SEMICLASSICAL SIMULATIONS OF HOT ELECTRONS IN GATE-ALL-AROUND SILICON MOSFETS

By

Mahmud Reaz

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Interdisciplinary Materials Science

August 13, 2021

Nashville, Tennessee

Approved:

Ronald D. Schrimpf, Ph.D.

Sokrates T. Pantelides, Ph.D.

Robert A. Reed, Ph.D.

Stephanie L. Weeden-Wright, Ph.D.

Greg Walker, Ph.D.

Copyright © 2021 Mahmud Reaz
All Rights Reserved

To my grandparents: Shajeda Khatun & Md. Fazlur Rahman
parents: Bilkis Akhter & Md. Shirajul Islam,
sister: Kakoli, uncle: Manik,
wife: Nawsheen, and daughter: Zenobia

ACKNOWLEDGMENTS

This dissertation is the outcome of 'golden years' spent in the '*Radiation Effects and Reliability Group*' at Vanderbilt University. Throughout the journey, I have had the opportunity to meet and learn from several leading scholars of the microelectronics field. Interaction with them enriched me as a researcher and as a person in ways that cannot be overstated.

My sincere appreciation goes towards my advisor, Professor Ronald D Schrimpf, who guided me every step of the way. He is an exceptionally inspiring human being who kept an open-door policy for 24/7. The foundation of this dissertation work is laid upon many of his invaluable lessons. I am indebted to him for introducing me to the field of solid-state electronics and stimulating interests in me towards fundamental (first-principle) thinking. This dissertation, prepared over years of research, would not be possible without his remarkable kindness, patience, and encouragement.

A very special acknowledgment goes to Professor Sokrates T Pantelides, a pioneer in the field of density functional theory. He was truly inspirational in all aspects, and his wisdom on the physics of solids, first-principles simulation, and their implication to technology greatly benefited my research.

I am thankful to my remaining committee members: Professor Robert Reed, Dr. Stephanie L. Weeden-Wright, and Professor Greg Walker, for their continuous support and encouragement. Professor Reed contributed to my understanding of charge-energy transfer phenomena in semiconductors and many others. I have highly enjoyed conversations with Professor Walker regarding the phonon transport and Monte Carlo simulations in materials. Chapter four of this dissertation is primarily based on a proposal by Dr. Stephanie L. Weeden-Wright, and many thanks to her for including and funding me in her project.

The primary code of the simulator is written by Professor Massimo V Fischetti from the University of Texas at Dallas. He is a pioneer in device simulation, and I am truly thankful to him for assisting me in every possible way throughout my Ph.D. Several discussions in the method section are adapted from his publications.

I am grateful to Professor Fleetwood for helping to tie the theoretical results to experiments. It was a true privilege to learn the intricacies of device and interface imperfections from him. Above all, I am thankful to all aforementioned researchers for schooling me in the craft of persuasive academic writing — invaluable teaching offered by the very best in it.

I am thankful to Dr. Jingtian Fang for teaching me countless concepts of electron transport simulation and providing me initial guidance in getting started with the Fortran Programming language. I do appreciate the conversations with Dr. Michael Alles about the technological aspect of the study and Dr. Enxia Zhang and Dr. Andrew Sternberg about the experimental aspects.

I am grateful to Prof. Witulski for being flexible —he allowed me independence in managing competing projects that I had been fortunate to be a part of. The insights that I have gained from Professor Bhuva about electronic circuits have helped me better understand the broader implications of my work. I am fortunate for the assistance of Fenglai Liu in setting the simulation environment in the Accre Cluster. Without him, the project would stretch much longer.

The comprehensive studies presented in this dissertation were not possible without the help of many of my colleagues. In particular, Andrew Tonigan (now a Dr.) aided in devising the simulation studies and validating the initial results. The code for several illustrations presented in this dissertation was developed with the extensive help of Brandon Smith. Mohammad W Rony ran TCAD simulations, which enabled the comparison of the results of the Monte Carlo simulations to the concurrent industry standards. Experimental investigation of the electrical characteristics by Mariia Gorchichko was used in calibrating the oxide thickness of the device. In addition to helping me with the modeling, I am thankful to Kan Li for being a friend throughout. I am grateful to imec for sharing their state-of-the-art devices with us. In addition, the derived insights from imec's reliability study of the GAA devices were instrumental in understanding the implications of the results of this study and aided greatly in reaching conclusions. Finally, I would like to give special regards to my MS supervisor, Professor Kartik C Ghosh, for his mentorship and lessons on spectroscopy techniques that prepared me well for graduate school.

I do appreciate the encouragement from my family throughout — without them, no achievement is meaningful enough.

TABLE OF CONTENTS

Page

DEDICATION .... **iii**

ACKNOWLEDGMENTS .... **iv**

LIST OF TABLES .... **viii**

LIST OF FIGURES .... **ix**

Chapter 1 .... **1**

1.1 Preface 1
1.2 Reliability of Microelectronic Devices and IC 2
1.3 Simulation Methods for Studying Electron Transport 4
1.4 Suitability of Semiclassical Simulations for Scaled Transistors 6
1.5 Hot-Carriers in Materials 7
1.6 Hot-Carriers in GAA Devices 8
1.7 Barriers to Continuous Scaling 9
1.8 Summary 10

Chapter 2 .... **12**

2.1 Introduction 12
2.2 Assumptions 13
2.3 Transport and Field Equations 13
2.4 Electronic Energy Bands 16
2.5 Calculation of Scattering Rates 16

2.6 Device Boundaries and Interfaces 18

2.7 Initial Conditions 19

2.8 Ensemble Monte Carlo Method for Solving BTE 19

2.9 Assigning Particle Properties to Mesh 21

2.10 Solution of the Poisson Problem 22

2.11 Coulomb Scattering 23

Chapter 3 28

3.1 Preface 28

3.2 Relaxation of Hot-carriers in Materials 28

3.3 Electron-Hole Pair Creation Energy: Experimental Findings 30

3.4 Thermalization Process of High-Energy Carriers 31

3.5 Background: Energy Loss Mechanisms 31

3.6 Synopsis of the Work 33

3.7 Full band Monte Carlo simulations 34

3.7.1. Boltzmann transport equation 34

3.7.2. Carrier Scattering Mechanisms 36

3.8 Simulation Results 37

3.8.1. EHP production in silicon via photoabsorption 38

3.8.2. Dominant Scattering Mechanisms 42

3.8.3. EHP creation in silicon via electrons and holes 43

3.8.4. Pair creation in germanium 44

3.9 Summary 45

Chapter 4 46

4.1 Preface 46

4.2 Key Aspects of Three-Dimensional Transistors 46

4.3 Addressing Open Questions for GAA Devices....................48
4.4 Degradation due to Short-Channel Effects in GAA Devices....................49
4.5 Simulation Methods and Scattering....................52
4.6 Carrier Transport in a Si Nanowire MOSFET....................58
4.7 Spatial and Energy Distribution of Hot Electrons....................60
4.8 Comparison with Experimental results....................63
4.9 Summary....................67

Chapter 5....................69

5.1 Introduction....................69
5.2 Scattering Barrier to Ballistic Transport....................70
5.3 Summary....................76

Chapter 6....................77

6.1 Introduction....................77
6.2 Discussions....................79
6.3 Summary....................85

Chapter 7....................86

References....................88

LIST OF TABLES

Page

Table I. The impact-ionization model parameters for electrons and holes in silicon and germanium. The threshold energy is relative to the conduction and valence band edge for electrons and holes, respectively. 35

Table II. Quantum Efficiency Dependence on the Incident Photon Energy in Silicon 37

Table III. Quantum Efficiencies for Various Incident Photon Energies in Germanium 43

LIST OF FIGURES

Page

Fig. 1. Optimum separation of carriers (left axis) and charge wight (right) for treating long range plasma scattering. The characteristic length for Landau damping become smaller at high densities and deviates from $n - 1/2$ behavior due to degeneracy. Fig. is adapted from ref. [171]. 22

Fig. 2. (a) The electronic band structure of silicon calculated with empirical pseudopotentials. Five conduction bands and three valence bands are shown. The conduction band minimum along the path Γ → X is set as the reference energy of zero. Degeneracy of electronic states is included. (b) The joint density of states for photon absorption with different incident photon energies in silicon is shown with the solid blue line. Only vertical transitions are considered for photon absorption to create electron-hole pairs. The electron-hole-pair creation energies for silicon devices excited with different photon energies studied below are indicated with the red dots. 30

Fig. 3. (a) Carrier-phonon and impact ionization rates for electrons and holes. The conduction-band minimum is set at 0. By convention, the electron (hole) energy is assigned a positive (negative) sign. (b) The spatial distribution of Monte Carlo particles superimposed on the energy band diagram under illumination of 6.3 eV photons. The colormap indicates the carriers' kinetic energy in eV. By convention, electrons are assigned a positive kinetic energy, whereas holes are assigned a 'negative' kinetic energy sign..... 32

Fig. 4. (a) Calculated photocurrent collected on the left contact for various photon energies. In all cases, the illumination starts from equilibrium at t = 0, and lasts until t = 7 ps, as indicated by the dashed line. Here the currents are averaged over a time span of 1 ps for each moment. (b) The dependence of the quantum efficiency on the incident carrier energy. The horizontal blue dashed line indicates the unity of the quantum efficiency. The red

dashed line indicates the fitted linear relationship between the quantum efficiency and the incident particle energy. The inverse of the slope corresponds to the EHP creation energy of 3.69 eV for silicon. Quantum efficiencies for photon energies below 3.6 eV have not been simulated. 38

Fig. 5 The number of events (a) and the energy loss or gain (b) for electron scattering with different phonon modes, and the number of events (c) and the energy loss or gain (d) for hole scattering with different phonon modes. The phonon emissions result in carrier energy loss and phonon absorptions result in carrier energy gain. Some of the phonon modes considered in the simulations are indicated in the inset of the figure, with LA indicating the longitudinal acoustic phonon, TA indicates the transverse acoustic phonon, and NOP indicates the nonpolar optical phonon mode. 39

Fig. 6. (a) The dependence of the EHP creation energy on the incident electron and hole energy in silicon. (b) The dependence of the quantum efficiency on the incident carrier energy. The horizontal blue dashed line indicates the unity of the quantum efficiency. The red dashed line indicates the fitted linear relationship between the quantum efficiency and the incident carrier energy. The inverse of the slope corresponds to the EHP creation energy of 2.62 eV for germanium.. 41

Fig. 7. The dependence of the EHP creation energy on the incident electron and hole energy in germanium. 42

Fig. 8. Short-channel characteristic of FinFET versus horizontally-integrated GAA-NW (after [17]). (a) Better subthreshold characteristic for linearly stacked GAA devices (b) Unacceptable drain-induced barrier lowering for scaled FinFETs. 49

Fig. 9. TEM images of a GAA Si nMOSFET from imec [4] depicting a (a) NW array and (b) single NW vertical stack. WF metal represents the work-function metal used for the gate electrode. (c) The simplified device used for simulation is shown schematically (not

drawn to scale). The insulating oxide in the S/D access region and gate contact extends radially outward until the end of the tensor mesh (not shown). 56

Fig. 10. (a) Experimental and simulated gate control at 50 mV, 0.5 V, and 0.9 V drain bias. Curves are shifted along the x-axis by -0.5, 0.0, and 0.5 V, respectively. The experimental drain current is divided by 2 to match the simulated *I-V* characteristics of the single NW structure depicted earlier. (b) Electron potential energy along the x-direction of the cylinder, illustrated with a color map. The density of the equipotential lines reflects the magnitude of the electric field. Metallurgical junction boundaries are at x = 10 nm and 34 nm. The dashed-red lines separate the oxide and semiconductor region of the device. The aspect ratio is different than actual device dimensions for clarity..................................... 58

Fig. 11. Carrier kinetic energy (KE) distribution above the conduction-band edge along the x-direction of the device. The conduction-band edge is averaged over the cross section of the channel. Statistical carrier enhancement within the dashed vertical lines ensures adequate sampling of interactions in an otherwise sparsely populated region. The kinetic energy of the carriers is represented by both particle color and the distance above the conduction band.. 59

Fig. 12. Energy distribution of electrons for several drain biases at $V_{GS} - V_{Th} = 1.3$ V. (a) Electron-energy histogram within 2 nm in each direction of the metallurgical drain-junction. (b) Density of electrons above 0.5 eV in the NW. The dotted curves show results without including EES.. 61

Fig. 13. Density of electrons above 0.5 eV in the NW for several overdrive biases at $V_{DS} =$ 0.9 V. Dashed lines show the hot-electron distribution without EES................................ 62

Fig. 14. (a) Experimentally derived time-to-failure (TTF) maps of the GAA device, adapted from Chasin *et al.* [7]. (b) The simulated density of hot carriers in a small volume near metallurgical junction ~ within ±2 nm in the x-direction. (c) The simulated average kinetic

energy of these hot electrons. Only the electrons with energy higher than 0.5 eV is considered for subplots (b) and (c). The scales of the color bars in (a) is logarithmic, and those in (b) and (c) are linear. 64

Fig. 15. Dependance of the drive current to the elastic scattering processes in the channel for several drain biases. Points with cross and circles respectively indicates the simulation results with and without accounting for Coulomb processes. 70

Fig. 16. Trend for phonon emission (loss) by electrons in the channel and drain region with respect to the applied drain-voltage. Phonon emissions are calculated as a sum of transverse acoustic, longitudinal acoustic, and non-polar optical phonon modes. Both curves are individually normalized by their maximum value since the number of particles in the channel and drain region is different in the simulation. 71

Fig. 17. Histograms showing the errors in selecting the final state in k-space for elastic (a) electro-electron and (b) electron-ionized impurity scattering. The mean and standard deviation of the error while selecting the final state of the electrons are much smaller than the energetics of the simulation. 72

Fig. 18. (a) The number of events (b) and the energy loss or gain for electron scattering with different elastic and phonon modes. The plotted data were temporally accumulated over some period of steady state simulation in the entire device. Some of the elastic and phonon modes considered in the simulations are indicated in the x-axis of the figure, with LA indicating the longitudinal acoustic phonon, TA indicates the transverse acoustic phonon, NOP indicates the nonpolar optical phonon, e-imp indicates the electron-impurity, and e-e indicates the electron-electron scattering mode. The red and green color correspondingly indicates the emission and absorption of energy by the electrons. Note: The capital typefaces in x-axis indicate inelastic events and differentiate to those from the

Coulomb processes in small typefaces. The energy loss-gain by the ee and e-imp processes are not real but only provided to illustrate the numerical aspect of energy conservation. . 73

Fig. 19. (a) The number of events (b) and the energy loss or gain for electron scattering with different elastic and phonon modes in the channel region only. Similar conditions to that of Fig. 17 apply. Except e-e scattering, all other momentum transfer events are significantly reduced in the channel regions compared to those plotted including the access regions in Fig. 18. 74

Fig. 20. The energy loss/gain spectrum of the electrons in the channels with the phonon modes considered in the simulation. The range of the interaction in LA, TA, and NOP modes follow the phonon energy bands, as considered in the simulation. For acoustic phonon modes in (a)-(d), electrons emit almost twice energy compared to the energy absorbed. Thus, emission spectrum for the acoustic phonons skewed to higher energies resulting in net loss of energy for the electrons. Energy loss by the NOP phonons in (e) are five times higher than the absorption spectra in (f). NOP interactions occur at a single energy due to the choice a flat energy band for the optical phonons................................. 75

Fig. 21. Two-dimensional approximation of an idealistic circular NW nMOSFET. Effective mass Schrodinger equation is solved only on the region shown in beige (light-brown). The metallurgical junction boundaries in the source and drain side are indicated with dotted lines. The illustration is not drawn to scale. 81

Fig. 22. Simulated (a) electrostatic potential (b) charge density and (c) distribution of the charge carriers in the device. The maximum of the potential in (a) is arbitrary chosen as zero. Source-junction barriers is highest along the mid-channel for electron conduction from source (left) to drain (right). This barrier drops from the center towards the interface. The charge density for the electrons in (b) show two separate inversion layers associated to the second subband. The electric field in the transverse direction is less than needed for

volume inversion. Illustrated in (c), electrons in the quantized region $(12 < x < 32)$ are distributed in real space accordingly to the squared wavefunction and in k-space accordingly to their energy. Electrons populate the first and second subband up to 0.15 eV and are treated as bulk at higher energies. The different sizes of the particles reflect their differences in charge-weight on logscale.. 82

Fig. 23. (a) Potential for the electrons in the quantized region along the transverse direction at mid-channel. Corresponding wavefunctions for the ground (red) and first excited state (green) unprimed subbands... 84

Chapter 1

Introduction

1.1 Preface

Three-dimensional semiconductor devices are expected to reduce power consumption, increase computational block density, and provide sufficient drive-current. The reliability risks of these devices due to minority carrier injection or impact ionization events have been reduced by a considerable margin since both logic and memory transistors of present Si technology operate at supply-voltages below the silicon bandgap. But hot-electron effects due to majority carriers remain a barrier to the ultimate scaling of Si transistors [1]. In this regard, the apparent mobility of electrons and the lifetime of highly-scaled devices are known to be sensitive to several factors, including doping, potential profile, electron velocities, and momentum relaxation processes. Understanding these key processes is important both for device designers and assurance engineers to better design and analyze the devices. The aim of this dissertation is to quantitatively relate the microscopic electron transport phenomena to the performance and lifetime of a short-channel n-MOSFET Si transistor. The underlying simulation model employed in this article employs classical (Newtonian) trajectories for electrons while including quantum physics for their kinetic energies and for treating collisions in momentum space.

In this work, the theoretical contexts for semiclassically treating electron transport are illustrated first to provide the basis for the presented results. The literature is rich in this topic with decades of development for silicon, and hence, significant areas of relevance are identified for the discussion. Second, phonon and impact-ionization processes on large-area semiconductors are investigated to identify the origin of charge-energy conversion in technologically important semiconductors. Excellent agreement with experimental results substantiates that the ensemble Monte Carlo electron transport method using full electronic

dispersion is well suited to explain the energy transfer of the hot-electrons to the lattice, at least up to 3 eV energies. Third, additional energy-transfer processes are identified for nanodevices and included for simulating hot-carrier effects in gate-all-around (GAA) silicon nanowire FETs. An agreement of the simulated energetics to experiments reveals that the mechanisms for electron-induced degradation in the scaled devices might differ from what was popularly accepted previously for the devices and bias-conditions of the past era.

Fifth, Coulomb interactions were studied and found to be responsible for significant mobility reduction in low-field conditions. The physics of the confined carriers, as found in the current literature, is presented. Mobility-limiting scattering events in the reduced dimensions are identified from the contemporary literature, and common aspects of geometric confinement are illustrated by a semiclassical simulation using simplified assumptions. At the very end, routes for a more rigorous simulation study of the carriers in an ultra-thin channel are offered.

