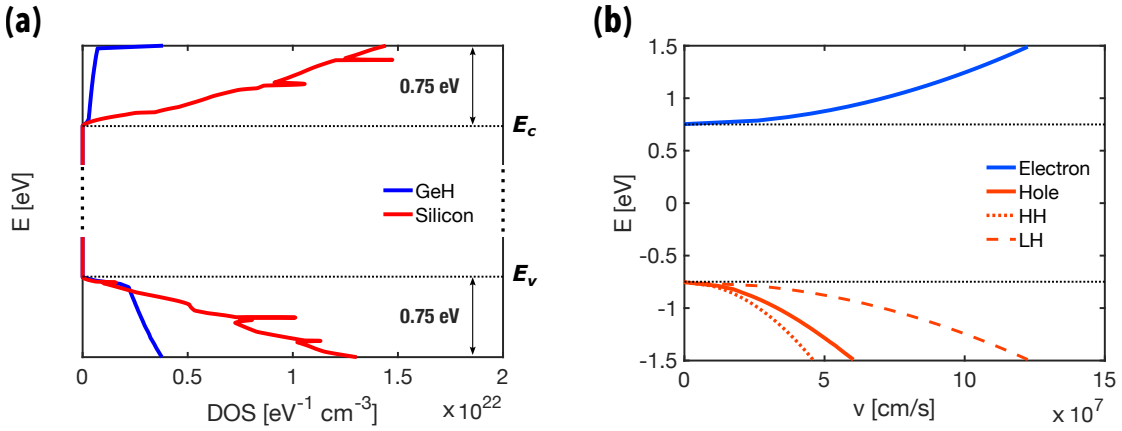


Figure 5.8: (a) Energy-resolved density of state (DOS) for GeH and silicon. For comparison, the unit of DOS for GeH was converted to $\mathrm{eV^{-1}cm^{-3}}$ by dividing 2D DOS of GeH with its thickness (0.55 nm). (b) The bandstructure-limited velocity of GeH is plotted for both electrons (blue solid line) and holes (red solid line). The red dashed line shows the velocity of light holes and that of heavy holes by the red dotted line.

In the MOSFETs based on low-DOS materials like GeH, the channel potential can be more easily lowered as the higher gate voltage is applied, allowing gate capacitance value energy to range from the edge of the conduction/valence bands. As a result, higher energy levels with respect to E_C for NMOS (or lower energy states with respect to the valence band top E_V for PMOS) are involved. Due to the large injection velocity at higher (lower) energy levels for NMOS (PMOS), as shown in Fig. 5.8(b), the average carrier velocity (within the transmission energy window) of 2D GeH FETs becomes higher when a large gate voltage is applied. In contrast, for silicon, the majority of current flows near E_C for NMOS (E_V for PMOS), making the average carrier velocity almost constant.

5.4.2 Quantum Capacitance

The improvement of charge fitting is achieved by altering the value of the gate capacitance in the model. Capacitance extracted from NEGF simulation is not limited by the oxide capacitance, which is usually adopted as the gate capacitance for conventional silicon devices, showing a small deviation from the oxide capacitance (C_{ox}) at the ON state. This deviation from C_{ox} is due to the effect of quantum capacitance (C_Q), an essential phenomenon observed in 2D semiconductors, such as graphene.

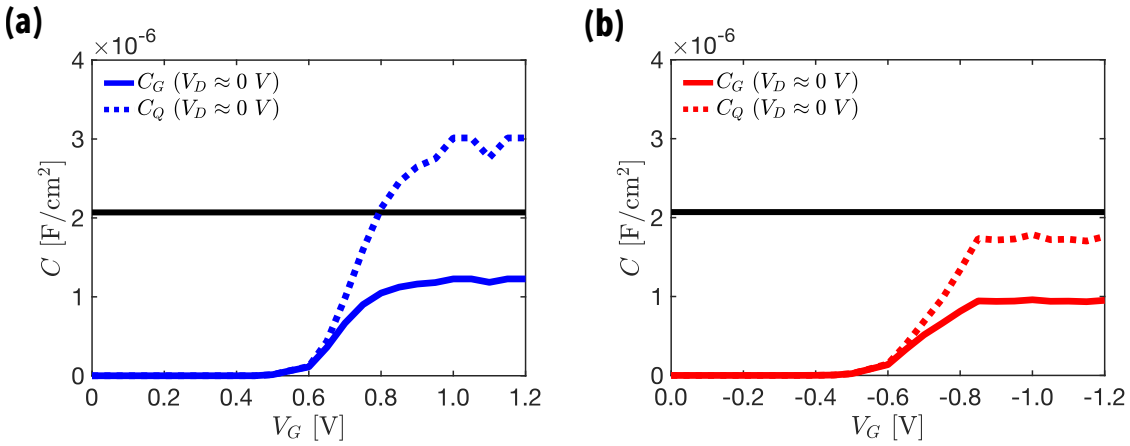


Figure 5.9: Oxide capacitance (C_{ox}, black solid line), quantum capacitance (C_Q, colored solid line), and total gate capacitance (C_G, dash-dot line) for (a) NMOS and (b) PMOS at $V_D \approx 0V$.

We extract C_Q from numerical simulations at $V_{DS} \approx 0\,\mathrm{V}$. First, intrinsic gate capacitance is computed numerically using

$$C_G = \frac{\partial Q_{Ch}}{\partial V_{GS}},$$

where Q_{Ch} is the steady-state channel charge. Then, the quantum capacitance can be calculated with

$$C_Q = \frac{C_{ox} C_G}{C_{ox} - C_G}$$

according to the Eq. (5.19) in section 5.4. The capacitance (C_G, C_Q, and C_{ox}) are plotted in Fig. 5.9 for both NMOS and PMOS. In this study, C_{ox} is calculated to be $2.1 \times 10^{-6}\,\mathrm{F/cm}$ due to a 3.85-nm-thick Al_2O_3 gate dielectric. For NMOS, the total gate capacitance C_G extracted from NEGF is around $1.2 \times 10^{-6}\,\mathrm{F/cm}$, which leads to $C_Q = 3.0 \times 10^{-6}\,\mathrm{F/cm}$ at the ON state. Similarly, for PMOS, C_G is saturated around $1.0 \times 10^{-6}\,\mathrm{F/cm}$, and C_Q is

1.7×10^{-6} F/cm at high (negative) gate voltages. As a result, the gate capacitance of GeH FET is a strong function of quantum capacitance due to its low density of state from the 2D effect, whereas the gate capacitance of conventional silicon MOSFET can be approximated with the oxide capacitance because of the large DOS. For our modified MVS model, the gate capacitances are evaluated by fitting the charge density at the virtual source point, as illustrated in Fig. 4(a), resulting in the gate capacitance $C_G = 1.5 \times 10^{-6}\,\mathrm{F/cm^2}$ for NMOS and $C_G = 1.7 \times 10^{-6}\,\mathrm{F/cm^2}$ for PMOS, both of which fall close to the value from NEGF.

5.4.3 Fitting Parameters

Table 5.1: Fitting Parameters for NMOS and PMOS.