1.2 Reliability of Microelectronic Devices and IC

Developing reliable chips with smaller silicon transistors is a problem of paramount scientific and economic importance. FinFETs, with their three-dimensional geometry and excellent gate-control, have helped the microelectronics industry to deliver integrated circuits to the 10 nm node and beyond (smaller sizes). However, further scaling has become increasingly difficult due to the higher sensitivity of the transistors' performance to the semiconductor design, fabrication processes, integration methods, and material properties. Disruptive technology changes were/are necessary to scale down these devices closer to their fundamental limits — how small transistors can be made while still offering acceptable performance. Short-channel effects (SCE) are important in these nanodevices when the channel length becomes comparable to the depletion layer widths of the source and drain junctions. SCE for Si devices are widely studied at sub-30 nm dimensions and GAA devices are preferred. An ultimate transistor design is expected to be a derivative of the gate-all-around (GAA) geometry due to the better electrostatic control of the gate over the carriers in

the channel [1], [2]. Densely stacked GAA devices can provide three times larger current than their planar counterparts once the process variability during the fabrication is reduced [3]. Vertically-stacked (VS) GAA devices, fabricated by Interuniversity Microelectronics Centre (imec) as a breakthrough in the field in 2016 [4], [5], are indeed what scientists dreamed of over the last decades as a flawless Si MOSFET.

Nevertheless, these devices were found to be comparable to FinFETs in electrical performance and, in some cases, have shown inferior reliability compared to FinFETs [2], [6]–[9]. GAA transistors would eventually outperform FinFETs once the process technology is better optimized, but the extent of improvement might be smaller than expected. Simplified modeling techniques and misinterpretation of data that previously gave birth to a widespread optimism for GAA devices must be tested with experiments and physics-based simulations because electrons in nanomaterials behave differently and are sensitive to a wide range of circumstances.

The performance and reliability of transistors depend on several factors, including process technology, material properties, bias temperature instability (BTI), hot electron-induced processes, electrostatic control, device doping, geometry, channel length, scattering modes, incoherent processes, and quantum effects [9]–[18]. Significant efforts are ongoing to reduce/assess the unpredictability from random dopant fluctuation (RDF) [3], [19]–[21], work function variation (WFV) from the metal gate granularity (MGG) [12], [21], [22], metal thickness variation (MTV) [23]–[25], line edge roughness (LER), and variation in oxide structure (shape, thickness, underlap/ overlap) [21]. The extreme sensitivity to process variability for the devices with reduced dimensions is straightforward to understand. For example, the presence of a single unintentionally implanted phosphorus atom can raise the doping of the access regions by ~2% and doping of the channel by orders of magnitude. In addition, a monolayer of oxide may constitute up to 33% of the effective oxide thickness

(EOT) of a SiO_2-HfO_2 gate structure. Although RDF and WFV are found to influence the device properties the most, the apparent limitation is not fundamental [20], [21], [25].

To investigate the absolute limiting mechanisms in Si, modeling the reliability and performance data with equilibrium statistics or fitting parameters is not sufficient [17], [26]–[30]. These popular analytical/empirical models are based on the experimentally derived parameters (average velocity, apparent-mobility, channel current, etc.) [15], [31], [32] and/or a multi-vibrational mechanism for the electrons to collectively surmount the barrier at the oxide-semiconductor interface. Multi-vibrational effects (MVE) were initially introduced to justify electrons crossing interface barriers higher than their individual energy. While doing so, the direct Si-H bond dissociation by the hot carriers at the interface was considered the most critical step [7], [8], [33]–[37]. Although the MVE process might still be valid for the bias-regime (1.0- 4 V) and dimensions (30- 200 nm processes) of the pre-millennial devices, the pathways for hot-carrier-induced degradation in the sub-30 nm devices can be different. Such differences occur because logic transistors fabricated in recent times are operated at biases smaller than 1 V, and the probability for several electrons of less than 1 eV energy to jointly excite Si-H bonds to sufficient energy (require > 2.5 eV) is not determined and is expected to be low. Before these cross-sections are calculated or experimentally determined for sub-30 nm devices, the simulated electron energy distribution from a high-fidelity approach remains the best bet to uncover the critical steps of device degradation. Besides widespread adoption (and controversy) of MVE mechanisms within the community, no physics-based evidence exists to conclusively validate one mechanism or other for the devices of 10 nm process nodes with ~ 20 nm gate-length.

1.3 Simulation Methods for Studying Electron Transport

Compact Modeling of billions of transistors together on a die during the integrated circuit development employs several approximations, as found in the drift-diffusion (DD) or virtual source (VS) models. These tactics are highly useful and no more complicated than data-

fitting with existing knowledge — the parameters of the fit are mined from high precision experiments to high-fidelity simulations using density functional theory, semiclassical transport method, and non-equilibrium Green's function (NEGF). Short-channel devices are susceptible to such parametric simulation approaches as both carriers and material properties (e.g., velocity saturation, coupled horizontal and vertical electric field), which can be averaged in larger dimensions, diverge at nanometer length scales [38]–[40]. As such, using DD and VS models for simulating novel nano-devices is worrying since parameters no longer remain transferable from other devices of different designs and dimensions.

Conventionally, a hydrodynamic (HD) approach with energy conservation provides the parameters for density-based simulations, but the constraints of HD simulation itself are derived from a more rigorous semiclassical electron transport simulation [41]–[44]. HD models are useful for their low computational overhead, but the extent of velocity overshoot observed from HD simulation is overly sensitive to the accuracy of the energy relaxation time. Again, energy relaxation time is not known for devices of the new design unless a high-fidelity semiclassical or quantum transport simulation has been performed on an equivalent device beforehand. Semiclassical simulations are superior because detailed information about the full electronic energy band and energy-dependent scattering rates (quantum-mechanically derived) are integrated into the workflow [10], [45]–[47]. The subsequent chapters display the application and usefulness of semiclassical techniques in analyzing electron transport for large and small-volume Si devices.

Even better, perhaps, is the application of NEGF to transport problems at sub-10 nm dimensions because confinement and tunneling are addressed appropriately. From Keldysh to Landauer–Buttiker [48], [49], the NEGF technique is a rigorous approach to include quantum phenomena, incoherent scattering with phonons, and device imperfections [10], [48]–[52]. The NEGF goes beyond the semiclassical approach by directly incorporating quantum behavior and normalizing the density of states, unlike the semiclassical approaches

with Fermi's golden rule [53]. However, comprehensively treating inelastic/incoherent interactions within the NEGF formalism adds an extra iteration cycle to an already formidable computational task [54]. Utilizing the computational power of modern times, several research groups have been able to (often separately – one at a time) demonstrate the effects of confinement, tunneling, Coulomb-blockade, Coulomb, and phonon interactions w.r.t length and thickness of the channel [10], [50]–[52]. Carrier mobility is found to non-linearly depend on the length and thickness of the channel [10], [50]–[52]. Furthermore, geometric confinement was substantial for devices with a smaller cross-section than $6 \times 6\,\mathrm{nm}^2$ [55]. In addition, NEGF simulation reveals that additional thinning of the channel would cause electron-phonon and electron-roughness scattering to overshoot, triggering a significant reduction in mobility. The tunneling is relevant only below 10 nm gate-length [56]–[58] but is considered irrelevant now since scaling along the long-channel direction for Si has already stopped in view of diminishing performance gain [59]. Overall, electrostatic improvement in the GAA geometry may open doorways for scaling, except the effects of scatterings in these devices are largely unknown [18], [60].

1.4 Suitability of Semiclassical Simulations for Scaled Transistors

GAA transistors so far are scaled down to around ~ 20 nm channel length and 6-8 nm diameter by imec. Device designs in 5 nm forksheets [1] technology by imec, 7 nm processes of Taiwan semiconductor manufacturing company (TSMC), and the 10 nm nodes of Intel are largely unknown to academics due to proprietary reasons and are primarily in Fin geometry (except forksheets [1]). However, variants of GAA devices are only considered promising to scale down to 3 nm nodes [18]. In the thin GAA transistor, a fair number of electrons travel in the atomic layers closer to the interface and may experience strong scattering from phonon processes, many-body effects, surface roughness, interface and/or border traps, Coulomb forces, plasma interactions, impurity scattering [17], [47], [61]–[63]. Considering the dimensions (L_G ~ 24 nm, diameter ~ 8 nm) of the vertically stacked GAA

devices fabricated by imec, the semiclassical three-dimensional simulation framework is still the most comprehensive tool to understand the long-range scattering, hot-carrier triggered mechanisms, induced vulnerabilities, and atomic-level nuances (e.g., carrier-carrier interaction, source starvation, bias dependence, impact ionization, charge generation, etc.) [64]. Stochastic approaches to numerically solving the BTE treat hot-carrier dynamics more accurately as they do not depend upon parameters that are defined predominantly for equilibrium carrier distributions. Moreover, energy-exchange processes among hot-carriers, cold-carriers, and the lattice are better accounted for by using a semiclassical approach [39], [52], [65], [66]. In addition, several quantum phenomena can be included semiclassically without solving the Schrodinger equation in real time— phonon scattering rates and electronic dispersions can be pre-derived from first-principles calculations and experimental form factors.

NEGF is not preferred since quantum effects are not strong for these dimensions, and treating the long-range interactions comprehensively with NEGF in the entire transistor is still an insurmountable numerical problem [11], [67]. TCAD simulation has evolved and now incorporates the six-band k•p method to include electronic band structure, kinetic velocity model for ballistic transport, bandgap narrowing, and quantum alteration in the inversion layer. However, the inherent model does not directly account for scattering phenomena that control energy transport in the channel of a quasi-ballistic device [18].

1.5 Hot-Carriers in Materials

Beyond constraints due to device design, the behavior of the hot carriers fundamentally depends on the structural and electronic properties of a semiconducting material [52]. Before treating several complex scattering modes altogether while simulating an entire device, semiclassical simulations and the treatment of the hot carriers should be verified first at the materials-level with limited complexity [52]. Since energy-loss by hot carriers limits the speed, current-drive, and mobility of semiconducting materials and devices [9], electron

transport in high energies and subsequent thermalization have been simulated using a full-band Monte Carlo (FBMC) method and presented in ***chapter III***. An excellent agreement of the calculated EHP creation energy to experiments supports the accuracy of the employed semiclassical method and reveals the phonon interaction cross-sections and impact ionization processes as a function of carrier energy and time. This comprehensive approach contrasts with traditional TCAD simulations incorporating the DD formalism, where electron-hole pair (EHP) creation energy is a user-defined input.

The details of the problem statement of the work, methods, and contribution of scattering processes to the energy-loss of the hot-carriers are presented in ***chapter III*** and compared to that of a scaled GAA device in ***chapter V***. The simulations and derived conclusions are obtained for semiconducting materials with large dimensions where quantum confinement effects are negligible. The diode structure, as chosen for the simulations, enables charge-collection at the contacts, similar to the experimental setups used in measuring the ionization energy of carriers. The workflow of the semiclassical simulator used in this work is discussed at length in ***chapter II***.

1.6 Hot-Carriers in GAA Devices

Potential fluctuations, source starvation, ionized impurity scattering, plasma, and short- and long-range interactions affect the carriers in the nanodevices and are influenced by the limited number of carriers, reduced dimensions, abrupt change in doping, and electrostatic conditions. In addition to similar treatment for the phonon and impact ionization scatterings as presented in ***chapter III***, hot-carrier processes above are rigorously treated at the device level and presented in ***chapter IV***.

The simulated electron energy distributions primarily depend on the drain (V_{DS}) and overdrive (V_{GS} - V_{Th}) field. Some of the electrons gain energies well above the ballistic peak due to phonon processes, localized momentum transfer, and carrier-carrier interactions.

Trends in the simulated electron energy distributions match well trends in experimental studies of device degradation reported in [9]. Our results suggest that the presence of a high concentration of energetic carriers in the drain may lead to the creation of interface traps via an additional pathway initiated by the release of hydrogen from P-H complexes in the drain.

Quantum mechanical (QM) confinement and tunneling along the channel direction depend on the width of the nanowire and length of the active/ channel region, respectively. Considering the findings of the earlier studies [52], [55], quantization and tunneling were found to be less significant for the device under study (DUS) and hence not included in the simulation. However, the QM phenomena would become important for additional scaling and discussed later in ***chapter VI***.

1.7 Barriers to Continuous Scaling

A recent theoretical study for Si GAA devices forecasted acceptable subthreshold characteristics (SS slope $\sim 70\ mV/decade$) up to 8 nm gate-length [18]. Decreasing the gate length even smaller than 8 nm would present a sizable theoretical and technological challenge as reducing leakage current due to quantum mechanical tunneling from source to drain would become challenging, a phenomenon already well-studied for silicon [56]–[58]. In addition, scaling the device along the channel direction beyond 20 nm provides diminishing returns in achieving better effective current density. Contrary to the ballistic dream many had for ultimately scaled devices, we, along with a few other studies, have previously shown that electrons always scatter in the channel regardless of scaling, reducing the current by approximately 20% under moderate applied field conditions [40], [68]. Considering the slight advantage of scaling along the lateral dimensions, it is unlikely that channel-length will get reduced significantly for the next-generation GAA devices or their nanosheet variants. In a quest to continually improve the chip yield, there are still margins in reducing the channel thickness. However, below 6 nm channel diameters, significant current reduction and threshold voltage shifts have been simulated [18], [55], [60]. When the

sub-band splitting for narrower devices becomes comparable or even greater than $\sim k$T, the number of channels for conduction gets significantly reduced.

Consequently, a threshold voltage shift of 0.5 V and a six-fold reduction in channel current has been observed when diameters are shrunk further [69]. Although appropriate gate-metal can account for the change in the threshold voltage, significant current reduction appears inevitable, as already observed in earlier studies. Therefore, scattering phenomena contributing to the mobility reduction in GAA devices are investigated in chapter V. The effect of momentum relaxation and inelastic scattering events on the electron energy distribution functions as obtained in the channel and access region is compared for low-field conditions and it is shown that the mobility reduction is considerable. Furthermore, momentum relaxing elastic events are found to accelerate the phonon-initiated energy transfer processes at low-bias conditions.

Since GAA devices or their nanosheet variants are considered a solution for next-generation technology nodes, future scaling in circular geometry must be comprehensively investigated to assess the remaining degrees of freedom in designing such devices. Hence, the method for treating carrier transport in confined GAA geometry is discussed in chapter VI. For doing so, a device with a channel diameter of 2 nm but otherwise equivalent to the one in chapter IV is selected. Because of the extreme geometric confinement, additional considerations for computing scattering rates and deriving subband structure using the effective-mass Schrödinger equation are highlighted. The surface roughness, phonon-assisted, and Coulomb scattering are more influential at the considered dimensions, and routes are provided to include them appropriately in the simulations.

1.8 Summary

This document addresses the origin of short-channel effects (SCE) in scaled GAA transistors via high-fidelity semiclassical simulation. The theory for semiclassical electron transport is

discussed in ***chapter II***. The treatment of phonon processes, electronic band structure, and impact ionization are validated for the large-area semiconductors in ***chapter III***. The simulation study explains the origin of carrier-generation energy in semiconductors via impact ionization processes. Inelastic phonon processes occurring at or smaller than 5 eV energy dominate the electron-hole pair creation events.

The experimentally validated FBMC approach is extended in ***chapter IV*** to simulate a scaled transistor, including other scattering processes for studying confined volume and high doping devices. Relevant inelastic interactions such as Coulomb scattering, source starvation, potential fluctuations, and intervalley scatterings are included during the electron-transport simulation of a quasi-ballistic device, considering GAA-SNW-FETs as an example. In addition, we provide evidence of why semiclassical approaches are better suited to study SCE in GAA devices with an active region longer than 10 nm and a radius of 4 nm. The simulation results suggest that the basis for MVE mechanisms can be questioned at low operating biases where dopant-induced mechanisms may play a vital role over direct hot-carrier interaction with passivated hydrogen ligands in the oxide-semiconductor interfaces. This conclusion is drawn by correlating the experimentally derived device-lifetime data to the simulated hot carrier distribution in the device.

In ***chapter V***, Coulomb scattering is shown to limit the mobility of GAA devices in the low-field conditions where electron-ionized impurity and electron-electron interactions are strong.

Chapter VI provides thoughts on how to explore the quantum confinement effects by coupling Schrodinger, Poisson, and Boltzmann equations. Relevant physics to be included to study SCE in the nanowires FETs with diameters of 2 nm are presented. ***Chapter VII*** summarizes the key findings of this dissertation work.

Chapter 2

Physical Models and Method for Solution

2.1 Introduction

In this chapter, important theories of semiclassical electron transport are discussed and applied conditions for simulations are elaborated. Adaptation of the Monte Carlo method in solving the Boltzmann transport equation is reviewed, emphasizing the treatment of the energy and momentum transfer processes. Justifications are provided for choosing appropriate physics models in treating transport in the low- and high-density regimes. The employed electron-phonon scattering rate, many-body effects, screening effects, electron-electron and electron-impurity scattering are based on experimentally derived deformation potentials and form factors to pure theoretical studies of a high-density electron gas. Numerical aspects of self-consistently coupling the BTE and Poisson equation are provided in detail, in conjunction with methods that include the complete dispersion of the energy bands for computing scattering rates.

The goal of this chapter is to provide a background for the modeling environments considered while simulating a large area diode and short-channel transistor. The results of these simulations are presented in subsequent chapters. As such, topics that are pivotal for electric transport but constitute a common practice are either shortened or left out. Only the discussions relevant to the presented research in the subsequent chapters are included. Several theoretical aspects of the simulation, as presented in this chapter, are adapted from the content in ref. [70] with permission from the author. The presented information is not new but necessary to establish the results within context as many choices made while treating physics in this work are correct but non-traditional.

2.2 Assumptions

Sharp variation in potential across a single unit-cell cannot be treated well enough without directly invoking first-principles quantum-mechanical treatment [52], [67]. The semiclassical and stochastic method of solving the Boltzmann equation with rigid particles (carriers) is considered valid under the notion that the particle wavelength is much smaller than the size of the semiconductors in all directions [71]. The prerequisite is that the confining potential varies slowly and can be considered roughly constant within each unit cell, defined for computation. This allows electron wave packets to span over several unit cells without experiencing excessive potential variation. Thus, electrons are allowed to be localized simultaneously in real and reciprocal space.

Instantaneous, uncorrelated, and complete collisions by scattering potentials are specified as localized in space and time. In doing so, matrix elements for the scattering strength become zero for larger distance and time [70]. Several scattering centers are added classically to the interaction matrix and valid for weak coupling between the electrons and deformation potentials [52]. Retardation effects from the time-varying charge distribution are neglected, and the Poisson equation is used instead of a complete set of Maxwell's relation. Both short and long-range Coulomb interactions are semiclassically treated. Mesh elements, particle density, and frequency of Poisson solution are chosen such that long-range interactions are treated best within the computational limits of the current era. Empirically determined form factors are used in constructing electronic bands that can evaluate electron transport up to 5 eV. Deformation potentials are used to treat phonon scattering for acoustic and optical phonon branches.

2.3 Transport and Field Equations

Off-equilibrium and time-variant treatment of the charge transport is computed according to the Boltzmann transport equation (BTE)

$$\frac{\partial f(E_n(\mathbf{k}),\mathbf{r},t)}{\partial t} = -\frac{\mathrm{d}\mathbf{k}}{\mathrm{d}t}\cdot\nabla_{\mathbf{k}} f(E_n(\mathbf{k}),\mathbf{r},t) - \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\cdot\nabla_{\mathbf{r}} f(E_n(\mathbf{k}),\mathbf{r},t) + \left(\frac{\partial f(E_n(\mathbf{k}),\mathbf{r},t)}{\partial t}\right)_{coll}, \quad (2.1)$$

where f denotes the probability of carrier distribution as a function of carriers' wavevector, energy state, real space coordinate, and time. $E_n(\mathbf{k})$ denotes the energy of a carrier with wavevector $\mathbf{k}$ in band n and position $\mathbf{r}$. The terms on the right side of eq. 2.1, respectively, account for the drift, diffusion, and scattering of the carriers. The following computes the contributions from scattering of the carriers

$$\left(\frac{df(\mathbf{k},\boldsymbol{r},n,t)}{dt}\right)_{coll} = \sum_v \int \frac{d\boldsymbol{k}'}{(2\pi)^3} S(\mathbf{k},n;\mathbf{k}'v;\boldsymbol{r})\, f(\mathbf{k}',\boldsymbol{r},v,t)\,[(1-f(\mathbf{k},\boldsymbol{r},n,t)]$$

$$-\sum_v \int \frac{dk'}{(2\pi)^3} S(\mathbf{k}',v;\mathbf{k},n;\mathbf{r})\, f(\mathbf{k},\mathbf{r},n,t)[1-f(\mathbf{k}',\mathbf{r},v,t)], \quad (2.2)$$

here $S(\mathbf{k},n;\mathbf{k}'v;\boldsymbol{r})$ provides the scattering rate for a transition from initial $(\mathbf{k},n;\boldsymbol{r})$ to final state $(\mathbf{k}'v;\boldsymbol{r})$. These equations of the inhomogeneous transport are self-consistently coupled to the density and potential relationship, known as the Poisson equation:

$$\nabla.[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r},t)] = \rho(\mathbf{r},t) - \rho_d(\mathbf{r},t), \quad (2.3)$$

where $\phi(\mathbf{r},t)$ is the electrostatic potential, $\varepsilon(\mathbf{r})$ is the position-dependent static permittivity, and $[\rho(\mathbf{r},t) - \rho_d(\mathbf{r},t)]$ is the combined charge density from ionized dopants, acceptors, and free carriers. The spatial coordinate **r** can inhabit two or three-dimensional space depending on treating the transport problem in six or seven dimensions, including the **k**-space dimensions and time. The Poisson and BTE equations are connected by Hamiltonian mechanics, equivalent to Newton's law of motion within the formalism of classical mechanics. The derived relationship links the motion of the carriers in k-space to electrostatic potential through the following equations

$$\frac{d\mathbf{k}}{dt} = \frac{e}{\hbar} \nabla_{\mathbf{r}} \phi(\mathbf{r}, t) \tag{2.4}$$

$$\frac{d\mathbf{r}}{dt} = \frac{1}{\hbar} \nabla_{\mathbf{k}} E_n(\mathbf{k}). \tag{2.5}$$

Here, e is the charge of an electron, $\hbar$ is the reduced Planck constant and $E_n(\mathbf{k})$ is the energy-momentum dependance of the electrons. $E_n(\mathbf{k})$ is provided as a numerical table and interpolated in k-space as needed. These tabulated dispersion data are also used for the computation of the density of state and velocity. The combined charge-density of the free carriers is linked to the probability distribution function by

$$\rho(\mathbf{r}, t) = 2e \sum_{n_{cb}} \int \frac{d\mathbf{k}}{(2\pi)^3} f(\mathbf{k}, \mathbf{r}, n, t) - 2e \sum_{n_{vb}} \int \frac{d\mathbf{k}}{(2\pi)^3} f(\mathbf{k}, \mathbf{r}, n, t). \tag{2.6}$$

Here, n_{cb} and n_{vb} refer to the valence and conduction band states, respectively. For unipolar simulation with majority carriers only, the distribution function for minority carriers is approximated with the Fermi-Dirac statistics by changing the effective temperature of the local distribution. The distribution function over all bands is normalized using the ensemble particle density through

$$2 \sum_{n} \int \frac{d\mathbf{k}}{(2\pi)^3} f(\mathbf{k}, \mathbf{r}, n, t) = n(\mathbf{r}, t). \tag{2.7}$$

As can be seen above in eqs. 2.1-2.2, the BTE contains both differential and integration terms, prohibitive for analytical solutions without gross approximation such as equilibrium statistics. However, under moderately applied biases, carriers move away from the equilibrium/ thermalized distribution and occupy the high energy states [40]. The shape of these high-energy states is complex and simplified analytical formulas to present them do not exist. As such, numerically solving the BTE equation with the full dispersion relations is more appropriate to simulate the transport of the hot carriers.

2.4 Electronic Energy Bands

An electronic band structure for Si with five conduction and three valence band is employed for solving the electronic Boltzmann transport equation, presented in eq. 2.1. The band structure is derived using local empirical pseudopotential and fitted to the available experimental data, such as the density of states and energy splitting in the symmetry points for silicon [72], [73]. The band structure below 50 meV is approximated using the **k.p** method by retaining the $\mathbf{k}^4$ terms since non-parabolic corrections are accepted as valid at low energy [71], [74]. Above 50 meV, carrier transitions are computed in the irreducible unit-cell in k-space, much smaller than the first Brillouin zone (BZ), thanks to the symmetry of the silicon crystal. The energy of the particles during each transition is conserved up to $1meV/_{event}$ or smaller, depending on the density of the chosen grid points in reciprocal space. The k vectors are traced back to the first BZ by adding or subtracting the unit vectors of reciprocal space.

2.5 Calculation of Scattering Rates

The scattering rate of the carriers with crystal momentum **k** in band v is computed using Fermi's golden rule−

$$\frac{1}{\tau_s(\mathbf{k},v)} = \sum_{\mathbf{k}',v'} S_s\,(\mathbf{k}'v';\mathbf{k}v) = \frac{2\pi}{\hbar}\frac{1}{(2\pi)^3}\sum_{v'}\int d\mathrm{k}'|M_s(\mathbf{k}'v';\mathbf{k}v)|^2\,\delta[E_v(\mathbf{k}) - E_{v'}(\mathbf{k}') \pm \hbar\omega_s(\mathbf{k}'-\mathbf{k})]. \quad (2.8)$$

Here, the energy exchange with the scattering potential of type s is $\hbar\omega_s(\mathbf{k}'-\mathbf{k})]$. The delta function accounts for the energy conservation and the density of the states of the crystal. As can be seen in eq. 2.8, no momentum conservation is considered. The type of the scattering potential, denoted with s, can be attributed to phonon modes, electron-electron scattering, short- and long-range Coulomb components, ionized impurities, etc.

The matrix elements between the initial and final states, $M_s\ (\mathbf{k}'v';\mathbf{k}v)$, from a scattering potential $V_s(\mathbf{r}) = \sum_{s,q} V_{\mathbf{q}}\, e^{i\mathbf{q}.\mathbf{r}}$ are expressed using standard and Dirac notation in the following

$$M_s(\mathbf{k}'v';\mathbf{k}v) = <\mathbf{k}'v'|\boldsymbol{V_s}|\mathbf{k}v> = \frac{1}{N_c\Omega_c}\int d\mathbf{r}\, e^{-i\mathbf{k}'.\mathbf{r}} u^*{}_{k'v'}(\mathbf{r}) V_s(\mathbf{r}) u_{kv}(\mathbf{r}) e^{-i\mathbf{k}.\mathbf{r}}, \quad (2.9)$$

here, Ω_c is the volume of a single cell, N_c is the total number of cells, and $u_{kv}(\mathbf{r})$ is the periodic component of the Bloch function repeating over each unit cell. Utilizing the periodicity argument of the Bloch unit vector such that $\mathbf{r} \longmapsto \mathbf{r} + \mathbf{R}_\alpha$, and expressing $\mathbf{r}$ as a sum over unit cells indexed by α, and spanning the integration over one unit cell, we get

$$M_s(\mathbf{k}'v';\mathbf{k}v) = \frac{1}{N_c}\sum_q \sum_\alpha e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q}).\mathbf{R}_\alpha} V_{s,\mathbf{q}} \frac{1}{\Omega_c}\int dr_{cell}\, e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q}).\mathbf{r}} u^*{}_{k'v'}(\mathbf{r}) u_{kv}(\mathbf{r}), \quad (2.10)$$

here, q denotes the plane wave component of the scattering potential. The sum over α yields a nonzero value when $e^{i\mathbf{G}.\mathbf{R}_\alpha} =$ unity, where G is the reciprocal lattice vector $[\mathbf{G} = (\mathbf{k} - \mathbf{k}' + \mathbf{q})]$. The resulting matrix elements can be expressed as a discrete sum of components factored into the scattering potential and the overlap factor of the Bloch waves. Then,

$$M_s(\mathbf{k}'v';\mathbf{k}v) = \sum_G V_{s,\,\mathbf{k}'-\mathbf{k}+\mathbf{G}}\ I(\mathbf{k}'v';\mathbf{k}v,\mathbf{G}), \quad (2.11)$$

where the overlap integral, $I(\mathbf{k}'v';\mathbf{k}v,\mathbf{G}) = \frac{1}{\Omega_c}\int dr_{cell}\, e^{i\mathbf{G}.\mathbf{r}} u^*{}_{k'v'}(\mathbf{r})\, u_{kv}(\mathbf{r})$

The reciprocal vector, $\mathbf{G}_u$,which maps $\mathbf{k}' - \mathbf{k}$ to the first BZ, is the only contributing component to the sum for phonon processes. $\mathbf{G}_u$ are zero for *normal* phonon processes since the resultant wavevector resides within the first Brillouin zone. For *umklapp* processes, unit $\mathbf{G}_u$ vectors trace back the resultant wavevectors to their equivalent representation in the first Brillion zone. In theory, all components of *G* vectors may contribute to the electron-electron, ionized impurity, and other Coulomb interactions. But the squared matrix element for Coulomb interaction drops quite fast ~ G^{-4} and only the first component is kept [70]. Both

energy and momentum are conserved in *normal* processes, whereas k-vectors in *umklapp* processes are rescaled and violate momentum conservation. Previous simulations with both energy and momentum conservation yielded equivalent results to those with energy conservation alone (random-k approximation) [59], [75]. Thus, considering the computational aspects, simulation results in this study are presented with energy conservation alone.

2.6 Device Boundaries and Interfaces

Electrostatic potential in the semiconductor and insulator contacts, equivalent to their applied biases, acts as an electrostatic boundary condition for solving the Poisson equation. Contacts in the device are defined as lines and areas respectively for 2D and 3D transport. In addition to establishing boundary conditions for the Poisson equation, contacts on the semiconductor set the specifications for solving the BTE. In doing so, thermal equilibrium and charge-neutrality in the regions close to the contacts are ensured by injecting/removing carriers through contacts. Injection or removal of the carriers takes place only if the net error in charge neutrality is reduced as a result. Carriers are injected randomly from a local Fermi distribution. The local Fermi distribution is a function of the temperature and target charge density in the mesh nodes adjacent to contacts. The selection of a wave-vector and band index for the injected carrier adheres to the following probability density function

$$P_{th}(\mathbf{k}, \upsilon) \propto f_0(k, \upsilon, T_0, E_F)\, \theta[\mathrm{v}^{g\mathrm{n}}{}_{\upsilon}]\, |\mathrm{v}^{g\mathrm{n}}{}_{\upsilon}|, \quad (2.12)$$

where $f_0(\mathbf{k}, \upsilon, T_0, E_F)$ is the Fermi-Dirac distribution function in the non-degenerate regime for temperature T_0, band n, and Fermi level E_F. A Heaviside function, $\theta[\mathrm{v}^{g\mathrm{n}}{}_{\upsilon}]$, selects the portion of the distribution with the group velocity component directed into the device. The normal component of the group velocity, $|\mathrm{v}^{g\mathrm{n}}{}_{\upsilon}|$, helps match the incoming flux of the injected particles to that of the existing thermal distribution near the contacts. Particle currents are computed in each semiconductor contact by tallying the flux of the injected vs.

removed carriers over time. Floating material edges, which are not defined as contacts, are treated as Neumann boundary conditions (fixed value of the function derivative along the edges). The derivative of the electric field in the normal direction of the Neumann boundary is set to zero and assists in solving the Poisson problem.

The semiconductor-oxide interfaces are treated as ideal (without surface roughness, tunneling, and oxide/interface traps), and carriers are secularly reflected from these interfaces. During this reflection, only the sign of the group-velocity component in the direction perpendicular to the interface gets reversed. No diffusive transport or thermalization of the carriers in the semiconductor-insulator interfaces is considered.

2.7 Initial Conditions

MC particles are initially allotted in real space by the net local dopant density of a region. All donors and acceptors are considered fully ionized from the start, which is considered an acceptable approximation for room temperature ionization of the dopants. The resultant density of the free carriers is used to compute the local Femi-level. In the end, each carrier is provided a random **k** and band index ν, selected from the Fermi-Dirac probability density function. Upon restarting a device simulation from existing solutions, the initial distribution functions of the carriers are read from the existing simulations.

2.8 Ensemble Monte Carlo Method for Solving BTE

A stochastic Monte Carlo technique is employed for synchronously solving the BTE for an ensemble of carriers [73]. Using an ensemble instead of a single particle facilitates the conservation of average energy and average momentum. In addition, interparticle interactions such as electron-electron scattering can be realistically treated with multiple carriers present [61]. Ensemble particles are dispersed to the real and reciprocal space mesh according to an initial distribution. As mentioned in the earlier section, the initial distribution

of the carriers is derived using the net charge density from dopants or an existing solution if available.

The following Newtonian equation of motion sets each particle within the ensemble to a free (ballistic) flight in the real and k space for a predefined duration, called MC-timestep (Δt_{MC}).

$$\mathbf{k}(t + \Delta t_{MC}) = \mathbf{k}(t) + \frac{\Delta t_{MC}}{2}(\boldsymbol{p_1} + \boldsymbol{p_2}) \tag{2.13}$$

$$\mathbf{r}(t + \Delta t_{MC}) = \mathbf{r}(t) + \frac{\Delta t_{MC}}{2}(\boldsymbol{x_1} + \boldsymbol{x_2}) \tag{2.14}$$

Here, subscripts 1 and 2 correspondingly associate the quantities at time t and $(t + \Delta t_{MC})$. The position and momentum are derived using the eqs. 2.4-2.5, respectively, from the electronic energy bands and the obtained potential using the Poisson equation. The motion of the particles is numerically computed since no analytical expression reasonably represents the force or dispersion well.

After a predefined number of ballistic time steps, the probability of each particle scattering is estimated from the total scattering rate ($\frac{1}{\tau}$). The probability of a particle scattering after Δt_{SC} is

$$P_{sctr} = 1 - exp\left\{-\int_{t}^{t+\Delta t_{SC}} dt \frac{1}{\tau[\mathbf{k}(t), v]}\right\}. \tag{2.15}$$

Δt_{SC} is chosen small such that $\mathbf{k}(t + \Delta t_{SC}) \approx \mathbf{k}(t)$. Then, the scattering probability for each particle becomes

$$P_{sctr} \approx \frac{\Delta t_{SC}}{\tau[\mathbf{k}(t+\Delta t_{SC}), v]}. \tag{2.16}$$

Consequently, using eqs. 2.8-10, the total scattering rate for each particle is evaluated on their final wavevector $\mathbf{k}$ (after time Δt_{SC}) and band index n. The ratio of Δt_{SC} to the mean time for scattering (τ) is compared to a random number (ζ_1) for each particle. If $\zeta_1 < \frac{\Delta t_{SC}}{\tau}$,

then a scattering of type s may happen when $P_{s-1} \leq \zeta_1 \leq P_s$, where P_s is the cumulatively calculated scattering rate up to type s —

$$P_s = \sum_{r=1}^{s} \frac{\Delta t_{SC}}{\tau_r(\mathbf{k},v)}. \quad (2.17)$$

Once a particular scattering process is selected, two additional random numbers are generated to select a final state from the numerical energy band. All energy-conserving shells with wavevector $\mathbf{k}_j$ and band v_j are considered as an acceptable final state. The probability of scattering to each state is assigned according to the square of the matrix element and density of states —

$$P_j = |M_s(\mathbf{k}_j \mathrm{v}_j; \mathbf{k}v)|^2 \mathcal{D}os_j. \quad (2.18)$$

Here, $\mathcal{D}os_j$ is the density of states associated with j^{th} electronic state. If N_{tot} number of states satisfy the energy conservation threshold, and then a random state j is selected using a random number ζ_2 such that $j = \zeta_2 N_{tot}$. The transition to final state j is ultimately accepted using an additional random number ζ_3 if $P_j > \zeta_3 P_{max}$, where $P_{max} = \max[\langle P_j \rangle_{j=1,N_{tot}}]$. Degeneracy effects are considered before approving the transition to the selected final state.

2.9 Assigning Particle Properties to Mesh

Quantities of interest are computed from the particle location in phase space after computing the scattering rates and statistically selecting the scattering events following the free flights. Attributes (e.g., charge, energy) of the particles are mapped to the adjacent Poisson mesh element (smallest cube of mesh nodes in real space) in the simulation space. This allows the evaluation of average charge density, average current density (charge density × drift velocity), crystal momentum, normal components of the group velocity, potential, average statistical weight, etc. In addition, each particle can be assigned a variable weight on their charge based on their location in real coordinates to enhance rare events. This statistical

enhancement technique allows us to assign carriers to low-carrier density regions and investigate the transport around the tail of the energy distribution. The extent of the enhancement is considered during (1) charge-density is assigned to the mesh elements from the particle distribution and (2) current is computed at the contacts from particle injection/removal.