Fitting Parameters	**15-nm NMOS**	**15-nm PMOS**
Contact resistance R_s [k$\Omega \cdot$ µm]	0	0
Gate capacitance C_G [F/cm^{-2}]	1.53×10^{-7}	1.70×10^{-7}
VS velocity v_{vs} [cm/s]	3.50×10^{7}	1.75×10^{7}
Mobility μ [cm^2/(V $\cdot$ s)]	74	30
Threshold voltage V_{TH} [V]	0.78	0.76
DIBL δ [mV/V]	17	11
Punch-through factor n_d [1/V]	0.135	0.292

Key parameters are listed in Table 5.1 for intrinsic transistors with $R_s = 0$. The virtual source velocities of NMOS and PMOS are 9.23×10^7 cm/s and 4.09×10^7 cm/s, respectively, which match the band structure-limited velocity as shown in Fig. 5.8(b). Other parameters, such as subthreshold swing and threshold voltage, also closely agree with the values observed in the NEGF transport simulation. In addition, the carrier mobility used in the compact model for PMOS ($= 30\,\mathrm{cm^2/(V \cdot s)}$) is lower than that of NMOS ($= 74\,\mathrm{cm^2/(V \cdot s)}$) which is supported by the band structure of GeH where heavier effective mass can be observed for PMOS. All parameters were adopted from the original MVS model with careful consideration to minimize the size of the parameter set. To give users more flexibility, the MVS model chooses these directly as input instead of calculating through numerical simulation or experimental measurement. Most of the fitting parameters in our compact model have clear physical interpretations, feeding the circuit simulation with reliable input parameters. When we simulate circuit performance, external/series resistance, and parasitic capacitance will be incorporated.

5.5 HSPICE Simulation

Circuit simulation can be conducted through HSPICE using a developed modified MVS model with fitted parameters. A 15-stage FO4 CMOS inverter chain is used for a benchmark circuit. Circuit parameters such as delay, total charge, and leakage charge are extracted at a certain circuit node in one clock period. Circuit-level figures of merit such as propagation delay (τ_p), static power dissipation (P_{static}) and dynamic power dissipation ($P_{dynamic}$) are extracted as:

$$\tau_p = N_d \tau \tag{5.22}$$

$$P_{static} = f N_d Q_{leak} V_{DD} \tag{5.23}$$

$$P_{dynamic} = \alpha f N_d (Q_{tot} - Q_{leak}) V_{DD} \tag{5.24}$$

where α is an activity factor and N_d is the logic depth. τ is measured as the average delay per stage. Q_{tot} and Q_{leak} are the total charge supplied by the power supply voltage under a pulsing input and a static input tied at the ground, respectively. f is the input pulse frequency. Energy-delay product (EDP), as a figure of merit for circuit performance, can be optimized by carefully engineering power supply voltage (V_{DD}) and threshold voltage (V_{TH}). A lumped sum resistances of $1\,\mathrm{k\Omega\,\mu m}$ and $0.1\,\mathrm{k\Omega\,\mu m}$ are used for contact resistance in the circuit simulation.

To evaluate technology and optimize device-circuit performance, HSPICE simulations are conducted using the modified MVS model that was developed in this work for the compact modeling of GeH FETs. A 15-stage FO4 inverter chain was selected as a benchmark digital circuit, which was designed to switch at the maximum possible frequency. An activity factor of 10^{-2} was used, as detailed in previous work, to balance the ratio of dynamic power and static power at a reference point ($V_{DD} = 1.0\,\mathrm{V}$ and $V_{TH} = 0.8\,\mathrm{V}$). To investigate the effect of contact resistance on circuit performance, two different contact resistances of $1\,\mathrm{k\Omega\,\mu m}$ and $0.1\,\mathrm{k\Omega\,\mu m}$ were employed, as opposed to the previous assumption of perfect contact at the metal and source/drain interface of GeH FET. The simulation is assumed at temperature of 300 K with MOSFETs of $L_{ch} = 18\,\mathrm{nm}$, considering the effect of varying contact resistance will be most pronounced with short-channel devices. Circuit parameters including propagation delay, total power dissipation (P_T), and energy-delay product (EDP) were evaluated while varying V_{DD} and V_{TH}.

Figure 5.10 is used to illustrate the relation of different reference voltages (V_{DD}, V_{TH}, V_{ON}, and V_{OFF}) in the operation window. To change V_{DD} at a fixed V_{TH}, the right edge of the grey window (V_{ON}) is shifted, while the position of the left edge (V_{OFF}) remains

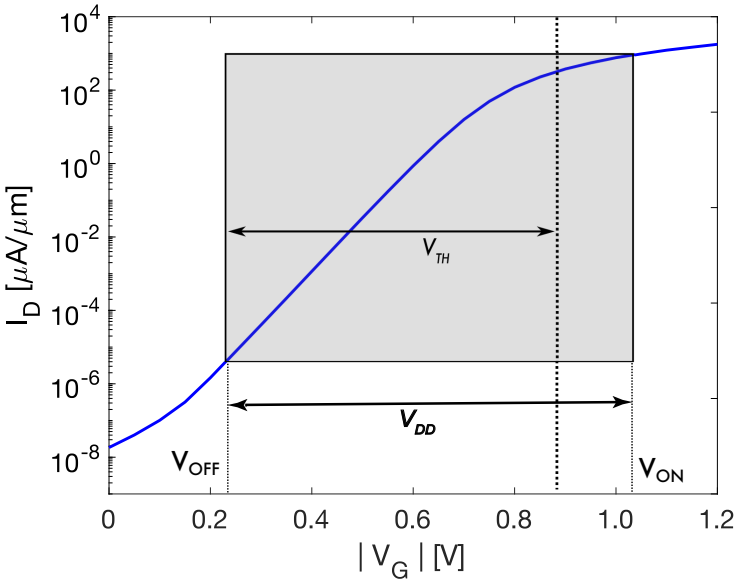


Figure 5.10: A power supply voltage window (grey box) is plotted to illustrate the relation of the reference voltages at which circuit parameters, such as propagation delay, power dissipation, and energy-delay product, are evaluated.

fixed. On the other hand, to vary V_{TH} under a fixed V_{DD}, both V_{ON} and V_{OFF} are shifted simultaneously while maintaining the width of the grey window (V_{ON} – V_{OFF}) in Figure 5.10.

The impact of V_{DD} and V_{TH} on P_T and τ_p was shown in Fig. 5.11(a) and (b), respectively, for both 1.0 kΩ µm and 0.1 kΩ µm contact resistances, illustrating a trade-off between the speed and the energy cost. EDP as a function of threshold voltage was plotted in Fig. 5.11(c) and (d), showing an optimal operation point for each power supply voltage. By sweeping both V_{DD} and V_{TH}, the minimum EDP can be found at the peak value in the $1/EDP$ color map in Fig. 5.11(e) and (f). In the HSPICE simulation, V_{TH} shift was performed identically for NMOS and PMOS to improve the computational efficiency.

In Fig. 5.11(e), with $R_c = 1.0\,\text{k}\Omega\,\mu\text{m}$, the minimum EDP of 9.8×10^{-3} fJ · ps was achieved at $V_{DD} = 0.8$ V and $V_{TH} = 0.5$ V, which is approximately 50 times lower than the EDP for a non-optimized device (e.g., with $V_{DD} = 0.7$ V and $V_{TH} = 0.3$ V). Similarly, the trend of circuit parameters with $R_c = 0.1\,\text{k}\Omega\,\mu\text{m}$ is also shown in Fig. 5.11(f). In this case, the optimal operation point shifted to $V_{DD} = 0.9$ V and $V_{TH} = 0.5$ V, indicating that lower V_{DD} are required to achieve the minimum EDP of 3.1×10^{-3} fJ · ps. This shift can be attributed to the steeper τ_p curve and transition point moving toward higher V_{TH} with lower contact

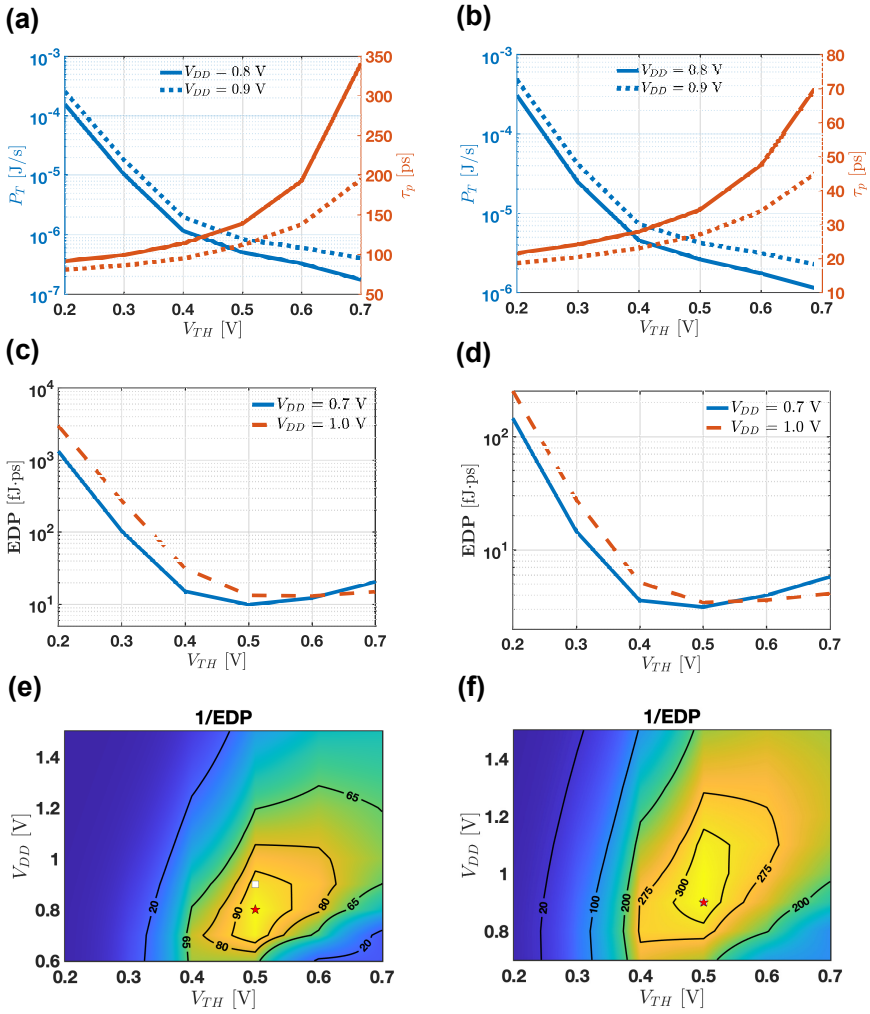


Figure 5.11: (a) Total power dissipation (P_T), propagation delay (τ_p) and (c) energy-delay product (EDP) vs. threshold voltage (V_{TH}) at different power supply voltages of $V_{DD} = 0.7\,\mathrm{V}$ and $1.0\,\mathrm{V}$ and (e) Color map of 1/EDP as a function of V_{DD} and V_{TH} are plotted with a higher contact resistance ($R_c = 1\,\mathrm{k\Omega \cdot \mu m}$). The corresponding results with a lower contact resistance $R_c = 0.1\,\mathrm{k\Omega \cdot \mu m}$ are plotted in (b), (d), and (f). In colormap, a warmer color indicates a smaller EDP. The optimal operation point (i.e, the minimum EDP point) is located at $V_{DD} = 0.8\,\mathrm{V}$ and $V_{TH} = 0.5\,\mathrm{V}$ with $EDP = 9.8 \times 10^{-3}\,\mathrm{fJ \cdot ps}$ for $1\,\mathrm{k\Omega \cdot \mu m}$ in (e) and at $V_{DD} = 0.9\,\mathrm{V}$ and $V_{TH} = 0.5\,\mathrm{V}$ with $EDP = 3.1 \times 10^{-3}\,\mathrm{fJ \cdot ps}$ for $0.1\,\mathrm{k\Omega \cdot \mu m}$ in (f).

resistance ($0.1\,\mathrm{k\Omega\,\mu m}$) while the P_T curve becomes flatter. The impact of delay and power behavior should be considered to understand the shift of the minimum EDP point.

5.6 Summary

To investigate the performance of 2D-based devices, a multi-level simulation was proposed in this study. Our study has significantly enhanced the multi-level simulation framework proposed initially in prior research to evaluate the performance of 2D-based devices. This comprehensive approach began with a quantum transport simulation to accurately model the intrinsic characteristics of GeH FETs. Subsequently, we developed a compact model tailored to match the numerical data obtained from the quantum transport simulation. Finally, HSPICE circuit simulations are conducted to optimize the circuit performance of GeH FETs, with the modified MVS model playing a pivotal role in seamlessly bridging the device and circuit simulations.

The numerical simulation results showed that GeH NMOS exhibits larger ON current and higher carrier velocity, while carrier density is more elevated with PMOS. The former is due to a smaller effective mass of electrons, and the latter is because of the larger DOS of holes. Unique characteristics of GeH MOSFET were incorporated in the compact model, which includes a voltage-dependent virtual source velocity and quantum capacitance to account for the unique nature of 2D materials. In this work, NMOS and PMOS are separately simulated and analyzed, which is essential in CMOS circuit-level design, simulation, and optimization. The effect of contact resistance at the source/drain for GeH MOSFET was also evaluated. It was found that the minimum EDP point shifts to a lower power supply voltage and higher threshold voltage with more considerable contact resistance, and the optimal region becomes significantly smaller. Although the channel length and EOT scaling were left for future work, the authors are confident that the modified MVS model will scale well for different technologies, as the original MVS has shown excellent scaling. While our primary focus has been GeH MOSFETs, the models elucidated in this study can be applied to ballistic MOSFETs made from other 2-D semiconductors with minimal adjustments.

Overall, this work focuses on the multi-level simulation framework and highlights the advantages of utilizing NEGF simulation and the MVS model in a combined approach. Notably, the outcomes predicted by the modified MVS model demonstrated exceptional agreement with NEGF simulation results. Our findings collectively underscore the robustness and accuracy of the enhanced simulation framework, making it a valuable tool for further advancements in 2D-based device research.

Chapter 6

Conclusion

6.1 Summary

The presented thesis has mainly focused on two parts: 1) specifically in developing a more accurate numerical simulation of MOSFETs based on 2D materials operating at cryogenic temperatures and 2) bridging the gap between intrinsic simulation and circuit analysis of 2D-based MOSFETs.