2.10 Solution of the Poisson Problem

A finite-difference tensor mesh is used for solving the Poisson equation to calculate the potential. The unwanted portion of the computational mesh is defined with a very low dielectric constant, close to that of vacuum. The quantities required to solve the Poisson equation are (1) free carrier density, (2) ionized dopant density, (3) boundary conditions on the potential, and (4) piecewise dielectric constant. For the bipolar simulation presented in ***chapter III***, free carrier density for both majority and minority carriers is obtainable from the particle distribution in the device. For unipolar simulation with only one type of carriers

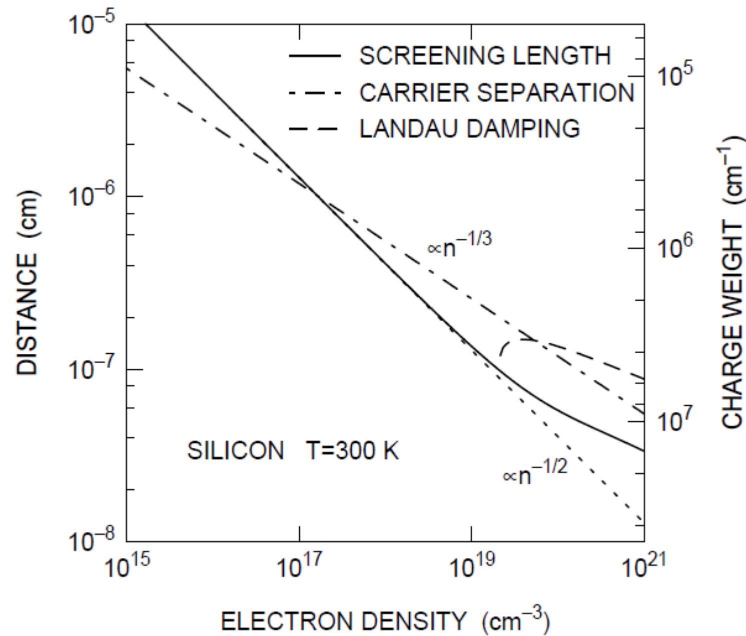

Fig. 1. Optimum separation of carriers (left axis) and charge wight (right) for treating long range plasma scattering. The characteristic length for Landau damping become smaller at high densities and deviates from $n^{-1/2}$ behavior due to degeneracy. Fig. is adapted from ref. [171].

represented as particles, presented in ***chapter IV***, the density of the minority carriers (holes) is calculated using a constant quasi-Fermi level approximation. For such cases, density is calculated by multiplying the Fermi-Dirac integral by the effective density of states in each valence band. Potentials are defined at the mesh points, and the resultant force on the particles is interpolated from the adjacent nodes.

2.11 Coulomb Scattering

Mobile carriers exposed to the background jellium of oppositely charged ions may form standing waves and jointly oscillate because of the resultant (background) electric field [76]. In quantum formulation, a single unit of these collective (plasma) oscillations is called a plasmon, and they can exchange energy with a single particle like photons/phonons. If the average displacement (phase velocity) of the particles is smaller than the plasma oscillation, then the energy is damped from the wave to the particles and vice versa [46], [77]. Particles with velocities very different from the plasma wave do not participate in these energy exchanges. The highest damping occurs when the wave's phase velocity coincides with the electrons' thermal velocity. Consequently, this mode is not sustainable. Plasma oscillations are spatially resolved by choosing mesh sizes small enough so that the shorter wavelength of the collective wave is damped properly and are resolved in the time domain by updating the resultant change in potential quickly enough, according to the Nyquist theorem.

For silicon, as shown in Fig. 1, the characteristic length for Landau damping becomes smaller than the Debye screening length in degenerate densities. The smallest length between the two is chosen to resolve the plasma oscillations accurately for a specified carrier density. In order to treat plasma scattering in large-area diodes with moderate doping, as presented in chapter III, mesh sizes follow the Debye curve. For a short-channel MOSFET device simulation presented in chapter IV, the Landau curve is chosen accordingly in view of the large doping in the access region. The selected mesh spacing also acts as a boundary between short- and long-range interaction. Short-ranges forces are active when the particles are within

the screening lengths of each other — inside the borders of the same mesh element. The component of the long-range forces diminishes if the interacting particles reside within the same mesh element.

The oscillating electrons in the frequency domain create fluctuations in charge densities [78]. Subsequent potential instabilities may elevate electrons to high energies in regions where the potential becomes depressed. In addition, degeneracy heating elevates electrons to high energies in regions where the density becomes higher due to collective motion. Free carriers in semiconductors oscillate with the following plasma frequency

$$\omega_p = \left(\frac{e^2 n}{\epsilon_s m_c}\right)^{1/2} . \tag{2.19}$$

Here, ϵ_s is the permittivity of the material, n is the electron density and m_c is the conductivity mass. To circumvent the potential fluctuation and associated plasma catastrophe, potentials are updated rapidly enough — quicker than $1/2\omega_p$ seconds, approximately on the same order of the frequency of scattering events ($\frac{1}{\Delta t_{MC}}$). For a dopant density of $2 \times 10^{20} cm^{-3}$, potentials are updated every $1/10\omega_p\,(s) \sim 0.2\,fs$ while simulating the GAA devices in ***chapter IV***. Plasma effects are negligible for the moderate doping of the work in ***chapter III*** and were not considered there [38], [39].

A final consideration in treating plasma interactions is the charge weight of the particles in the simulation, and tackling it well is a major numerical struggle [79]. When particles are assigned with sub-electron charges, the potential fluctuation from plasma-electron interaction gets depressed as interparticle distances become shorter than that of the real electrons. In this scenario, the movement of a few of them to adjacent mesh elements does not significantly affect the potential — the charge-weights of the electrons are too small. To avoid this, particles can be assigned charges similar to that of an electron. But in doing so, collecting sufficient statistics with only hundreds of electrons requires unrealistically large

computational overhead, not yet manageable with modern computing clusters. As such, the short-channel device $(24\,nm\ \times 8nm \times 8nm)$ in ***chapter IV*** is simulated with particles presenting charges ~ ten times smaller than that of an electron.

Electrons in semiconductors respond to charge perturbation and the Poisson equation is the connection between the charges to the field (potential) [73]. But the situation is not analogous to Coulomb interaction of two-point charges in a vacuum [39]. The long-range force between two-point charges (electrons) in a semiconductor is greatly reduced due to the shielding of the electric field by the remaining electrons in the jellium. In other words, the electric field is reduced due to the repulsion of the electrons by each other, creating a shielding effect at larger distances [39], [78]. The phase shift due to the attractive and repulsive for are accounted for in the code within the semiclassical framework [70]. The subsequent reduction in the electric field can be included in the modeling by using the screened Poisson equation —

$$-\nabla_{\mathbf{r}}.(\varepsilon\nabla_{\mathbf{r}}\varphi) + \beta_s^2 = Q\delta(\mathbf{r}) \tag{2.20}$$

Here, β_s is the screening parameter, and $\delta(\mathbf{r})$ is the Kronecker delta function. The solution to eq. 2.20 introduces an exponential damping factor for the potential via β_s term —

$$\varphi(\mathbf{r}) = \frac{Q_{point}}{4\pi\varepsilon|\mathbf{r}|}e^{-\beta_s|\mathbf{r}|}. \tag{2.21}$$

Since electrons are interacting fermions, the correlation between them elevates the average energy of the ensemble, especially at high densities [80]. These many-body effects are accounted for in the simulation by using the modified *Hartree* model [81] —

$$< \Delta E > (n) = 0.90\,(\frac{4\pi n}{3})^{1/3}\frac{e^2}{4\pi\varepsilon} \tag{2.22}$$

Here, *n* represents the density of free carriers, accepted equal to that of the dopants. The resultant change in the average energy is greater than the thermal energy (*kT*) in the high-

density regime, and the carrier distribution is updated accordingly by changing the effective temperature of the Fermi-Dirac distribution.

The scattering rate for short-range Coulomb force is computed considering that every particle can scatter with the remaining ensemble within its screening parameters. The screening radius (β_s^{-1}) is computed from the electron density at every $\Delta_{t,MC}$ interval —

$$\beta_s^2 = \frac{e^2}{\varepsilon_s}\frac{\partial n(\mathbf{r},t)}{\partial E_F(\bar{E},\mathbf{r},t)} \tag{2.23}$$

In eq. 2.23, n is the local density, ∂E_F is the local Fermi level at n and average energy $\bar{E}$. Each plausible pair within the screening volume is considered using the theoretical framework of ref. [61], [61]. As scattering rate rises with the increasing density of the electron gas, interactions are dynamically screened such that the screening parameter (β) become a complex function of both energy (ω) and wavevector (q). Then, the resultant dielectric function become —

$$\varepsilon(q,\omega) = \varepsilon_s[1 + \frac{\beta(q,\omega)^2}{q^2}]. \tag{2.24}$$

The detailed dependance of the dynamic parameter (β) to the wavevector, valley parameters, and components of plasma frequency is provided in ref.

A condensed theoretical discussion on the rate for ionized impurity and electron-electron scattering is presented here so that results presented later in ***chapter V*** can be clearly understood. The impurity scattering rate for an electron with wavevector **k** of the crystal momentum in band v and at location **r** is expressed as

$$\frac{1}{\tau_{imp}} = \frac{2\pi}{\hbar} N_{imp} \sum_{G,v'} h(y\pm,s) \int \frac{\partial \mathbf{k}'}{(2\pi)^3} |M(\mathbf{k}v,\mathbf{k}'v',G)|^2 \delta[E_v(\mathbf{k}) - E_v(\mathbf{k}')][1 - f(\mathbf{k}',v',\mathbf{r})],$$

and the matrix element, $|M(\mathbf{k}v,\mathbf{k}'v',G)| = \frac{Ze^2}{\varepsilon_s}\frac{I(\mathbf{k}'v';\mathbf{k}v)}{|(\mathbf{k}-\mathbf{k}'+\boldsymbol{G})^2+\beta^2(q,0)|}$. (2.25)

Here, G is the reciprocal-lattice wavevector, N_{imp} is the density of ionized impurities, $I(\mathbf{k}'v';\mathbf{k}v)$ is the overlap between the Bloch states, delta function accounts for energy conservation, $(1-f)$ term addresses the degeneracy, and β_s is the static screening parameter. In addition, h tackles for the phase shift correction as a function of $y \sim k$ and $s \sim \beta_s^{-1}$; both y and s are rescaled to ensure that the excess charge from the impurity is accounted for. The scattering rate for electron-electron scattering is similar to eq 5.1, except (i) the matrix element includes double overlap integrals, one for each scattered electron, (ii) screening parameter is dynamic and depends on both energy and momentum, and (iii) sum of the scattering probability is calculated using triplet state including the spin effects on symmetric and antisymmetric matrix element, similar to [70].

Band to band tunneling due to the short-range force is considered, and the theory on the treatment of the impact ionization process is provided in ***chapter III***. Ionized impurity scattering is similarly treated within the first-Born approximation, using static screening parameter with wavevector dependence only, and tabulated phase-shift parameters. The treatment of the phonon scattering with deformation potentials and overlap integrals is discussed in chapter ***III***. Flat dispersion is assumed for the optical branches, whereas the acoustic branches are treated as isotropic. Linear dispersion is assumed for the acoustic branches for small q, and tabulated values for the energy of are provided the short-wavelength phonons. Static screening is used up to the plasma frequency.

Chapter 3

Interaction of Hot-carriers with Cold-Carriers and Lattice

3.1 Preface

A physics-based justification of the ionization energy in the semiconducting materials had long been overdue in both the microelectronics and physics communities. Ionization energy in semiconductors has been determined with high precision by experimental means -- an empirical relation for energy-charge conversion is available from the time of W. Shockley [82]. The experimentally derived ionization energy is found to be exclusively associated with the material's bandgap. However, carriers in very high energies mostly experience nuclear interaction events that are not related to the band gap of the materials. The carriers' trajectories are invariant to the dispersion (E vs. k) characteristic above ~ 10-20 eV energy since the available energy-states become almost continuous at high-energies, like a free particle moving in the air. As such, the connection of the ionization energy to a material's property (bandgap) is likely to be concealed in the energy-exchange processes within a few eV beyond the band edge. To verify this, a semiclassical simulation of the thermalization of EHPs from 5 eV to the band edge were performed.

This collaborative work has been done by several researchers from i) EECE, Interdisciplinary Materials Science, and Physics Department of Vanderbilt University ii) Materials Science Department of the University of Texas at Dallas and iii) ECE Department of Libscomb University in Nashville. My individual contribution was to simulate and validate the energy exchange processes initiated by energetic electrons and holes. The findings of this work are published in [83].

3.2 Relaxation of Hot-carriers in Materials

Interaction of hot-carriers with cold-carriers and lattice during the electron-hole pair (EHP) generation processes is important for determining the amount of free charge liberated by the

ionizing particles, avalanche breakdown in p-n junctions [82], and photoemission efficiency [84]. Energy release from the ionizing particles to free carriers in semiconducting materials has been historically investigated by experimental techniques alone [85]. Simulating the energy deposition events in the entire range (MeV – 0.25 eV) is challenging because the physics of carriers navigating above and below ~ 10 eV (approximately) is different. In the low-energy regime, the generated electrons/ holes are influenced by the non-parabolic shape of the energy bands, impact ionization, and phonon processes [44], [86].

This study aims to show that the energy required to generate EHPs in semiconductors can be explained by free carriers' interactions only up to a few eV above the band edge. The average energy lost (3.69 eV in Si and 2.62 eV in Ge) by the incident high-energy particles to create a thermalized electron-hole pair, obtained from the simulations, is very close to the experimentally measured radiation-ionization energies of Si and Ge. Indeed, an excellent agreement between this simulation results and the experimentally derived carrier-ionization energy demonstrates that accounting for phonon process and impact ionization events at the low energies are sufficient to understand energy-charge conversions. In particular, the radiation-ionization energy for semiconducting materials exclusively depended on the electron transport mechanisms below ~5 eV energy or less. Overall, this study validates the treatment of hot carriers within the semiclassical framework, especially since there are no fitting parameters in the ensemble Monte Carlo approach to solving the Boltzmann transport equation.

It is well understood that around 1/3 of the ionization energy is absorbed by valence electrons for jumping to the conduction band and creating an electron-hole pair. In addition to an energy loss equal to the band gap energy via impact ionization, acoustic-phonon emission, which has been omitted in prior work, contributes 30% of the remaining carrier-energy loss, while optical-phonon emission contributes the other 70%. The term ionization in this section

refers to the creation of nonequilibrium free charge carriers in a semiconductor after irradiation with high-energy fermions/ bosons.

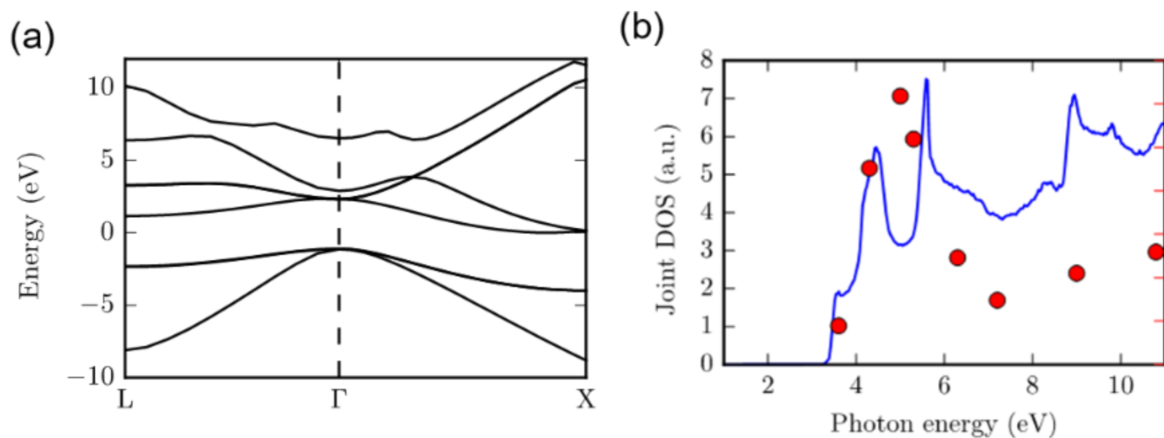

Fig. 2. (a) The electronic band structure of silicon calculated with empirical pseudopotentials. Five conduction bands and three valence bands are shown. The conduction band minimum along the path Γ → X is set as the reference energy of zero. Degeneracy of electronic states is included. (b) The joint density of states for photon absorption with different incident photon energies in silicon is shown with the solid blue line. Only vertical transitions are considered for photon absorption to create electron-hole pairs. The electron-hole-pair creation energies for silicon devices excited with different photon energies studied below are indicated with the red dots.

3.3 Electron-Hole Pair Creation Energy: Experimental Findings

It is generally accepted that the average energy required to create a thermalized EHP is approximately three times the band-gap energy, based on experiments conducted on a wide variety of semiconductors [82], [84], [87], [88]. In silicon, with a band gap of 1.12 eV at room temperature, the average energy to produce an electron-hole pair is about 3.6 eV (or 22.5 keV/fC) for an ionizing event [82], [87]–[89]. Lowe and Sareen have reported experimental measurements of the temperature dependence of the electron-hole-pair creation energy with X-rays [90]. For germanium, Antman *et al.* have measured a value of 2.98 eV for the average energy per electron-hole pair at 77° K [91] and Klein has reported 2.80 eV at room temperature [92]. Subtracting the backscattering of 25% of the incident electron-

beam energy, the effective energy transferred is 2.10 eV, which is about three times of the band gap at this temperature, 0.74 eV.

3.4 Thermalization Process of High-Energy Carriers

Various research teams have developed Monte Carlo radiation transport methods for computing energy loss of charged particles in semiconductor materials; in Ref. [93] 26 researchers from 16 different organizations provide a summary of each of these tools. Physical processes included in these tools range from simple transport via electronic stopping to nuclear-interaction processes [93]. These tools are typically limited to ballistic transport of the incident and secondary particles, i.e., particle energies > 250 eV. Therefore, they do not directly consider the EHP thermalization process. Instead, the energy deposited by the incident particle is converted to free charge using the experimentally measured radiation-ionization energy. Two groups have developed tools and techniques for transporting electrons down to ~20 eV [94], [95]. However, these approaches do not consider the impact of the electronic band structure and the carrier-phonon scattering on electron transport which are indeed critical to create the thermalized carriers in the final stage of particle energy deposition. The Monte Carlo carrier transport tool used in this current work utilizes the full electronic band structure to simulate carrier motion between 10 eV and the band edge, which allows for estimation of the EHP creation energy.