At the core of this work is the establishment of an improved simulation framework that leverages the NEGF formalism. This framework accounts for critical temperature-dependent phenomena such as electrostatics, scattering, dissipation effects, carrier freeze-out, and band tail phenomena, offering an in-depth understanding of 2D materials in cryogenic conditions. The successful alignment with experimental data, especially for HfS_2 MOSFETs, highlights the framework's accuracy and potential in enhancing the performance of cryogenic electronic devices. In-depth explorations into the cryogenic behavior of MOSFETs, specifically addressing phenomena like the Fermi distribution function, incomplete ionization, interface trapping, and the band tail effect, were conducted. These investigations revealed the complex interplay between decreased temperature and semiconductor physics, providing a comprehensive framework for optimizing device performance under cryogenic conditions. The thesis systematically dissected the behavior of HfS_2 MOSFETs at cryogenic temperatures, covering aspects from scattering and mobility to ionization and channel length scaling. This rigorous analysis not only deepens the understanding of HfS_2 MOSFETs in low-temperature environments but also contributes to the design and optimization of devices for quantum computing applications.

Another important part is the development of a multi-level simulation approach tailored for 2D material-based nanoelectronics. Through an exhaustive case study on GeH MOSFETs, this approach integrates device simulation, physics-based compact modeling, and circuit benchmarking. The revision of the virtual source model to encapsulate the unique characteristics of 2D-material FETs and the subsequent HSPICE simulations underscore the significant improvements in the energy-delay product, facilitating the design of efficient next-generation nanoelectronic circuits. Further, the work extends to the performance analysis and optimization of GeH FETs through a multi-level simulation framework. The introduction of a modified virtual source model, accounting for the unique properties of 2D materials, alongside detailed device and circuit simulations, demonstrates the potential of GeH MOSFETs in CMOS technology.

6.2 Outlook

The advancements and insights gleaned from this thesis are instrumental in furthering the development and optimization of two-dimensional (2D) material-based nanoelectronics, especially under cryogenic conditions and at the circuit level. The potential application of cryogenic MOSFETs in quantum computing opens up exciting avenues for future research, particularly in integrating these devices within quantum circuits. A pivotal direction for future endeavors includes incorporating cryogenic simulations into the multi-level simulation framework to examine the circuit-level performance of 2D material-based transistors at cryogenic temperatures. Developing a cryogenic-physics-based compact model within the multi-level simulation framework is essential for advancing this research. Additionally, efforts could be directed toward creating hybrid systems that combine 2D material-based devices with qubits, thereby optimizing the interface and communication between classical and quantum computing components. Such integration is crucial for achieving scalable quantum computers.

Although this thesis has predominantly focused on germanene (GeH) and hafnium disulfide (HfS_2) for MOSFET applications, the exploration of 2D materials is far from exhaustive. Future investigations could delve into other 2D materials exhibiting unique properties, such as transition metal dichalcogenides (TMDs) beyond HfS_2, black phosphorus, and graphene. Identifying and characterizing materials with enhanced electronic, thermal, and mechanical properties is vital for improving device performance, particularly in high-stability and efficiency-demanding applications. Furthermore, the cryogenic simulation framework could be leveraged to explore advanced device architectures, including tunnel field-effect transistors (TFETs), gate-all-around (GAA) transistors, and stacked 2D heterostructures.

These innovative architectures hold the promise of mitigating short-channel effects, augmenting electrostatic control, and reducing power consumption, which are imperative for miniaturizing device dimensions.

While significant progress has been made in simulation and theoretical analysis, moving forward to experimental validation and prototyping of devices based on the developed models is paramount. Collaborating with experimental research groups and industry partners will be crucial for prototype fabrication, facilitating the empirical testing of theoretical predictions. Such collaborative efforts will not only serve to validate the developed models but also unearth new challenges and avenues for optimization.

6.3 Future Work

This thesis lays a foundational framework for the advancement of 2D material-based nanoelectronics, emphasizing their potential in cryogenic applications. To extend the research frontier, the following detailed avenues for future work are proposed, aimed at addressing critical challenges and leveraging novel opportunities:

Deep Cryogenic Environment Exploration

- *Development of Enhanced Simulation Stability:* Creation of advanced simulation techniques that demonstrate increased stability at ultra-low temperatures, beyond the cryogenic range in this thesis (50 K). This entails the formulation of novel numerical methods or the adaptation of existing ones to ensure accurate, reliable simulations down to millikelvin temperatures. Investigating the quantum effects that become significant at these temperatures will also be crucial for a comprehensive understanding of 2D material behavior in deep cryogenic conditions.

Cryogenic Compact Modeling and Quantum Circuit Integration

- *Physics-Based Compact Models for Cryogenic Conditions:* Focus on developing and refining compact models that encapsulate the unique quantum phenomena observed in 2D materials at cryogenic temperatures. Such models are essential for the predictive simulation of device and circuit behavior, enabling the design of optimized circuits that exploit the quantum mechanical properties of 2D materials for enhanced performance.

- *Seamless Integration with Quantum Circuits:* Undertake targeted research to seamlessly integrate cryogenic MOSFETs into quantum circuits, emphasizing the optimization of interfaces for efficient signal transfer between classical and quantum domains. This research will explore novel methodologies for qubit control and readout facilitated by 2D material-based devices, potentially setting new benchmarks for quantum computing architecture.

Multi-Level Simulation Framework Enhancement

- *Cryogenic Simulation Capabilities:* Augment the existing multi-level simulation framework to incorporate detailed cryogenic simulation functionalities. This enhancement will support the end-to-end design and evaluation of 2D nanoelectronic circuits tailored for operation in cryogenic environments, addressing the need for devices that are both highly efficient and compact.

- *Benchmarking and Material Characterization:* Embark on comprehensive benchmarking studies to contrast the performance of cryogenic 2D circuits against conventional technologies, emphasizing their energy efficiency and suitability for specific applications. Concurrently, extend the simulation framework to include a broader spectrum of 2D materials, focusing on their characterization at cryogenic temperatures to uncover materials with superior electronic, thermal, and mechanical properties.

- *Exploration of Advanced Device Architectures:* Investigate cutting-edge device architectures such as TFETs, GAA transistors, and stacked 2D heterostructures for their operational viability and performance at cryogenic temperatures. This exploration aims to identify design strategies that minimize power consumption and mitigate short-channel effects, essential for the continued miniaturization of electronic devices.

Bridging Theory with Experimental Practice

- *Collaborative Development and Empirical Testing:* Establish strong partnerships with experimental research groups and industry stakeholders to transition from theoretical models and simulations to the fabrication and testing of real-world devices. These collaborative efforts are vital for validating the practical applicability of research findings and for refining theoretical models based on empirical data. Such synergies will facilitate the rapid advancement of 2D material-based electronics, steering the field towards the realization of next-generation computing technologies.

References

[1] S. J. Pauka, K. Das, R. Kalra, A. Moini, Y. Yang, M. Trainer, A. Bousquet, C. Cantaloube, N. Dick, G. C. Gardner, M. J. Manfra, and D. J. Reilly. A cryogenic CMOS chip for generating control signals for multiple qubits. *Nature Electronics*, 4(1):64–70, January 2021. Number: 1 Publisher: Nature Publishing Group.

[2] Avid Kamgar. Subthreshold behavior of silicon MOSFETs at 4.2 K. *Solid-State Electronics*, 25(7):537–539, July 1982.

[3] Arnout Beckers, Farzan Jazaeri, and Christian Enz. Cryogenic MOS Transistor Model. *IEEE Transactions on Electron Devices*, 65(9):3617–3625, September 2018.

[4] Jing Wang. *Device physics and simulation of silicon nanowire transistors*. PhD thesis, Purdue University, 2005.

[5] Zhibin Ren. *NANOSCALE MOSFETS: PHYSICS, SIMULATION AND DESIGN*. PhD thesis, Purdue University, 2001.

[6] Demin Yin. *Theoretical Investigation of Contact Effects on the Performance of 2D-Material Nanotransistors*. PhD thesis, University of Waterloo, 2021.

[7] Supriyo Datta. Nanoscale device modeling: the Green's function method. *Superlattices and Microstructures*, 28(4):253–278, October 2000.

[8] Kai Miao. *QUANTUM THERMAL TRANSPORT IN SEMICONDUCTOR NANOSTRUCTURES*. PhD thesis, Purdue University, 2017.

[9] Eva Pavarini. 6 Crystal-Field Theory , Tight-Binding Method and Jahn-Teller Effect. 2012.

[10] AbdulAziz Almutairi, Yiju Zhao, Demin Yin, and Youngki Yoon. Performance Limit Projection of Germanane Field-Effect Transistors. *IEEE Electron Device Letters*, 38(5):673–676, May 2017.

[11] Francesco Bova, Avi Goldfarb, and Roger G. Melko. Commercial applications of quantum computing. *EPJ Quantum Technology*, 8(1):2, December 2021.

[12] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 1 edition, June 2012.

[13] Nathalie P. De Leon, Kohei M. Itoh, Dohun Kim, Karan K. Mehta, Tracy E. Northup, Hanhee Paik, B. S. Palmer, N. Samarth, Sorawis Sangtawesin, and D. W. Steuerman. Materials challenges and opportunities for quantum computing hardware. *Science*, 372(6539):eabb2823, April 2021.

[14] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[15] Jay M. Gambetta, Jerry M. Chow, and Matthias Steffen. Building logical qubits in a superconducting quantum computing system. *npj Quantum Information*, 3(1):2, January 2017.

[16] John Clarke and Frank K. Wilhelm. Superconducting quantum bits. *Nature*, 453(7198):1031–1042, June 2008.

[17] V. M. Schäfer, C. J. Ballance, K. Thirumalai, L. J. Stephenson, T. G. Ballance, A. M. Steane, and D. M. Lucas. Fast quantum logic gates with trapped-ion qubits. *Nature*, 555(7694):75–78, March 2018.

[18] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, January 2001.

[19] Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, J. Pablo Bonilla Ataides, Nishad Maskara, Iris Cong, Xun Gao, Pedro Sales Rodriguez, Thomas Karolyshyn, Giulia Semeghini, Michael J. Gullans, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, February 2024.

[20] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, January 2003.

[21] Bishnu Patra, Rosario M. Incandela, Jeroen P. G. van Dijk, Harald A. R. Homulle, Lin Song, Mina Shahmohammadi, Robert Bogdan Staszewski, Andrei Vladimirescu, Masoud Babaie, Fabio Sebastiano, and Edoardo Charbon. Cryo-CMOS Circuits and Systems for Quantum Computing Applications. *IEEE Journal of Solid-State Circuits*, 53(1):309–321, January 2018.

[22] E. Charbon, F. Sebastiano, A. Vladimirescu, H. Homulle, S. Visser, L. Song, and R M. Incandela. Cryo-CMOS for quantum computing. In *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 13.5.1–13.5.4, San Francisco, CA, USA, December 2016. IEEE.

[23] Ruben Asanovski, Alexander Grill, Jacopo Franco, Pierpaolo Palestri, Arnout Beckers, Ben Kaczer, and Luca Selmi. Understanding the Excess 1/f Noise in MOSFETs at Cryogenic Temperatures. *IEEE Transactions on Electron Devices*, 70(4):2135–2141, April 2023.

[24] Youngki Yoon, Kartik Ganapathi, and Sayeef Salahuddin. How Good Can Monolayer MoS2 Transistors Be? *Nano Letters*, 11(9):3768–3773, September 2011.

[25] Likai Li, Yijun Yu, Guo Jun Ye, Qingqin Ge, Xuedong Ou, Hua Wu, Donglai Feng, Xian Hui Chen, and Yuanbo Zhang. Black phosphorus field-effect transistors. *Nature Nanotechnology*, 9(5):372–377, May 2014.

[26] Yiju Zhao, AbdulAziz AlMutairi, and Youngki Yoon. Assessment of Germanane Field-Effect Transistors for CMOS Technology. *IEEE Electron Device Letters*, 38(12):1743–1746, December 2017.

[27] Yiju Zhao, Demin Yin, and Youngki Yoon. Intrinsic Performance of Germanane Schottky Barrier Field-Effect Transistors. *IEEE Transactions on Electron Devices*, 65(10):4188–4195, October 2018.

[28] Rubaya Absar, Hazem Elgabra, Dylan Ma, Yiju Zhao, and Lan Wei. Cryogenic CMOS for Quantum Computing. In *Design and Applications of Emerging Computer Systems*, pages 591–621. Springer Cham, 2024.

[29] Xiaolong Liu and Mark C. Hersam. 2D materials for quantum information science. *Nature Reviews Materials*, 4(10):669–684, October 2019.

[30] Xiaohe Huang, Chunsen Liu, and Peng Zhou. 2D semiconductors for specific electronic applications: from device to system. *npj 2D Materials and Applications*, 6(1):51, August 2022.

[31] Saptarshi Das, Amritanand Sebastian, Eric Pop, Connor J. McClellan, Aaron D. Franklin, Tibor Grasser, Theresia Knobloch, Yury Illarionov, Ashish V. Penumatcha, Joerg Appenzeller, Zhihong Chen, Wenjuan Zhu, Inge Asselberghs, Lain-Jong Li, Uygar E. Avci, Navakanta Bhat, Thomas D. Anthopoulos, and Rajendra Singh. Transistors based on two-dimensional materials for future integrated circuits. *Nature Electronics*, 4(11):786–799, November 2021.

[32] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis. Single-layer MoS2 transistors. *Nature Nanotechnology*, 6(3):147–150, March 2011.

[33] M. M. Atalla, E. Tannenbaum, and E. J. Scheibner. Stabilization of Silicon Surfaces by Thermally Grown Oxides. *Bell System Technical Journal*, 38(3):749–783, May 1959.