3.5 Background: Energy Loss Mechanisms

It has been reported that carriers lose their energy mainly via the emission of optical phonons and impact ionization. Shockley made the "assumption of uniform population" and proposed the simple model equation $E_{\mathrm{pair}} = 2.2E_{\mathrm{i}} + rE_{\mathrm{R}}$ where E_{I} is an impact-ionization threshold energy (1.1 eV for silicon [82]) and rE_{R} is the average optical phonon energy loss between impact ionization events. The 2.2 multiplying factor of E_{i} is empirically determined by a fit to experimental data and considers the energy loss associated with the band gap and the residual energy of the generated particles. Here $r = L_{\mathrm{ii}}/L_{\mathrm{R}}$, where L_{ii} is the impact-

ionization mean free path (~ 88 nm for silicon [82]) and L_R is the electron-optical phonon scattering mean free path (~ 4 - 6 nm for silicon [82], [88]) and E_R is the optical phonon energy (~ 0.063 eV for silicon) [82]. Klein proposed a similar phenomenological model for the pair-creation energy as $E_{pair} = 2.8E_g + rE_R$, assuming that the intrinsic band gap, E_g, optical phonon losses, and the residual kinetic energy contribute to the energy consumption when generating EHPs [88]. Again, the 2.8 multiplying factor of E_g is empirically determined by a fit to experimental data. Alig *et al.* performed theoretical calculations for the pair-creation energy based on the "scattering rate assumption". A free-particle parabolic band structure was used and a similar expression was derived in Alig's approach [87]. Unfortunately, in all of these reports, the energy loss associated with further relaxation of secondary charge carriers with energies below the threshold energy E_i down to the thermal energy range has not been reported.

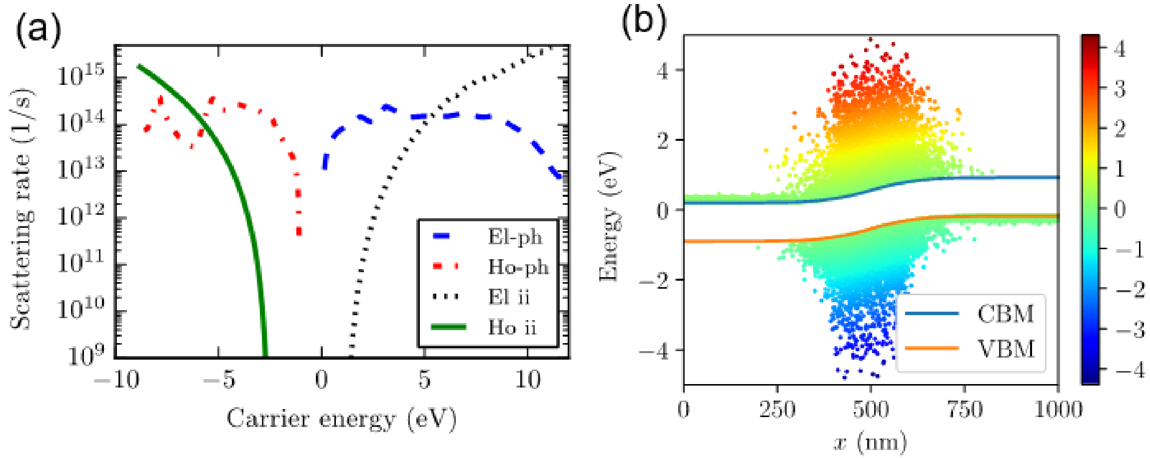

Fig. 3. (a) Carrier-phonon and impact ionization rates for electrons and holes. The conduction-band minimum is set at 0. By convention, the electron (hole) energy is assigned a positive (negative) sign. (b) The spatial distribution of Monte Carlo particles superimposed on the energy band diagram under illumination of 6.3 eV photons. The colormap indicates the carriers' kinetic energy in eV. By convention, electrons are assigned a positive kinetic energy, whereas holes are assigned a 'negative' kinetic energy sign.

3.6 Synopsis of the Work

The energy-exchange mechanisms of the induced hot-carriers from photon absorption, and electron- and hole-initiated impact ionization events are studied to investigate the pair creation processes in Si and Ge [71]. A high-fidelity physics code, presented earlier in section II, is used to simulate the carrier dynamics of the sub-10 eV energetic particles in silicon and germanium. The introduced hot-carriers distribution in the device mimics the experimentally derived results from incident photons or impact ionization events. The Boltzmann transport equation is solved using a synchronous ensemble Monte Carlo technique, along with the full electronic band structure (*i.e.,* multiple bands energy dispersion) to obtain the carrier kinetics (e.g*.,* energy-momentum relation and group velocity) and tabulate the thermalization steps of the hot carriers. The scattering mechanisms are examined in detail, accounting for the energy dependence of carrier-phonon scattering rates and the impact-ionization rates. Similar to the experimental findings, the ionization energy of the EHP generation processes is observed to be independent to the source of excitation (e.g., photon absorption, impact ionization by electrons or holes).

We show that the experimental values of the radiation-ionization energies can be reproduced using FBMC carrier-transport simulations. Simple isolated p-n junctions similar to those used in the experiments are modeled. The photocurrents at the contacts of the simulated diodes are monitored during and after excitation. The collected charge and thus the quantum efficiency are derived from the photocurrents. These simulations demonstrate that the final stage of the carrier relaxation via carrier-phonon scattering and impact ionization, for carrier energies less than 5 eV (i.e., electrons or holes occupying the lowest energy bands), completely accounts for the experimentally determined values of the radiation-ionization energy for Si and Ge. We find that, in addition to an energy loss equal to the band gap energy via impact ionization, acoustic-phonon emission, which has been omitted in prior work,

contributes 30% of the remaining carrier-energy loss, while optical-phonon emission contributes the other 70%.

This chapter is organized as follows. Sec. 2.7.1 briefly describes the method used to solve the Boltzmann transport equation with full electronic band structure. The employed models for calculating the electron-phonon scattering and impact ionization rates are discussed in Sec. 2.7.2. Finally, based on the Monte Carlo simulations, results regarding pair creation events are presented in Sec. 2.8.1, 2.8.2, and 2.8.3 for Si and in Sec. 2.8.4 for Ge. The electron-hole pair creation energies for the excitation by sub-10 eV energetic particles are compared. Conclusions are presented in Sec. 2.9.

3.7 Full band Monte Carlo simulations

3.7.1. Boltzmann transport equation

The Boltzmann transport equation, namely

$$\frac{\partial f(E_n(\mathbf{k}),\mathbf{r},t)}{\partial t} = -\frac{\mathrm{d}\mathbf{k}}{\mathrm{d}t}\cdot\nabla_{\mathbf{k}} f(E_n(\mathbf{k}),\mathbf{r},t) - \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\cdot\nabla_{\mathbf{r}} f(E_n(\mathbf{k}),\mathbf{r},t)$$

$$+ (\partial f(E_n\ (\mathbf{k}),\mathbf{r},t)/\partial t)_coll \qquad (3.1)$$

describes the spatial and temporal change of the carriers' distribution function, $f(E_n(\mathbf{k}),\mathbf{r},t)$, which depends on the carriers' energy $E_n(\mathbf{k})$, position $\mathbf{r}$, and time t [13]. It is based on the overall changes originating from the processes of carrier drift, diffusion, and scattering. The carriers' energy, $E_n(\mathbf{k})$, depends on the electronic band index, n, and crystal momentum, $\mathbf{k}$, which is well described using the full electronic band structure. The carriers drift under the effect of the electric field, resulting in a change of their momenta. Given an electrostatic potential distribution throughout the device, the resulting electric field and the carriers' group velocity, $\partial E_n(\mathbf{k})/\partial\mathbf{k}$, are used to solve the equations of motion for the carriers' free flight between scattering events. Carriers diffuse when a carrier-concentration

gradient is present. Carriers can also scatter with phonons with a resulting change of their electronic states [71].

The implementation of the Monte Carlo computer program, called Anduril, used here (similar to the Damocles code in [13], [71]) is based on an ambient temperature of 300 K, i.e., the lattice is equilibrated at room temperature, which sets the phonon distribution for the calculation of carrier-phonon scattering rates. The energy bands are based on the empirical pseudopotentials fitted to the experimental room-temperature band gap. The calculations employ ten conduction bands and three valence bands, as shown in Fig. 2(a) for silicon. Note that the lowest conduction bands and the highest valence bands are twofold degenerate. The use of five conduction bands allows us to treat accurately sub-10 eV electron dynamics and the three valence bands allow us to treat sub-8eV hole dynamics. The densities of states for both conduction bands and valence bands are calculated with the Gilat-Raubenheimer algorithm [96]. The joint density of states for photon absorption that controls the probability of creating vertical transitions is shown in Fig. 2(b). The Monte Carlo simulation step is set at 0.2 fs, so that all the scattering events discussed in Sec. II B can be monitored. Dirichlet boundary conditions (value of the solution of the differential equations are provided along edges) are applied to the contacts. At the device boundary, specular reflection is assumed.

Table I. The impact-ionization model parameters for electrons and holes in silicon and germanium. The threshold energy is relative to the conduction and valence band edge for electrons and holes, respectively.

	electrons			holes					
	P (1/s)	α	E_{th} (eV)	P_1	α_1	E_{th1}(eV)	P_2	α_2	E_{th2} (eV)
Si	2×10^{11}	4.6	1.11	2.0×10^{9}	6.0	1.11	1.0×10^{12}	4.0	1.45
Ge	1.66×10^{11}	5.5	0.66	4.75×10^{11}	4.5	0.66			

3.7.2. Carrier Scattering Mechanisms

Two dominant scattering mechanisms (multiple carrier-phonon scattering and impact ionization) are considered here in the sub-10 eV carrier transport in lightly doped devices. Carriers can scatter with longitudinal acoustic (LA) phonons, transverse acoustic (TA) phonons, and nonpolar optical phonons (NOPs). Both phonon-emission and phonon-absorption processes are considered, and the corresponding electron energy loss or gain is accounted for by changing the carrier energy by the emitted or absorbed phonons using the full phonon dispersion relation [71]. The carrier-phonon scattering rates at room temperature are calculated using the deformation potentials given in Ref. [71]. The impact-ionization rate is treated using the Keldysh formula with parameters previously calibrated to experimental data [13], [97], [98]. Thus, the impact-ionization rate for an electron of energy E is:

$$\frac{1}{\tau_{\text{ii}}(E)} = \begin{cases} 0, & E < E_{\text{th}} \\ P \times \left(\frac{E - E_{\text{th}}}{E_{\text{th}}}\right)^{\alpha}, & E \geq E_{\text{th}} \end{cases} \tag{3.2}$$

where E_{th}, the ionization threshold energy (which is approximately one energy gap above the conduction band minimum), and P and α are model parameters. In Eq. (3.2) the electron energy is measured from the conduction-band minimum. The impact-ionization rate for holes is expressed similarly as:

$$\frac{1}{\tau_{\text{ii}}(E)} = \begin{cases} 0, & E < E_{\text{th1}} \\ P_1 \times \left(\frac{E - E_{\text{th1}}}{E_{\text{th1}}}\right)^{\alpha_1}, & E_{\text{th2}} \geq E \geq E_{\text{th1}} \\ P_1 \times \left(\frac{E - E_{\text{th1}}}{E_{\text{th1}}}\right)^{\alpha_1} + P_2 \times \left(\frac{E - E_{\text{th2}}}{E_{\text{th2}}}\right)^{\alpha_2}, & E \geq E_{\text{th2}} \end{cases} \tag{3.3}$$

where P_1, P_2 α_1, α_2, E_{th1} and E_{th2} are all model parameters [13], [97], [98]. In Eq. (3.3) the hole energy is measured from the valence-band maximum. Table I lists the well-calibrated parameters used in the simulations. Equations (2) and (3) represent empirical fits to extensive calculations of impact-ionization rates for electrons and holes in several semiconductors using the full band structure obtained from empirical pseudopotentials and screened

Coulomb matrix elements and experimental data [13], [97], [98]. Figure 3(a) compares the total carrier-phonon scattering rates (for both acoustic and optical phonons) and impact ionization rates for electrons and holes of different energies in silicon. Above 5.3 eV (relative to the conduction band minimum or valence band maximum), either for electrons or for holes, the impact ionization processes dominate, compared with the carrier-phonon processes. Kane has shown that the free-electron model is not valid for calculating the impact ionization rate for electrons with energies below 7 eV and that the full electronic band structure must be considered [99].

Table II. Quantum Efficiency Dependence on the Incident Photon Energy in Silicon

$E_{photon}(eV)$	$N_{carriers}(cm^{-1})$	$QE = \frac{N_{carriers}}{N_{photons}}$
3.6	1.758×10^{7}	1.005
4.3	1.904×10^{7}	1.088
5.3	2.308×10^{7}	1.325
6.3	2.943×10^{7}	1.682
7.2	3.425×10^{7}	1.957
9.0	4.245×10^{7}	2.426
10.8	5.027×10^{7}	2.873

$N_{carriers}$: number of collected charge particles at contacts.

$N_{photons}$: total # of injected photons ~ 1.75×10^{7} cm^{-1}.

3.8 Simulation Results

This section describes two-dimensional simulation results for three primary modes of carrier excitation in p-n diodes, namely by photons, energetic electrons, and energetic holes. The silicon p-n diodes have a length of 1 μm along the x-direction and a width of 0.25 μm along

the y-direction. The donor and acceptor concentrations in the n-type and p-type regions are 10^{16} cm^{-3}, respectively. The diode terminals are grounded in the simulations. In this equilibrium state, a built-in potential of 0.72 V induces a depletion width of 0.4 μm. Carrier transport and pair creation are monitored in the diode after exciting part of the depletion region from 0.4 μm to 0.6 μm, which is the most charge-sensitive part of the device.

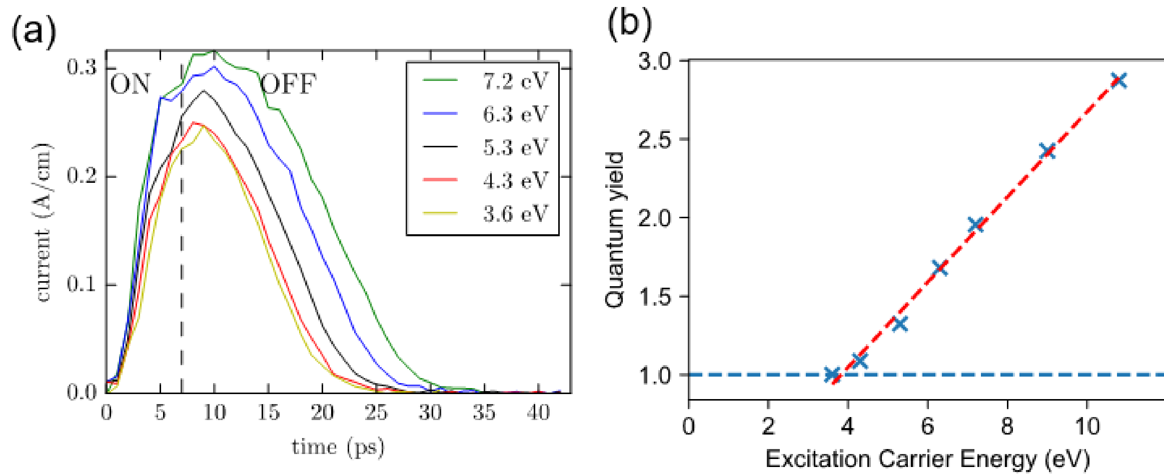

Fig. 4. (a) Calculated photocurrent collected on the left contact for various photon energies. In all cases, the illumination starts from equilibrium at t = 0, and lasts until t = 7 ps, as indicated by the dashed line. Here the currents are averaged over a time span of 1 ps for each moment. (b) The dependence of the quantum efficiency on the incident carrier energy. The horizontal blue dashed line indicates the unity of the quantum efficiency. The red dashed line indicates the fitted linear relationship between the quantum efficiency and the incident particle energy. The inverse of the slope corresponds to the EHP creation energy of 3.69 eV for silicon. Quantum efficiencies for photon energies below 3.6 eV have not been simulated.

3.8.1. EHP production in silicon via photoabsorption

We illuminate the devices with photons of different energies as listed in Table II and calculate the corresponding collected charge and quantum efficiency (QE). The simulated photon illumination lasts for 7 ps, after which the carriers undergo thermalization. As an example, the spatial distribution of the particles, aligned on top of the band diagram is shown in Fig. 3(b). The energetic secondary electrons and holes generated by the primary high-

energy (relative to the band gap) photon illumination can be seen. The EHPs are produced both in the primary collisions when the photon is absorbed and by the subsequent energy losses of the secondary EHPs. The generated electrons and holes in the depletion region drift, diffuse, or scatter after the photon absorption. To obtain accurate carrier statistics in the depletion region, in which the carrier density is orders of magnitude lower than the carrier density in the end regions, we employ a statistical enhancement technique [100]. A given number of Monte Carlo particles are assigned in the depletion region and each Monte Carlo particle represents a different number of electrons compared to what the particles in the end regions represent.

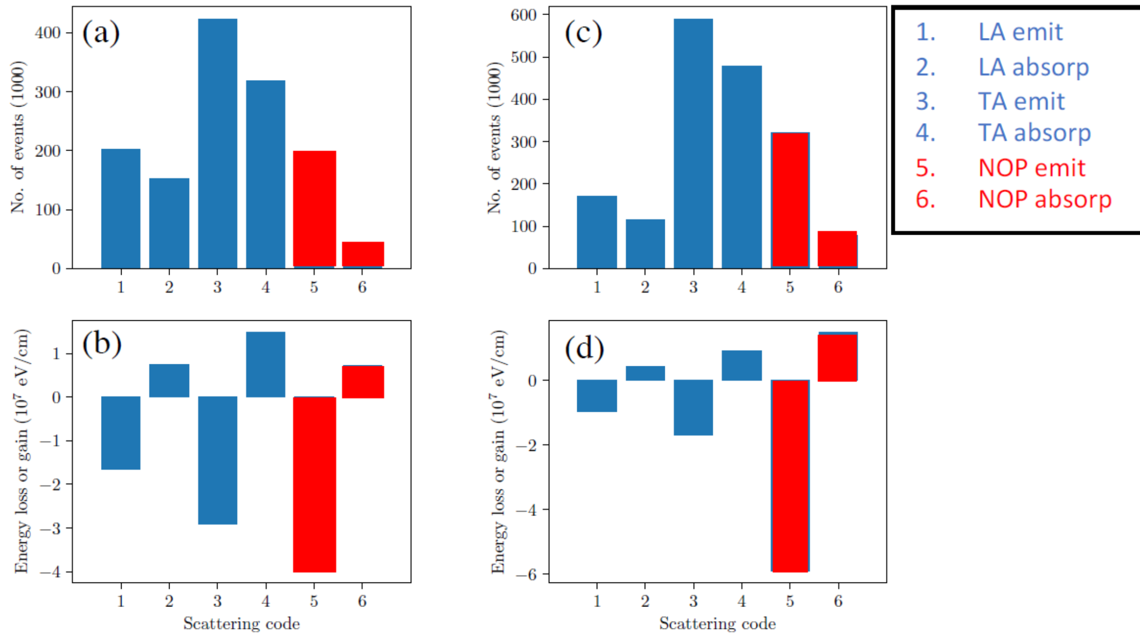

Fig. 5 The number of events (a) and the energy loss or gain (b) for electron scattering with different phonon modes, and the number of events (c) and the energy loss or gain (d) for hole scattering with different phonon modes. The phonon emissions result in carrier energy loss and phonon absorptions result in carrier energy gain. Some of the phonon modes considered in the simulations are indicated in the inset of the figure, with LA indicating the longitudinal acoustic phonon, TA indicates the transverse acoustic phonon, and NOP indicates the nonpolar optical phonon mode.

Carrier transport is simulated for ~ 40 ps. Figure 4(a) illustrates the transient photocurrent at the contact that collects electrons for excitation by different photon energies, ranging from 3.6 eV to 7.2 eV, at a given photon creation rate of $5 \times 10^{27}\ \mathrm{cm}^{-3} \cdot \mathrm{s}^{-1}$ (the current is defined as positive if it is directed into the device domain and vice versa). The current starts from zero at equilibrium and increases during the illumination, as more and more electron-hole pairs are generated and collected. Soon after the photon creation stops, the current reaches its peak and afterwards it begins to decay as the device returns to equilibrium. As expected, the device illuminated with increasingly energetic photons takes longer to relax to equilibrium and collects more charge. Shockley-Read-Hall recombination is not significant at the time scales considered here, as the excess-carrier lifetime is approximately 10 μs [101]. We have also ignored Auger recombination, given the relatively small density of free carriers assumed here, as the excess-carrier lifetime is approximately 10 μs [102]. Since the FBMC simulations are computationally time-consuming, it takes 2 CPU days to simulate 40 ps of transport of 10,000 MC particles (5,000 representing electrons and 5,000 representing holes). Therefore, we are limited to simulations no longer than 1 ns in an acceptable computational time-range.

The creation of excess carriers in a solid exposed to ionizing radiation is best characterized by the quantum efficiency (QE), namely the total number of electron-hole pairs generated per incident carrier. Integrating the current over the simulation time provides the total charge collected at the contacts and QE is calculated as the ratio of the collected charge to the total number of incident photons. Table II provides these values for each photon energy. Quantum efficiencies larger than unity can be understood in the following way. When the carrier energy is above 5 eV (relative to the conduction band minimum), pair-production (*i.e.*, impact ionization) dominates over carrier-phonon scattering events. The high-energy photons generate energetic electrons and holes with different energy distributions at different moments of the illumination, resulting in a net energy loss of 1.12 eV to the band gap for

each event. If the generated electrons or holes have energies above the threshold energy E_{th}, they may produce secondary ionization and generate additional electron-hole pairs. Once the hot electrons relax to an energy below E_{th}, they can only relax further through carrier-phonon scattering, dissipating all the kinetic energy to phonons, as discussed in the next section. Kane concluded that for electrons with an energy in the range of 3-5.5 eV, impact ionization dominates due to the related larger energy loss of 1.12 eV for each event. When the electron energies are below 3 eV, phonon processes dominate inelastic processes [98].

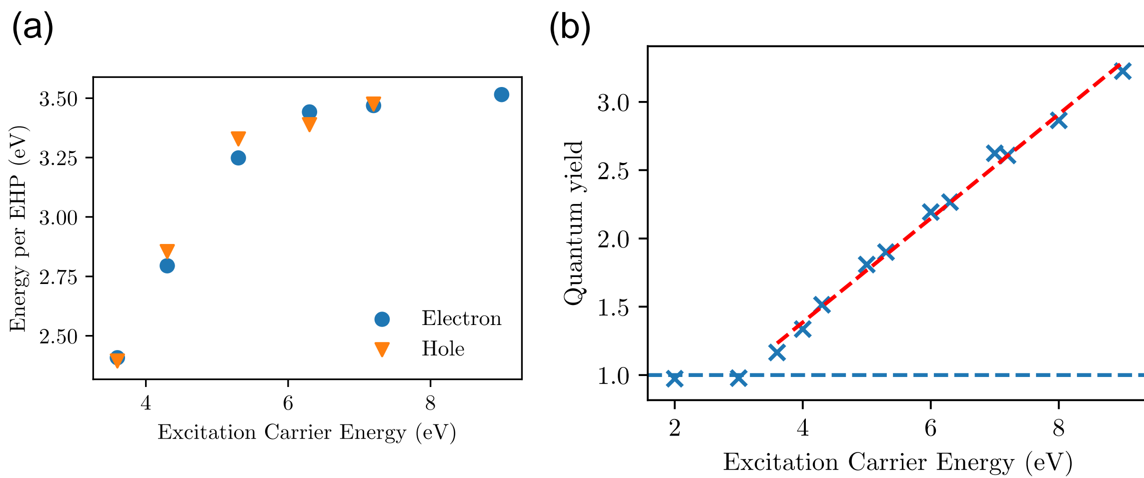

Fig. 6. (a) The dependence of the EHP creation energy on the incident electron and hole energy in silicon. (b) The dependence of the quantum efficiency on the incident carrier energy. The horizontal blue dashed line indicates the unity of the quantum efficiency. The red dashed line indicates the fitted linear relationship between the quantum efficiency and the incident carrier energy. The inverse of the slope corresponds to the EHP creation energy of 2.62 eV for germanium.

Fig. 4(b) plots the quantum efficiency versus photon energy. From the inverse of the slope of this line, we determine an EHP creation energy of 3.69±0.11 eV, which agrees closely with the experimentally determined radiation-ionization energies given in Refs. [82], [87]–[89]. Note that the slope of the fitting line is 0.271± 0.008 /eV. When the energy of the incident photon is at or above the EHP creation energy, the quantum efficiency is close to or larger than unity. This indicates that each photon can generate more than one electron-hole-pair. Mazziotta reported that the energy needed to generate an EHP approaches a constant

value for incident photon energies above the Si K-shell energy of 1.84 keV [103]. Below the K-shell energy, the pair-creation energy depends on the incident photon energy. He reported a maximum value of 3.8 eV for an incident photon energy ~4.5 eV and a minimum value of 3.5 eV for an incident photon energy ~6 eV. This is similar to the results shown in Fig. 2(b). Note that the red dots in Fig. 2(b) are obtained from Table II by dividing the photon energy by the QE, leading to the pair-creation energy at each injected photon energy. The simple parabolic band structure and a uniform density of states in the valence band employed by Mazziotta in [103] is likely responsible for the difference from our results. Furthermore, Mazziotta did not investigate the energy-loss mechanism for the thermalization process, which is studied in detail in the next section.

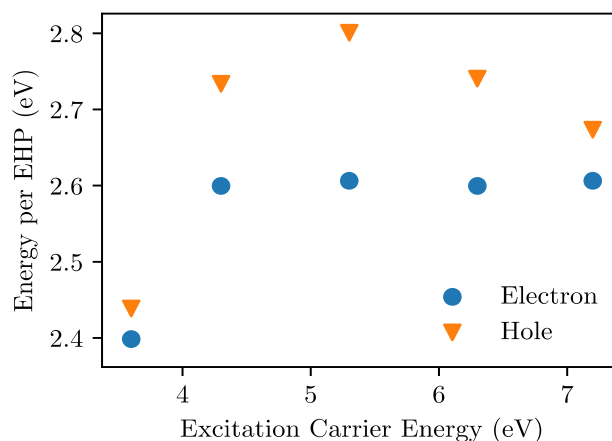

Fig. 7. The dependence of the EHP creation energy on the incident electron and hole energy in germanium.

3.8.2. Dominant Scattering Mechanisms

Figure 5 illustrates the total number of scattering events and the total energy loss or gain for complete thermalization of both electrons [Figs. 5(a,b) and holes (Figs. 5(c,d)] during the entire simulation time. Examining Figs. 5(a) and (c), both electron and hole thermalization indeed involve emission of nonpolar optical phonons during impact ionization, as discussed by Shockley [82] and Klein [88]. The new finding is that, for complete carrier thermalization

down to the band edge, emission and absorption of transverse acoustic phonons are the most common scattering events. For electrons, emission and absorption of longitudinal acoustic phonons are the next most common events. The total energy loss or gain corresponding to each of these scattering processes is shown in Figs. 5 (b) and (d). Electron scattering and hole scattering with nonpolar optical phonon emission lead to the most energy loss (~ 70%), since emitting a single nonpolar optical phonon consumes 0.0612 eV, as assumed based on the silicon phonon dispersion. Emission of acoustic phonons, lumping together both LA and TA phonons with smaller energies, leads to a comparable energy loss (~30%), which is omitted in most previous discussions [82], [88]. It is worth mentioning that different initial carrier energy distribution functions may lead to changes in the relative importance of different energy loss mechanisms. Fischetti *et al.* found that in n-type silicon, the off-equilibrium carriers subject to a steady and uniform electric field lose two-thirds of the energy to acoustic phonons and one-third to NOP phonons [104].

Table III. Quantum Efficiencies for Various Incident Photon Energies in Germanium

$E_{photon}(eV)$	$N_{carriers}(cm^{-1})$	$QE = \frac{N_{carriers}}{N_{photons}}$
2.0	1.703×10^7	0.973
3.0	1.710×10^7	0.977
4.0	2.340×10^7	1.337
5.0	3.167×10^7	1.810
6.0	3.840×10^7	2.194
7.0	4.592×10^7	2.624
8.0	5.015×10^7	2.866

3.8.3. EHP creation in silicon via electrons and holes

We have also performed similar studies when the excitation carriers are energetic electrons and holes. Energy conservation is ensured, whereas the momentum conservation can be

neglected because of the many Umklapp scattering-paths available. Indeed, Kane found that the so-called "random-k" approximation yields satisfactory results and the threshold for momentum-conserving pair creation is very close to that determined by energy conservation alone [99]. Following the same procedure employed in the case of photon illumination, the "nominal" average electron-hole-pair creation energy is found to be approximately 5.2 eV. However, it must be noted that impact ionization involves four carriers, i.e., three new carriers are produced per event [105], whereas photon illumination produces two new carriers. Thus, part of the energy of the initiating carrier is transferred to EHPs, while the remaining fraction is retained by the incident carrier itself. With the assumption of energy equipartition, the recoil carrier also consumes 1/3 of the incident carrier-energy and another 2/3 is transferred to the EHPs. The net pair-creation energy is thus 2/3 of the "nominal" value, i.e., 3.5 eV. The dependence of the EHP creation energy on the incident electron or hole energy is shown in Fig. 6(a). For energies above ~6 eV, the EHP creation energy is ~3.5 eV, which agrees well with the estimate from simulated photon excitation [103] and the experimentally determined values.

3.8.4. Pair creation in germanium

Similar Monte Carlo simulations have been performed for germanium diodes. The model parameters describing the impact ionization rates are listed in Table I. The simulated incident photon energies, the corresponding collected charges, and the quantum efficiencies, are listed in Table III. Figure 6(b) shows the dependence of the QE on the incident photon energies. From the inverse of the slope of the line in Fig. 6(b), we determine an EHP creation energy of 2.62$\pm$0.06 eV, similar to the results of Antman *et al.* [91] and Klein [92]. Note that the slope of the fitting line is 0.381$\pm$ 0.010 /eV.

We have also performed simulations when the incident carriers are electrons and holes. Figure 7 shows the dependence of the EHP creation energy on the incident electron and hole energy. For initiating electrons with energies above 4 eV, an EHP creation energy around

2.6 eV is also observed. For incident holes with energies above 4 eV, the EHP creation energy ranges from 2.6 eV to 2.8 eV, again agreeing well with simulated photon excitation and the experimental value. Klein has concluded that the mean energy dissipated in the creation of an electron-hole pair must be essentially the same for all types of radiation [88], [92].

The device temperature and doping concentration can affect the average electron-hole-pair creation energy. The electronic band structure and the phonon occupation numbers depend on the lattice temperature, which impacts the carrier-phonon scattering rates. Carriers further scatter with dopants. Inclusion of these effects is beyond the scope of this work.

3.9 Summary

The radiation-ionization energies of ~ 3.69 eV in silicon devices and ~ 2.62 eV in germanium devices are examined through full-band Monte Carlo simulations based on the detailed electronic band structures and carrier scattering models. This work extends previous understanding by showing that both acoustic phonon emission and optical phonon emission contribute significantly to energy loss. These are in addition to the energy loss associated with the band gap energy via impact ionization.

A primary finding of this work is that carriers in the lowest energy bands substantially contribute to the radiation-ionization processes, as indicated by the agreement between the calculated and experimental pair-creation energies. Charge carriers resulting from interactions of high-energy particles with the target material continuously lose energy through scattering processes, but the average energy required to create a thermalized electron-hole pair is determined only by interactions that occur after the carriers have energies of ~5 eV or less. In the next chapter, we show the influence of additional scattering processes to charge carriers, which become important for devices of confined volume.

Chapter 4

Hot Electron Effects in GAA Transistors

4.1 Preface

The work described in this section was performed through extensive collective efforts from several researchers at i) EECE, Interdisciplinary Materials Science, and Physics Department of Vanderbilt University, ii) Materials Science Department of the University of Texas at Dallas, iii) ECE Department of the Libscomb University, and iv) imec, Belgium. The findings of this work were published in ref. [68].

My individual contribution was to understand the workflow of the simulator, select relevant scattering processes for quasi-ballistic transport, develop the scripting tools for the 3D simulator in several interconnected thrusts, analyze the key conclusions, and prepare the initial draft for publication. In this project, I was fortunate to collaborate extensively with the experts to bring together the knowledge related to physics, semiconductors, fabrication processes, and quantum phenomena to utilize the full potential of the semiclassical transport model.

4.2 Key Aspects of Three-Dimensional Transistors

Without disruptive technology changes with esoteric 2D materials, the design of semiconductor devices at 7 nm node and beyond will continue in the non-planar direction [106] to reduce the short-channel effects (SCE). In that regard, silicon MOSFETs of Fin, nanosheet (NS), and nanowire (NW) architecture are preferred since they would not require disruptive changes to the existing fabrication processes worth billions of dollars. The design of these transistors focuses on delivering significant drive-current on a smaller process-node while keeping the leakage minimum. But these are competing requirements since transistors with smaller channel-area offer better electrostatic-behavior but increasing the resistance (~1/A). The current-drive for the devices with Fin and NS geometry is expected to be better

due to their larger-channel area over the transistors with circular-channel geometry [18]. Nevertheless, gate-all-around devices are likely to replace tri-gate and Fin variants in this decade, considering the performance requirements guideline by the *international technology roadmap for semiconductors* (ITRS). The drain-induced barrier-lowering (DIBL) and subthreshold-slope of the next-generation Fin devices are projected to be twice that of NWs, presented in Fig. 8.

Futuristic GAA CMOS technology, both NS and NW FETs, have competing advantages over each other. While NS provides higher on-current, NWs offer ultimate scaling in all directions and suffer less from short-channel effects (SCE) [18]. The reduced parasitic capacitance in the advanced forksheet NS structure, as proposed by imec [1], promises adequate subthreshold control like circular NWs while providing higher current-drive. Ultimately, as the dimensions keep shrinking more, Moore's law for silicon may stop with the smallest possible footprints of GAA devices since sufficient on-current within a limited die-area can also be achieved by vertically reorienting the devices similar to refs. [4], [5], [107] and/or solving process integration issues in three-dimensions. Indeed, once the process technology for GAA devices matures for extended stacking in non-planar directions, a higher density of nanowires per device would help in reducing the issues with the variability.

Regardless of the geometry, confined electrons in the highly scaled devices would no longer behave identically to what we know for larger devices [39]. Considering the GAA devices among the promising solutions to technology scaling [4], [5], [107], this study reproduces and analyzes the electron transport, reliability issues, and mechanism of degradation for imec fabricated circular NW transistors through high-fidelity simulations. Most importantly, this study answers why hot carriers' interaction with passivating hydrogens in the interface as the underlying basis for multi-carrier vibrational effects (MVE) needs to be revisited.

4.3 Addressing Open Questions for GAA Devices

Considering the devices of recent generations, the vertically-stacked (VS) GAA transistors of 8 nm diameters display better drive current, effective mobility, and resilience to temperature variation for over FinFETs of 5 nm width, fabricated on identical process flow [108]. For GAA-NW and Fin transistors of identical dimensions, an experimental study by Chasin *et al.* [6], [7], [9] ($L_G = 24$ nm, channel diameter/ width of 8 nm, effective oxide thickness (EOT) = 0.85 nm) reported similar bias-temperature instability of the GAA devices to FinFETs but threshold-voltage shifts, transconductance degradation, self-heating, and high-power dissipation remain major reliability concerns for GAA technology. These effects may reduce current drive, maximum operational frequency, and device lifetime. In addition, since the oxide surrounds the channel, power dissipation problems may be worse for GAA MOSFETs than for FinFETs. Again, these experimental comparisons do suffer from the intrinsic variability in the process steps [4]–[7], [9], [107], such as gate-thickness variation, interface properties, RDF, etc. The effects arising from random variabilities are different for GAA and Fin geometry. While GAA devices are better resistant to gate-length variation, FinFETs are reported far more immune to the variability in channel-width for 5nm node and beyond [109]. Before the intrinsic process variations are diminished to negligible scale or sensitivities to it are better understood, simulation of idealistic devices in conjunction with experimental results is the way forward to gain better insights into the smallest transistors' performance and reliability.

This chapter investigates the degradation of the VS-SGAA FETs in refs. [4], [5], [107] considering it a probable candidate for future generations and scalable down to at least 2 nm nodes [109]. Experimental degradation data of these devices are compared against the influence of phonon processes, Coulomb interactions, plasma oscillations, potential fluctuations, and bias dependence. In addition, hot electron kinetics versus operating conditions is discussed in light of the experimentally-derived reliability predictions from

earlier studies. Effects of momentum relaxing interaction to mobility reduction are presented in the next chapter. It is noteworthy to mention that esoteric forksheet transistors may provide better drive-current and may scale down to 2 nm node but was not considered in this work since the information required to simulate these devices is not available yet.

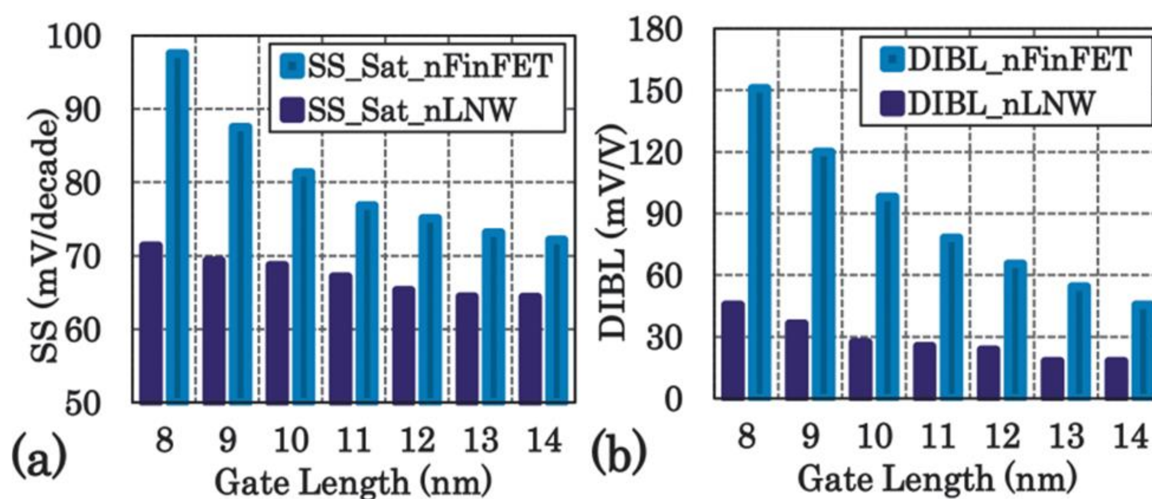

Fig. 8. Short-channel characteristic of FinFET versus horizontally-integrated GAA-NW (after [17]). (a) Better subthreshold characteristic for linearly stacked GAA devices (b) Unacceptable drain-induced barrier lowering for scaled FinFETs.