[34] C. Auth, A. Aliyarukunju, M. Asoro, D. Bergstrom, V. Bhagwat, J. Birdsall, N. Bisnik, M. Buehler, V. Chikarmane, G. Ding, Q. Fu, H. Gomez, W. Han, D. Hanken, M. Haran, M. Hattendorf, R. Heussner, H. Hiramatsu, B. Ho, S. Jaloviar, I. Jin, S. Joshi, S. Kirby, S. Kosaraju, H. Kothari, G. Leatherman, K. Lee, J. Leib, A. Madhavan, K. Marla, H. Meyer, T. Mule, C. Parker, S. Parthasarathy, C. Pelto, L. Pipes, I. Post, M. Prince, A. Rahman, S. Rajamani, A. Saha, J. Dacuna Santos, M. Sharma, V. Sharma, J. Shin, P. Sinha, P. Smith, M. Sprinkle, A. St. Amour, C. Staus, R. Suri, D. Towner, A. Tripathi, A. Tura, C. Ward, and A. Yeoh. A 10nm high performance and low-power CMOS technology featuring 3 rd generation FinFET transistors, Self-Aligned Quad Patterning, contact over active gate and cobalt local interconnects. In *2017 IEEE International Electron Devices Meeting (IEDM)*, pages 29.1.1–29.1.4, San Francisco, CA, USA, December 2017. IEEE.

[35] Shamiul Alam, Md Shafayat Hossain, Srivatsa Rangachar Srinivasa, and Ahmedullah Aziz. Cryogenic memory technologies. *Nature Electronics*, 6(3):185–198, March 2023.

[36] Farzan Jazaeri, Arnout Beckers, Armin Tajalli, and Jean-Michel Sallese. A Review on Quantum Computing: From Qubits to Front-end Electronics and Cryogenic MOSFET Physics. In *2019 MIXDES - 26th International Conference "Mixed Design of Integrated Circuits and Systems"*, pages 15–25, Rzeszów, Poland, June 2019. IEEE.

[37] S. Bonen, U. Alakusu, Y. Duan, M. J. Gong, M. S. Dadash, L. Lucci, D. R. Daughton, G. C. Adam, S. Iordanescu, M. Pasteanu, I. Giangu, H. Jia, L. E. Gutierrez, W. T. Chen, N. Messaoudi, D. Harame, A. Muller, R. R. Mansour, P. Asbeck, and S. P. Voinigescu. Cryogenic Characterization of 22nm FDSOI CMOS Technology for Quantum Computing ICs. *IEEE Electron Device Letters*, pages 127–130, 2019.

[38] R.L. Patterson, A. Hammoud, J.E. Dickman, S. Gerber, M. Elbuluk, and E. Overton. Electronics for deep space cryogenic applications. In *Proceedings of the 5th European Workshop on Low Temperature Electronics*, pages 207–210, Grenoble, France, 2002. IEEE.

[39] K.-H. Kao, C. Godfrin, A. Elsayed, R. Li, E. Simoen, A. Grill, S. Kubicek, I. P. Radu, and B. Govoreanu. Linking Room- and Low-Temperature Electrical Performance of MOS Gate Stacks for Cryogenic Applications. *IEEE Electron Device Letters*, 43(5):674–677, May 2022.

[40] Aykut Kabaoglu, Nergiz Sahin Solmaz, Sadik Ilik, Yasemin Uzun, and Mustafa Berke Yelten. Statistical MOSFET Modeling Methodology for Cryogenic Conditions. *IEEE Transactions on Electron Devices*, 66(1):66–72, January 2019.

[41] Chao Luo, Zhen Li, Teng-Teng Lu, Jun Xu, and Guo-Ping Guo. MOSFET characterization and modeling at cryogenic temperatures. *Cryogenics*, 98:12–17, March 2019.

[42] Jens Anders, Masoud Babaie, Joseph C. Bardin, Imran Bashir, Gérard Billiot, Elena Blokhina, Shai Bonen, Edoardo Charbon, John Chiaverini, Isaac L. Chuang, Carsten Degenhardt, Dirk Englund, Lotte Geck, Loïck Le Guevel, Donhee Ham, Ruonan Han, Mohamed I. Ibrahim, Daniel Krüger, Ka Meng Lei, Adrien Morel, Dennis Nielinger, Gaël Pillonnet, Jeremy M. Sage, Fabio Sebastiano, Robert Bogdan Staszewski, Jules Stuart, Andrei Vladimirescu, Patrick Vliex, and Sorin P. Voinigescu. CMOS Integrated Circuits for the Quantum Information Sciences. *IEEE Transactions on Quantum Engineering*, 4:5100230, 2023.

[43] M. Lundstrom. Elementary scattering theory of the Si MOSFET. *IEEE Electron Device Letters*, 18(7):361–363, July 1997.

[44] Suman Sarkar, Kazi Rafsanjani Amin, Ranjan Modak, Amandeep Singh, Subroto Mukerjee, and Aveek Bid. Role of different scattering mechanisms on the temperature dependence of transport in graphene. *Scientific Reports*, 5(1):16772, November 2015.

[45] A. Iribarren, R. Castro-Rodríguez, V. Sosa, and J. L. Peña. Band-tail parameter modeling in semiconductor materials. *Physical Review B*, 58(4):1907–1911, July 1998.

[46] Haojun Zhang, Wei Cao, Jiahao Kang, and Kaustav Banerjee. Effect of band-tails on the subthreshold performance of 2D tunnel-FETs. In *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 30.3.1–30.3.4, San Francisco, CA, USA, December 2016. IEEE.

[47] Arnout Beckers, Farzan Jazaeri, and Christian Enz. Inflection Phenomenon in Cryogenic MOSFET Behavior. *IEEE Transactions on Electron Devices*, 67(3):1357–1360, March 2020.

[48] Ioanna Kriekouki, Félix Beaudoin, Pericles Philippopoulos, Chenyi Zhou, Julien Camirand Lemyre, Sophie Rochette, Salvador Mir, Manuel J. Barragan, Michel Pioro-Ladrière, and Philippe Galy. Interpretation of 28 nm FD-SOI quantum dot transport data taken at 1.4 K using 3D quantum TCAD simulations. *Solid-State Electronics*, 194:108355, August 2022.

[49] Hung-Chi Han, Hung-Li Chiang, Iuliana P. Radu, and Christian Enz. Analytical Modeling of Source-to-Drain Tunneling Current Down to Cryogenic Temperatures. *IEEE Electron Device Letters*, 44(5):717–720, May 2023.

[50] Yiju Zhao, Youngki Yoon, and Lan Wei. 2D Material-Based MVS Model and Circuit Performance Analysis for GeH Field-Effect Transistors. *IEEE Transactions on Nanotechnology*, 22:792–799, 2023.

[51] Youngki Yoon, Giuseppe Fiori, Seokmin Hong, Gianluca Iannaccone, and Jing Guo. Performance comparison of graphene nanoribbon FETs with Schottky contacts and doped reservoirs. *IEEE Transactions on Electron Devices*, 55(9):2314–2323, 2008.

[52] Mathieu Luisier. Atomistic modeling of anharmonic phonon-phonon scattering in nanowires. *Physical Review B*, 86(24):245407, December 2012.

[53] Prasad Sarangapani, Yuanchen Chu, James Charles, Gerhard Klimeck, and Tillmann Kubis. Band-tail Formation and Band-gap Narrowing Driven by Polar Optical Phonons and Charged Impurities in Atomically Resolved III-V Semiconductors and Nanodevices. *Physical Review Applied*, 12(4):044045, October 2019.