4.4 Degradation due to Short-Channel Effects in GAA Devices

The device degradation, caused by multiple mechanisms being simultaneously active, is often measured and modeled using experimentally derived methods. The reduction of electrostatic control or channel-current are among the major reliability consideration for these devices, often arising from drain-induced barrier lowering (DIBL), velocity saturation, quantum confinement, and hot carrier degradation. However, dissociating the reliability measurements to the physical mechanisms remains an ongoing endeavor, although researchers have spent several decades working on this problem [9], [35], [46], [62], [97], [110]–[114]. Such a scenario exists because the device lifetime and performance have become increasingly sensitive to the changes in device design along with each generation of

process-technology, especially for smaller nano-devices since the sensitive regions (e.g., channel, drain, oxide) are only tens of atomic layer thick [9]. As such, device characteristics and path-to-failure should be revisited more frequently than ever for all promising next-generation technologies.

Short-channel effects (SCE) and device degradation have historically been attributed to high-field regions within the devices [115]. The electric field in the sensitive region of the device has risen continuously because of the aggressive scaling of the channel and access (S/D) region. In comparison, the scaling of the biases was sub-linear and has come to a stop for a decade. For earlier-generation devices operating at higher biases, all SCE were explained by the elevated-field concept, but a significant reduction in performance has also been observed for advanced devices operated at relatively low fields/ biases [114], [116]. Lateral electric fields remained stalled beyond 45 nm node because neither gate-length nor applied biases are scaled for advanced nodes of Si technology since 2008 [117]. However, device geometry, number of active channels, and channel width continue to exist among major design parameters. As a result of the continuous shrinking of channel width, the carrier movements (and injection) in real space are influenced by an enhanced interplay between lateral and vertical electric fields [114]–[116]. In other words, electron transport for modern devices is more strongly coupled to the biases on drain and gate contacts than previously.

In addition to Coulomb effects, scatterings in silicon transistors are a function of bias conditions, local electric field, doping, electron energy, interface properties, surface roughness, quality of the oxide material, random dopant fluctuation, parasitic effects, device geometry, and dimensions. This underlying dependence affects both the device performance and reliability. Fabrication processes, which influence a few of these factors, are continually evolving, whereas the bias condition, doping, and electric field are barely changing such that they can be considered static for the last decade. However, electric field has a strong geometry dependence $\left(-\frac{\partial V}{\partial \mathbf{r}}\right)$ and can accumulate or depreciate in view of the curvature of

the semiconductor-oxide boundaries. In that regard, the cylindrical device disperses potential better than all others, thanks to its curved surfaces throughout. The subsequent change in the field dispersion would be noticeable near the drain-junction region, where the electric field is highest within the semiconductor region. Since this drain-junction region is considered the most vulnerable (since Shockley and investigated heavily since then); in this chapter, energy distribution functions for the GAA devices in the drain-junction region are thoroughly investigated and compared to the experimental data.

The changes to the carriers' momentum in real space, on account of the coupled electric field, often determine the number of scattering carriers would experience in a quasi-ballistic device. The more elongated the effective path becomes, the chances get higher that an electron drifting in Si would encounter inelastic energy loss or gain due to collisions or phonon interactions. Both energy loss and gain from these scattering events are undesirable. Energy gain may extend the tail of the hot-electron distribution, triggering reliability concerns, whereas energy loss events contribute to heat and reduce the current drive. Electric fields in the channel and thin oxide of the logic-level Si MOS devices are not changing much anymore, but the sharpness of the corners, edges, and interfaces may skew the potential and change the electron velocities by a considerable margin. In this regard, GAA NWs with continuous curvature provides a continuously varying field without abrupt crests and valleys. Thanks to extensive research on the last few decades, it is now widely accepted that SCE is primarily a carrier-energy and carrier-density-driven phenomenon [34]–[36]. The following paragraph discusses the appropriateness of the ensemble MC method to study the induced degradation in short-channel devices.

The mechanisms for defect generation in CMOS devices due to hot carriers are broadly classified into (i) impact ionization [118], (ii) anode hole injection [119], and (iii) hydrogen release [120] events. Among these activities, only the probability of "hydrogen release" processes has remained relevant to logic-level Si devices of today's technology [16], [51],

[121]. The basis of popular multi-vibrational models s to explain the creation of interface traps via hydrogen release was primarily developed and validated for operating voltages higher than Si bandgap ~ 1.2 V [122]. Since logic-level transistors are no longer operated at such biases, it has become essential to review the ideas behind the MVE mechanisms for devices of this generation -- preferably by both experiments and physics-based simulations. Since direct measurements of hot-electrons through optical techniques have become difficult due to higher design complexity, oxide materials, metallization layers, and interconnects, comparing the simulated hot-electron distribution to device degradation data has become essential. This comparison would aid in determining whether hot electrons are available in sufficient energy and quantity for initiating the direct release of hydrogen by MVE processes.

Hot electrons, featuring an extended energy-tail, are known to peak in the sensitive drain-region where they primarily peak at energies equal to the sum of applied drain bias and the Fermi energy of the source-region [40]. The shape and magnitude of the distribution of the hot carrier in energy and physical space depend on additional factors including electrostatics, impact ionization events, plasma interactions, electron and phonon band structure, Coulomb interactions, short- and long-range carrier-carrier interactions, quantization of the inversion layer, surface roughness, source starvation, DIBL etc. [15], [39], [123], [124]. Subsequent sections of this study 1) discuss the atomistic processes in detail that may influence the electron energy distribution (EED), 2) compare the simulated EED to experimentally-derived device degradation data, 3) review mechanisms for device degradation, and 4) propose for the incorporation of quantum confinement in the device simulation as a next step.

4.5 Simulation Methods and Scattering

Lucky electron [17], [26], drift-diffusion [27]–[30], hydrodynamic [42], [43], and more recently, virtual source [15], [31], [32], [125] models have been used extensively to simulate electronic transport in Si devices. Often these approaches use simplified energy and

momentum relaxation times or averaged quantities, such as injection velocity, mean free path, and apparent (ballistic) mobility [31], which are useful but limited in analyzing high-energy transport effects [126], [127] because electrons in semiconductors travel below their ballistic limit regardless of the channel-length [128]. Using these simulation techniques, early predictions of device performance were optimistic for highly-scaled devices [123], [129]. In contrast, only a sublinear improvement of the current and speed is achieved in practice. In extreme cases, electron velocities have declined with additional downscaling of transistor dimensions [130]–[132]. Overly optimistic predictions of earlier simulation techniques occurred because they do not account for short/long-range interactions [39], [46], [77], [78].

Accounting for the elastic and inelastic interactions of the carriers in the channel, junction, and access region is the key to simulate a device of short-channel and high S/D doping. Several scattering processes, as introduced earlier, are considered within the semiclassical approach of this work for modeling a GAA device with ~ 20 nm channel-length and 8 nm diameter. Relevant scattering mechanisms and dispersion relationships are discussed below in view of their influence on the electrons in real and reciprocal space. Rationales are provided for excluding/ including scattering modes in the simulations of the devices of this study.

Energy-momentum dependence of electrons in Si above tenths of eV are complex and cannot be represented by analytical means [45], [133]. Agreement between semiclassical simulations and experiments in Si has been achieved previously at 77K and 300K with full electronic energy band without fitting parameters [86]. Without the complete dispersion data of the materials, scattering in momentum space at moderate energies (> 0.2 eV) would constitute significant errors, as understood from the shape of the conduction bands of silicon. The deformation potential or matrix element of carrier interactions in Fermi's golden rule modifies the electron configuration in k-space. After collisions, the allowed final states can

be selected from a parabolic dispersion to some degrees of analytic corrections using non-parabolicity parameters. In doing so, paramount errors for the energy of hot electrons would be accumulated through every cycle of iterations. Instead, a numerical and tabulated dispersion for electrons above 50 meV is used for the simulations presented in this chapter. The numerical energy bands are constructed from empirical pseudopotentials, fitted to room temperature data [127].

Long-range interactions enable electrons in the S/D regions to directly exchange energy with electrons in the channel through plasma oscillations. In sub-30 nm devices, plasma processes are more important than in larger devices [61] because of the proximity of the channel to the entire S/D regions. Due to this proximity, a significant fraction of the channel electrons exchange momentum with electrons within the S/D regions [39]. This substantial momentum transfer and associated enhancement of the hot-electron distribution can be reduced with lower doping levels within S/D regions [71]. But reducing S/D doping to bypass plasma effects is often not possible due to source starvation. Source starvation is a condition that occurs when there are not adequate numbers of carriers with a velocity component in the longitudinal direction; this is a known issue that restricts device scaling [38], [39]. Description of the treatments of plasma interactions from elevated doping and gate-length scaling is described in ref. [134].

Plasma interactions remain relevant even with metal gates and high-K materials used in modern technology. These increase the energy of plasma waves and weaken long-range scattering [46]. Moreover, short-range interactions modify the energy distribution and indirectly affect energy and momentum relaxation times [40], [61]. Thus, considering only long-range processes is not sufficient to simulate the energy distribution function.

2D MC simulations by Fischetti *et al.* [46] quantified the importance of treating short- and long-range carrier interactions and their non-trivial contributions to the resulting energy distributions and reliability. The combined effects of plasma and source-starvation are

articulated by Fischetti *et al*. [39], where static approximations (*e.g*., injection velocity) in virtual source models are shown to be inadequate. The virtual source region's potential adjacent to the source-channel junction is influenced by scattering events that occur throughout the channel. The resulting local potential fluctuations and changes in velocity are better captured using a particle-based simulation approach [39]. Uechi *et* al. [77], [78] treated both short- and long-range Coulomb interactions in 3D and showed that potential fluctuations in a highly-doped region ($> 1 \times 10^{20}$ cm^{-3}) may be as large as 0.20 eV and can be treated accurately using a semiclassical MC approach. The hot-electron influx from the channel towards the drain must be dynamically screened, considering the high-density of the free electron gas in the access region, as demonstrated earlier in refs. [45], [46].

Ionized impurity scattering is less significant in the unintentionally doped channel region, whereas the electrons in the access S/D region frequently scatter with the background charge from the ionized dopants. The resultant EED in the junction and the access regions can become a strong function of impurity scattering and additional heating from carrier degeneracy [71]. Despite being elastic in nature and rarer in the channel than the access region, ionized impurity scattering changes the momentum of the electrons in the channel and alters the direction of their drift velocity. This lengthens the electron path from source to drain and facilitates inelastic scattering with phonons—consequently, the current drops below the sub-ballistic limit.

In this work, we use a high-fidelity code similar to Damocles, developed by Fischetti and Laux [40], [47], [83], [113], [135], that solves Poisson's equation in 3D to appropriately characterize momentum relaxation in all dimensions. The 2D code has been validated extensively with experimental data [47], [71], [136]. The device is simulated using a finite-difference tensor mesh. Finer grids are used in k-space to ensure energy conservation in meV range during each stochastic scattering. Electrons in the Si NW are modeled as MC particles with effective electron charge weight smaller than one; holes are included in a constant

quasi-Fermi level (zero current) approximation. The detailed steps of our semiclassical approach are available in refs. [64], [70]. Simulations include scattering with acoustic phonons, optical phonons, impact ionization, and short/long-range interactions (plasma, potential fluctuations, carrier-carrier) with full electronic energy bands [127].

Electron-electron scattering is treated using the method described in appendix A of [45], appendix C1 of [46], and [14]. Short-range interactions of hot carriers in the channel with cold carriers of the drain are dynamically screened. Long-range interactions (plasma oscillations) are semi-classically treated using a self-consistent solution of Poisson's equation. Mesh spacing is set at 2.8 Å, as found to optimally represent the dependence of Landau damping on electron concentration [71]. Trade-offs are necessary due to conflicting requirements of mesh spacing and charge weight [137]. To reduce the complexity of the 3D inhomogeneous device simulations, plasma oscillations are treated as particles carrying 0.1 electron charge to achieve the necessary computational statistics [46], [137].

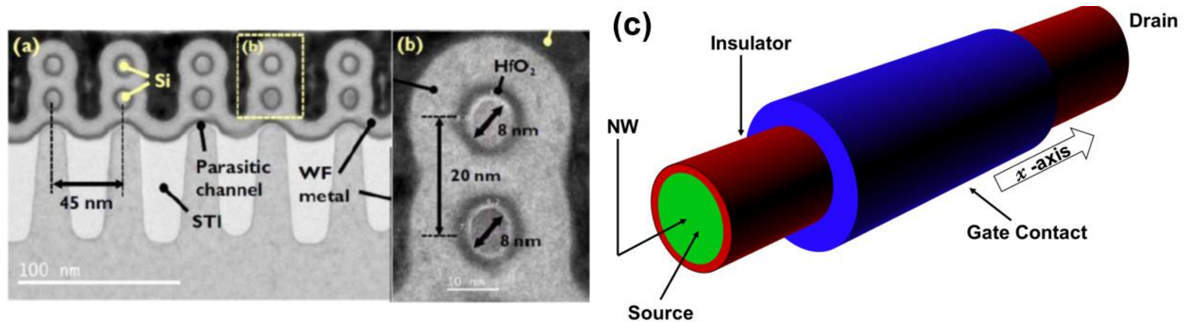

Fig. 9. TEM images of a GAA Si nMOSFET from imec [4] depicting a (a) NW array and (b) single NW vertical stack. WF metal represents the work-function metal used for the gate electrode. (c) The simplified device used for simulation is shown schematically (not drawn to scale). The insulating oxide in the S/D access region and gate contact extends radially outward until the end of the tensor mesh (not shown).

A Poisson time-step of 0.2 fs is used for all computations. The Gilat-Raubenheimer algorithm [96] is used to calculate densities of states for the conduction and valence bands. Carrier-phonon scattering rates are computed using the deformation potentials of [71]. The

impact-ionization model [13], [98] has been obtained by fitting a Keldysh-like energy-dependence to the ionization rate calculated using the empirical-pseudopotential energy band structure and calibrated to X-ray Photoelectron Spectroscopy and carrier separation experimental data [97]. Carrier-phonon scattering and impact ionization models have been validated in Si with experiments up to electron energy of 7.2 eV [83]. Ionized-impurity scattering is treated via the Conwell-Weisskopf model [138]. The oxide is treated as an ideal insulator, with no carriers present. Electron collisions at the oxide interface are specular and conserve momentum. The relative dielectric constant is adjusted to produce an effective oxide thickness of 0.85 nm, consistent with [4].

Effects of quantum confinement were found to be negligible in previous work on devices of similar or smaller dimensions, and were not included in the simulations [52], [55]. For example, calculations of the band structure of circular-cross-section Si nanowires using empirical pseudopotentials show that the lowest-energy sub-bands are spaced by about 0.2 eV for a diameter of 2.5 nm [52]. Rescaling by the square of the ratios of the respective diameters, this value drops to ~20 meV for the nanowires of this work. Moreover, MC simulations of rectangular cross-section nanowires by Donetti *et al.* [55] solved Poisson's equation in 3D, the Schrödinger equation in 2D, and treated transport as ballistic in 1D. For nanowire diameters larger than 4 nm, they find that, except for a V_{Th}-shift, the current is not affected by quantum effects in Si NWs. Hence, our simulations did not include quantum confinement. Source to drain tunneling is also not considered since it is not important for devices with gate lengths greater than ~10 nm [56]–[58]. For VS-GAA devices of upcoming generations, quantum confinement effects would become important at the reduced diameters of the nanowire. For the two-dimensional device simulation of such advanced devices, the effects of quantization of the lowest energy subbands in the inversion layer is presented in ***chapter v***. The semiconductor-oxide interface is conventionally chosen along a <100> plane and the direction of electron transport occurs parallel to the (100) directions of the device.

Next, we report 3D full-band Monte Carlo (MC) simulations for a GAA Si nanowire (NW) MOSFET, solving the Boltzmann transport equation by incorporating various scattering processes. Scattering processes such as intervalley, carrier-phonon, impact ionization, ionized-impurity, and carrier-carrier scattering [10], [45]–[47], included in the present MC approach, are critical for a complete description of the high-energy tails [44], [135], [136], [139]. Excellent agreement between the simulated and experimental current-voltage (*I*-*V*) characteristics is obtained.

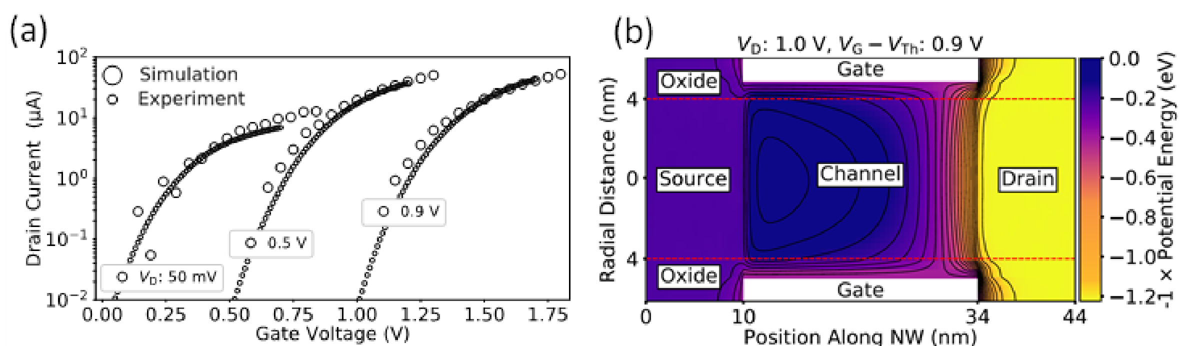

Fig. 10. (a) Experimental and simulated gate control at 50 mV, 0.5 V, and 0.9 V drain bias. Curves are shifted along the x-axis by -0.5, 0.0, and 0.5 V, respectively. The experimental drain current is divided by 2 to match the simulated *I*-*V* characteristics of the single NW structure depicted earlier. (b) Electron potential energy along the x-direction of the cylinder, illustrated with a color map. The density of the equipotential lines reflects the magnitude of the electric field. Metallurgical junction boundaries are at x = 10 nm and 34 nm. The dashed-red lines separate the oxide and semiconductor region of the device. The aspect ratio is different than actual device dimensions for clarity.