[54] Yiju Zhao, Youngki Yoon, and Lan Wei. A Multi-Level Simulation of GeH FETs: From Nanomaterial and Device Characteristics to Circuit Performance Optimization.

In *Proceedings of the 17th ACM International Symposium on Nanoscale Architectures*, pages 1–6, Virtual OR USA, December 2022. ACM.

[55] Elisabeth Bianco, Sheneve Butler, Shishi Jiang, Oscar D. Restrepo, Wolfgang Windl, and Joshua E. Goldberger. Stability and Exfoliation of Germanane: A Germanium Graphane Analogue. *ACS Nano*, 7(5):4414–4421, May 2013.

[56] Saptarshi Das, Richard Gulotty, Anirudha V. Sumant, and Andreas Roelofs. All Two-Dimensional, Flexible, Transparent, and Thinnest Thin Film Transistor. *Nano Letters*, 14(5):2861–2866, May 2014.

[57] Sunkook Kim, Jesse Maassen, Jiyoul Lee, Seung Min Kim, Gyuchull Han, Junyeon Kwon, Seongin Hong, Jozeph Park, Na Liu, Yun Chang Park, Inturu Omkaram, Jong Soo Rhyee, Young Ki Hong, and Youngki Yoon. Interstitial Mo-Assisted Photovoltaic Effect in Multilayer $MoSe_2$ Phototransistors. *Advanced Materials*, page 1705542, January 2018.

[58] Zhihao Yu, Zhun-Yong Ong, Songlin Li, Jian-Bin Xu, Gang Zhang, Yong-Wei Zhang, Yi Shi, and Xinran Wang. Analyzing the Carrier Mobility in Transition-Metal Dichalcogenide MoS $_2$ Field-Effect Transistors. *Advanced Functional Materials*, 27(19):1604093, May 2017.

[59] Zhinan Guo, Si Chen, Zhongzheng Wang, Zhenyu Yang, Fei Liu, Yanhua Xu, Jiahong Wang, Ya Yi, Han Zhang, Lei Liao, Paul K. Chu, and Xue-Feng Yu. Metal-Ion-Modified Black Phosphorus with Enhanced Stability and Transistor Performance. *Advanced Materials*, 29(42):1703811, 2017.

[60] Demin Yin, AbdulAziz AlMutairi, and Youngki Yoon. Assessment of High-Frequency Performance Limit of Black Phosphorus Field-Effect Transistors. *IEEE Transactions on Electron Devices*, 64(7):2984–2991, May 2017.

[61] Heekyeong Park, Gyuchull Han, Sang Woo Lee, Hyungbeen Lee, Seok Hwan Jeong, Muhammad Naqi, AbdulAziz AlMutairi, Young Jun Kim, Joonhyung Lee, Wan-joong Kim, Sunkook Kim, Youngki Yoon, and Geonwook Yoo. Label-Free and Recalibrated Multilayer MoS2 Biosensor for Point-of-Care Diagnostics. *ACS Applied Materials & Interfaces*, 9(50):43490–43497, December 2017.

[62] B. N. Madhushankar, A. Kaverzin, T. Giousis, G. Potsi, D. Gournis, P. Rudolf, G. R. Blake, C. H. van der Wal, and B. J. van Wees. Electronic properties of germanane field-effect transistors. *2D Materials*, 4(2):021009, February 2017.

[63] Juan P. Duarte, Sourabh Khandelwal, Aditya Medury, Chenming Hu, Pragya Kushwaha, Harshit Agarwal, Avirup Dasgupta, and Yogesh S. Chauhan. BSIM-CMG: Standard FinFET compact model for advanced circuit design. In *ESSCIRC Conference 2015 - 41st European Solid-State Circuits Conference (ESSCIRC)*, pages 196–201, 2015.

[64] Ali Khakifirooz, Osama M. Nayfeh, and Dimitri Antoniadis. A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters. *IEEE Transactions on Electron Devices*, 56(8):1674–1680, August 2009.

[65] Yogesh Singh Chauhan, Sriramkumar Venugopalan, Maria-Anna Chalkiadaki, Muhammed Ahosan Ul Karim, Harshit Agarwal, Sourabh Khandelwal, Navid Paydavosi, Juan Pablo Duarte, Christian C. Enz, Ali M. Niknejad, and Chenming Hu. BSIM6: Analog and RF Compact Model for Bulk MOSFET. *IEEE Transactions on Electron Devices*, 61(2):234–244, 2014.

[66] H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. *IEEE Transactions on Electron Devices*, 11(10):455–465, October 1964.

[67] Lan Wei, Omar Mysore, and Dimitri Antoniadis. Virtual-Source-Based Self-Consistent Current and Charge FET Models: From Ballistic to Drift-Diffusion Velocity-Saturation Operation. *IEEE Transactions on Electron Devices*, 59(5):1263–1271, May 2012.

[68] Wei Cao, Jiahao Kang, Wei Liu, and Kaustav Banerjee. A Compact Current–Voltage Model for 2D Semiconductor Based Field-Effect Transistors Considering Interface Traps, Mobility Degradation, and Inefficient Doping Effect. *IEEE Transactions on Electron Devices*, 61(12):4282–4290, December 2014.

[69] Biswapriyo Das and Santanu Mahapatra. An atom-to-circuit modeling approach to all-2D metal–insulator–semiconductor field-effect transistors. *npj 2D Materials and Applications*, 2(1):1–10, August 2018.

[70] Y. Jiang, N. Cucu Laurenciu, and S.D. Cotofana. Non-Equilibrium Green Function-based Verilog-A Graphene Nanoribbon Model. In *2018 IEEE 18th International Conference on Nanotechnology (IEEE-NANO)*, pages 1–4, July 2018.

[71] Raphaël J. Prentki, Fei Liu, and Hong Guo. Modeling of Ballistic Monolayer Black Phosphorus MOSFETs. *IEEE Transactions on Electron Devices*, 66(8):3668–3674, August 2019.

[72] C. H. Grein and Sajeev John. Temperature dependence of the Urbach optical absorption edge: A theory of multiple phonon absorption and emission sidebands. *Physical Review B*, 39(2):1140–1151, January 1989.

[73] A. Zaslavsky, C. A. Richter, P. R. Shrestha, B. D. Hoskins, S. T. Le, A. Madhavan, and J. J. McClelland. Impact ionization-induced bistability in CMOS transistors at cryogenic temperatures for capacitorless memory applications. *Applied Physics Letters*, 119(4):043501, July 2021.

[74] David Esseni, Pierpaolo Palestri, and Luca Selmi. *Nanoscale MOS Transistors: Semi-Classical Transport and Applications*. Cambridge University Press, 1 edition, January 2011.

[75] S.M. Sze and Kwok K. Ng. *Physics of Semiconductor Devices*. Wiley, 1 edition, October 2006.

[76] Supriyo Datta. *Quantum Transport: Atom to Transistor*. Cambridge University Press, 2005.

[77] M. P. Anantram, Mark S. Lundstrom, and Dmitri E. Nikonov. Modeling of nanoscale devices. *Proceedings of the IEEE*, 96(9):1511–1550, 2008. arXiv: cond-mat/0610247 ISBN: 0018-9219.

[78] D. S. Domanevskii, V. V. Krasovskii, M. V. Prokopenya, V. A. Vilkotskii, and S. V. Zhokhovets. The Origin of the Minority Impurity States in Heavily Doped Semiconductors. In *February 1*, pages 693–700. De Gruyter, December 1986.

[79] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. S. Lundstrom. A simple quantum mechanical treatment of scattering in nanoscale transistors. *Journal of Applied Physics*, 93(9):5613–5625, May 2003.

[80] R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom. Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors. *Journal of Applied Physics*, 95(1):292–305, January 2004.

[81] Aryan Afzalian. Computationally efficient self-consistent born approximation treatments of phonon scattering for coupled-mode space non-equilibrium Green's function. *Journal of Applied Physics*, 110(9):094517, November 2011.

[82] K. Miao, S. Sadasivam, J. Charles, G. Klimeck, T. S. Fisher, and T. Kubis. Büttiker probes for dissipative phonon quantum transport in semiconductor nanostructures. *Applied Physics Letters*, 108(11):113107, March 2016.

[83] Wenjuan Zhu, Tony Low, Yi-Hsien Lee, Han Wang, Damon B. Farmer, Jing Kong, Fengnian Xia, and Phaedon Avouris. Electronic transport and device prospects of monolayer molybdenum disulphide grown by chemical vapour deposition. *Nature Communications*, 5(1):3087, May 2014.

[84] Xiaolong Chen, Lin Wang, Yingying Wu, Heng Gao, Yabei Wu, Guanhua Qin, Zefei Wu, Yu Han, Shuigang Xu, Tianyi Han, Weiguang Ye, Jiangxiazi Lin, Gen Long, Yuheng He, Yuan Cai, Wei Ren, and Ning Wang. Probing the electronic states and impurity effects in black phosphorus vertical heterostructures. *2D Materials*, 3(1):015012, March 2016.

[85] V. Eyert. A Comparative Study on Methods for Convergence Acceleration of Iterative Vector Sequences. *Journal of Computational Physics*, 124(2):271–285, March 1996.

[86] Arnout Beckers, Farzan Jazaeri, Heorhii Bohuslavskyi, Louis Hutin, Silvano De Franceschi, and Christian Enz. Design-oriented modeling of 28 nm FDSOI CMOS technology down to 4.2 K for quantum computing. In *2018 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, Granada, March 2018. IEEE.

[87] Yiju Zhao, Youngki Yoon, and Lan Wei. A Multi-Level Simulation Scheme for 2D Material-Based Nanoelectronics. In *2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO)*, pages 388–392, July 2020. ISSN: 1944-9380.

[88] Serge Luryi. Quantum capacitance devices. *Applied Physics Letters*, 52(6):501–503, February 1988.

[89] Jilin Xia, Fang Chen, Jinghong Li, and Nongjian Tao. Measurement of the quantum capacitance of graphene. *Nature Nanotechnology*, 4(8):505–509, August 2009.

[90] J. C. Slater and G. F. Koster. Simplified LCAO Method for the Periodic Potential Problem. *Physical Review*, 94(6):1498–1524, June 1954.

[91] José M. Soler, Emilio Artacho, Julian D. Gale, Alberto García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. The SIESTA method for ab initio order-N materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745, March 2002.

[92] Gérard Ghibaudo, Mohamed Aouad, Mikaël Cassé, Sebastien Martinie, Francis Balestra, and Thierry Poiroux. On the modelling of temperature dependence of sub-threshold swing in MOSFETs down to cryogenic temperature. *Solid-State Electronics*, October 2020. Publisher: Elsevier.

[93] Nan Fang and Kosuke Nagashio. Band tail interface states and quantum capacitance in a monolayer molybdenum disulfide field-effect-transistor. *Journal of Physics D: Applied Physics*, 51(6):065110, February 2018. arXiv:1801.10345 [cond-mat].

[94] Arnout Beckers, Farzan Jazaeri, Alexander Grill, Subramanian Narasimhamoorthy, Bertrand Parvais, and Christian Enz. Physical Model of Low-Temperature to Cryogenic Threshold Voltage in MOSFETs. *IEEE Journal of the Electron Devices Society*, 8:780–788, 2020.

[95] Markus Kantner and Thomas Koprucki. Numerical simulation of carrier transport in semiconductor devices at cryogenic temperatures. *Optical and Quantum Electronics*, 48(12):543, December 2016.

[96] Akin Akturk, Jeffrey Allnutt, Zeynep Dilli, Neil Goldsman, and Martin Peckerar. Device Modeling at Cryogenic Temperatures: Effects of Incomplete Ionization. *IEEE Transactions on Electron Devices*, 54(11):2984–2990, November 2007.

[97] Lida Ansari, Scott Monaghan, Niall McEvoy, Cormac Ó Coileáin, Conor P. Cullen, Jun Lin, Rita Siris, Tanja Stimpel-Lindner, Kevin F. Burke, Gioele Mirabelli, Ray Duffy, Enrico Caruso, Roger E. Nagle, Georg S. Duesberg, Paul K. Hurley, and Farzan Gity. Quantum confinement-induced semimetal-to-semiconductor evolution in large-area ultra-thin PtSe2 films grown at 400 °C. *npj 2D Materials and Applications*, 3(1):33, September 2019.

[98] Azimkhan Kozhakhmetov, Riccardo Torsi, Cindy Y Chen, and Joshua A Robinson. Scalable low-temperature synthesis of two-dimensional materials beyond graphene. *Journal of Physics: Materials*, 4(1):012001, November 2020.

[99] Daniil Marinov, Jean-François De Marneffe, Quentin Smets, Goutham Arutchelvan, Kristof M. Bal, Ekaterina Voronina, Tatyana Rakhimova, Yuri Mankelevich, Salim El Kazzi, Ankit Nalin Mehta, Pieter-Jan Wyndaele, Markus Hartmut Heyne, Jianran Zhang, Patrick C. With, Sreetama Banerjee, Erik C. Neyts, Inge Asselberghs, Dennis Lin, and Stefan De Gendt. Reactive plasma cleaning and restoration of transition metal dichalcogenide monolayers. *npj 2D Materials and Applications*, 5(1):17, January 2021.

[100] Suyeon Cho, Sera Kim, Jung Ho Kim, Jiong Zhao, Jinbong Seok, Dong Hoon Keum, Jaeyoon Baik, Duk-Hyun Choe, K. J. Chang, Kazu Suenaga, Sung Wng Kim, Young Hee Lee, and Heejun Yang. Phase patterning for ohmic homojunction contact in $MoTe_2$. *Science*, 349(6248):625–628, August 2015.

[101] Jongyeol Baek, Demin Yin, Na Liu, Inturu Omkaram, Chulseung Jung, Healin Im, Seongin Hong, Seung Min Kim, Young Ki Hong, Jaehyun Hur, Youngki Yoon, and Sunkook Kim. A highly sensitive chemical gas detecting transistor based on highly crystalline CVD-grown MoSe2 films. *Nano Research*, 10(6):1861–1871, June 2017.

[102] Jiahao Kang, Wei Liu, Deblina Sarkar, Debdeep Jena, and Kaustav Banerjee. Computational study of metal contacts to monolayer transition-metal dichalcogenide semiconductors. *Physical Review X*, 4(3):1–14, 2014.