4.6 Carrier Transport in a Si Nanowire MOSFET

Figure 9(a, b) shows n-channel GAA MOSFETs fabricated at imec [4] ($L_{\mathrm{G}} =$ 24 nm, channel diameter 8 nm, EOT = 0.85 nm, S/D dopant density = 5×10^{20} cm^{-3}, and channel acceptor density = 1×10^{18} cm^{-3}). A single NW similar to Fig. 9(c) is considered in the simulations [4], but the results are also relevant to vertically stacked NWs in the absence of significant contact resistance and other parasitic effects, which are expected to be small in optimized processes of practical interest. Instead of crudely representing a NW with a square cross-section [55], we closely approximated the circular shape via a tensor mesh

with mesh-node density (defined by the constraints on grid-spacing to accurately account for long-range interaction) of 12.5 nm^{-2}. This approximation is acceptable since the resultant periodic fluctuation (0-0.14 nm) of the radius of the simulated NW is smaller than that of the physical device due to process-variations [9]. Mesh points outside the semiconductor and gate-oxide regions in the cuboidal volume are filled with spacer oxide materials of extreme dielectric constants, rendering their EOT to negligible value.

Fig. 10(a) compares I-V curves for experimental and simulation results. To bypass the noises in simulation, electron currents are temporally averaged during the steady-state simulation after the fall of the initial transient. The simulated characteristics agree well with experimental measurements over five orders of magnitude variation in current. The work function of the metal gate is adjusted in simulations to match the gate control of the physical device.

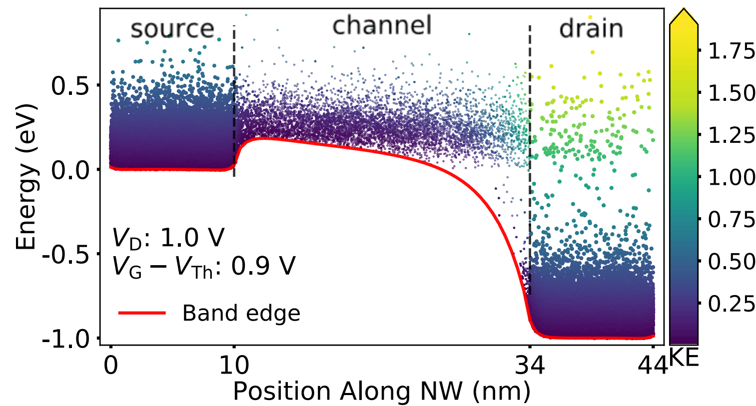

Fig. 11. Carrier kinetic energy (KE) distribution above the conduction-band edge along the x-direction of the device. The conduction-band edge is averaged over the cross section of the channel. Statistical carrier enhancement within the dashed vertical lines ensures adequate sampling of interactions in an otherwise sparsely populated region. The kinetic energy of the carriers is represented by both particle color and the distance above the conduction band.

For a device operating in deep saturation (V_{DS} = 1 V), the densest equipotential lines overlying the potential plot in Fig. 10(b) correspond to a localized field peak near the metallurgical drain-junction at x = 34 nm. Conversely, the electron energies that are plotted

vs. position along the channel in Fig. 11 display a dispersed hot carrier distribution similar to earlier MC simulations of a 30 nm long double-gate Si MOSFET [40]. In this work, we will refer to carriers with energies higher than 0.5 eV as "hot" [9], [33], [35]. Due to the limited volume of the drain access region (34 nm $< x <$ 44 nm), approximately 10% of the hot carriers in the distribution in Fig. 11 do not lose energy in the drain. Instead, they leave – still hot – the simulated region via the metallurgical contact to the drain. This is understandable because the drain dimension (~ 10 nm) is smaller than the mean free path for collisions of electrons in silicon.

4.7 Spatial and Energy Distribution of Hot Electrons

In designing reliable devices, it is critical to understand the sensitivity of the carrier energy to the drain and gate bias [34]–[36], [140]. Figure 12 shows the kinetic energy distribution of the electrons as a function of drain bias V_{DS} for a gate-voltage overdrive $V_{\mathrm{Ov}} = V_{GS} - V_{\mathrm{Th}}$ = 1.3 V. The solid and dotted curves correspond to simulation results obtained with and without electron-electron scattering (EES), respectively.

Calculated electron-energy distributions are obtained by considering carriers in a volume around the metallurgical drain junction, 32 nm $< x <$ 36 nm. The KE distribution in Fig. 12(a) displays a cold carrier peak at approximately 0.15 eV, followed by a dispersed hot-carrier population due to field-driven acceleration of electrons as they transport from source to drain. As shown in Fig. 12, the cold-carrier peak features a higher electron effective temperature due to degeneracy (5×10^{20} cm^{-3}) heating and long-range plasma interactions [39], [46], [141]. In addition to other Coulomb interactions [61], the tail of the cold-carrier distribution is enhanced due to EES, as evidenced by the closely differences among the solid and dotted lines.

In addition to the expected heating of the carriers with increasing V_{DS}, the hot-carrier population in Fig. 12(a) shows a broad secondary peak at an energy given roughly by $qV_{DS} + E_{\{Fs\}}$, where $E_{\{Fs\}}$ is the Fermi level of the source contact, approximately at 0.15 eV. Simple models of hot-electron degradation typically assume that no electron exceeds the Si bandgap energy (1.12 eV) when the supply voltage (V_{DD}) is lower than 1.12 V [9]. In contrast, at V_{DS} = 1.0 V, the hot electron distribution has a broad peak above 1.1 V, and a tail that extends past 1.5 eV. Higher V_D values lead to similarly shaped distributions that extend to even higher energies and exhibit longer tails.

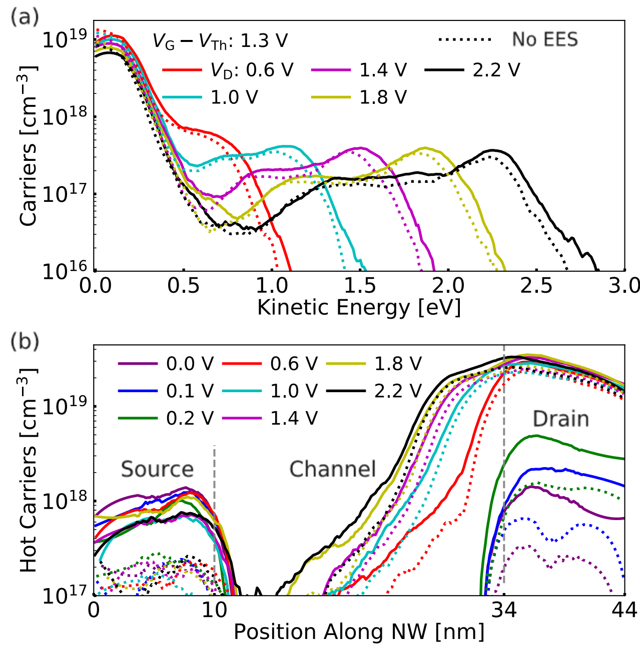

Fig. 12. Energy distribution of electrons for several drain biases at $V_{GS} - V_{Th}$ = 1.3 V. (a) Electron-energy histogram within 2 nm in each direction of the metallurgical drain-junction. (b) Density of electrons above 0.5 eV in the NW. The dotted curves show results without including EES.

The smearing of the high energy peak indicates that electron transport is not ballistic and electrons are strongly scattered by phonons, short- and long-range Coulomb processes, etc. The shape of the KE distribution is similar to previous studies that include these processes

[46], [61], [71]. EES contributes significantly to the hot-carrier population and extend the tail of the distribution [142].

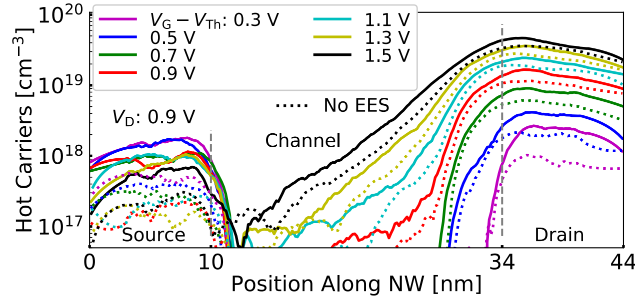

Fig. 13. Density of electrons above 0.5 eV in the NW for several overdrive biases at V_{DS} = 0.9 V. Dashed lines show the hot-electron distribution without EES.

For a constant overdrive voltage, Fig. 12(b) shows that the hot-carrier density at or near the drain junction increases up to deep saturation, V_{DS} = 0.6 V. However, the full-width at half-maximum (FWHM) continues to increase at high drain biases. The large FWHM affirms that a significant number of hot electrons do not thermalize upon reaching the 10 nm drain region. The effects of plasma interactions and potential fluctuations between the channel and S/D are particularly visible at V_{DS} = 0 V since the external field that accelerates the electrons along the longitudinal direction is absent. As shown in Fig. 12(b), a significant density of hot electrons at $V_{DS} \geq$ 0.2 V indicates that carriers energetic enough to cause damage by collective processes (e.g., multi-carrier, multiple activation energies, and trapped states) are generated at nominal bias conditions. Comparisons of the solid and dotted curves show that EES strongly increases the hot-electron population in the S/D. In addition, the hot-carrier population in the channel is broadened by EES, especially at higher biases. Moreover, comparisons of hot-electron densities in the source region at low and high drain biases indicate that source starvation is not a concern at the doping level (5×10^{20} cm^{-3}) of this work.

For the worst-case HCE in ultra-scaled devices, both the maximum energy and the carrier flux impinging on the interface play crucial roles [17], [86], [143]. The number of electrons in the channel of an ultra-scaled transistor depends on both V_{DS} and V_{GS} since the electrostatic coupling between the terminals becomes stronger as the device shrinks. Furthermore, the small S/D volume limits the maximum electron current, since the reservoir cannot keep up with carrier flow in the channel. Figure 13 shows that the magnitude and FWHM of the hot-electron density depend strongly on the gate-overdrive voltages. Strong enhancement of the hot-carrier population throughout the device by EES is again observed for the curves in Fig. 13. A significant hot-electron density is observed at V_{Ov} = 0.3 V, suggesting that a value of V_{DS} = 0.9 V can introduce enough energetic electrons to accelerate the device degradation even when the applied V_{Ov} is small. The damage caused by hot-electron processes becomes more significant with increasing V_{Ov} due to the subsequent increase of the electric field toward the interface. For the biasing conditions considered here, simulations of bipolar carrier transport show that the hot (> 1 eV) hole density in these structures is at least six orders of magnitude smaller than the density of hot electrons. Moreover, the electric field is not sufficiently strong to accelerate holes to high enough kinetic energies to cause significant degradation [9], [135], [144]–[146].

4.8 Comparison with Experimental results

Recently, Chasin *et al.* reported experimentally derived time-to-failure data (TTF) of the GAA device examined here. We show an adapted version of their figure in Fig. 14(a). The dependence of the failure rate on the energy and density of hot electrons is highly complex and highly nonlinear. Hot-carrier-induced degradation is driven by both the maximum energy and the density (flux) of the hot electrons [142], [147]. Within the framework of the theory of the inelastic scattering of hot electrons at defects [110], [148], which can lead to their activation, the pertinent cross-section depends on both the maximum energy that can

be dissipated in a scattering event and the density of hot carriers that can deliver the required amounts of energy.

In Figs. 14(b) and 14(c), we show the density and mean-energy of the hot-electrons, respectively. We focus first on the areas bounded by white lines in Figs. 14(b) and 14(c). In this regime, both the drain and gate voltages are relatively high, and hot carriers are abundant. In similar voltage ranges in Fig. 14(a), device lifetimes are greatly reduced. The overall similarity of our simulation results and data in Fig. 14(a) strongly suggests that HCE play a role in causing device failure in these voltage ranges. When the drain voltage is lower, HCE effects become negligible, and other mechanisms, such as positive bias-temperature instability (PBTI), limit device lifetime [9], [51], [121].

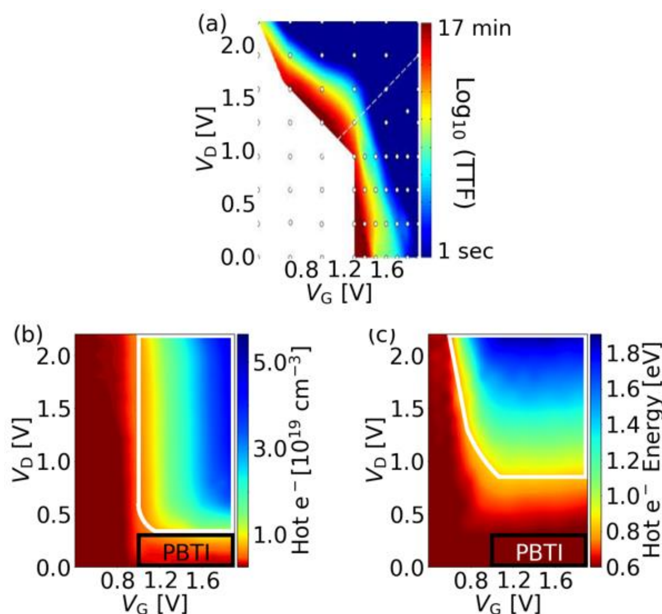

Fig. 14. (a) Experimentally derived time-to-failure (TTF) maps of the GAA device, adapted from Chasin *et al*. [7]. (b) The simulated density of hot carriers in a small volume near metallurgical junction ~ within ±2 nm in the x-direction. (c) The simulated average kinetic energy of these hot electrons. Only the electrons with energy higher than 0.5 eV is considered for subplots (b) and (c). The scales of the color bars in (a) is logarithmic, and those in (b) and (c) are linear.

We now consider how hot electrons near and/or within the highly-doped drain may lead to hot-carrier effects. One mechanism that has been considered extensively in the literature is

multi-vibrational excitation (MVE) [7], [8], [33]–[37]. The energy required to break a Si-H bond is ~2.6 eV [16], [149], [150]. The MVE model presumes that multiple scattering events at a hydrogen-passivated dangling silicon bond (Si-H) at the SiO_2 interface can provide enough energy to liberate the passivating hydrogen atom, leaving behind an interface trap [33]–[36].

The presence of a high density of hot electrons in the highly-doped drain, as observed in the simulation, suggests another potential mechanism for hot-electron induced interface-trap creation. The drains of these devices are highly doped with phosphorus [5], [150]. The injection of hydrogen in devices during processing, which is needed to passivate defects such as interfacial Si dangling bonds, inevitably passivates a significant fraction of dopant atoms as well [5], [16], [151]–[153]. The energy required to break a P-H bond associated with a passivated dopant complex in the drain is ~1.3 eV [154]. Once a hydrogen atom is liberated from a dopant atom, it can diffuse rapidly until it either encounters another H atom and dimerizes or encounters an energetically favorable reaction site. The energy for a diffusing hydrogen atom to react with a Si-H bond, leading to interface-trap creation, is typically of order ~0.3-0.4 eV [16], [153], [154]. Thus, the presence of a high concentration of carriers in the drain with energies above ~1.3 eV may lead to the release of atomic hydrogen, diffusion of a percentage of the liberated atomic hydrogen to the nearby channel, and its subsequent reaction with Si-H bonds to create interface traps [153]–[159]. In nanoscale devices, the presence of even a few interface traps near the drain junction can greatly degrade device performance [9], [10], [36], [45], leading to a significant reduction in lifetime.

Similar field- and/or carrier-induced release of hydrogen from passivated dopant complexes in the Si substrate has been proposed as a rate-limiting step in negative bias-temperature instability (NBTI) in *p*MOS devices [153], [155]. For *p*MOS NBTI, the barrier for hydrogen release from dopant atoms is reduced by the high densities of holes in the substrate [155]. For *n*MOS HCE in these nanoscale devices, the presence of a high density of energetic

electrons within the drain can be the initiating step in the process. The energetics of trap creation are otherwise similar after the hydrogen is released [16], [153]–[159]. In contrast to NBTI, the PBTI process that is expected to dominate device degradation at the high oxide electric fields present in these devices under the conditions of the experiments in Fig. 13(a) does not depend on hydrogen release and/or reaction. Instead, the PBTI in these devices is typically a result of electron tunneling into border traps, a process with very different energetics and kinetics. Whether HCE or PBTI is the rate-limiting process in device operation is determined by biasing levels and fractions of lifetime during which the device is conducting *vs*. the fraction of time during which the device is turned off. For cases in which HCE dominate, calculations similar to those shown in this work can be applied to assist in lifetime estimation. For cases in which PBTI effects dominate, alternative models have been developed and must be applied [9], [51], [160].

Due to the significant reduction in bond-breaking energies, hydrogen release from a P-H complex (~1.3 eV barrier [154]) in the drain can occur much more easily as the result of a single hot-carrier interaction than hydrogen release from a Si-H complex (~2.6 eV barrier [5], [16], [149]). Which of these mechanisms dominates for a particular device depends on a number of factors, including: (1) the channel current density, (2) densities of electrons within the drain with energies greater than ~1.3 eV, (3) relative densities of Si-H complexes along the channel interface and P-H complexes within the drain, (4) the respective interaction cross sections between hot electrons and Si-H complexes at the Si/SiO_2 interface and P-H complexes within the drain, and (5) probabilities that a hydrogen atom (charged or neutral) dimerizes with another hydrogen atom or forms another complex before reaching the channel/SiO_2 interface.

It is beyond the scope of this work to develop a quantitative model of hot-carrier degradation initiated by P-H bond breaking due to hot carrier interactions within the drain. Regarding the development of such a future model, we note that: (1) The channel current density can be

determined experimentally. (2) The calculational methods of this work provide the necessary information about carrier energies. (3) Si-H and P-H densities and spatial distributions depend on device processing and are generally not well known. Hence, at least initially, these would need to be treated as adjustable parameters. (4) Interaction cross sections for channel carriers and Si-H complexes have been calculated in MVE models (MVE) [7], [8], [33]–[37], [161], but interaction cross sections with P-H complexes are not well known at the energies relevant to this work. (5) The diffusion and reactions of hydrogen within Si MOS devices have been studied in great detail for more than 40 years, and diffusion barriers and rates for hydrogen transport and reactions have been calculated for multiple species under a wide range of experimental conditions [16], [34], [36], [111], [151], [152], [160]–[163]. We therefore expect that significant insight into reliability-limiting mechanisms for nanoscale devices could be obtained by evaluating the relative importance of hydrogen release from dopant complexes in hot-carrier-induced degradation in MOS devices.

4.9 Summary

A comprehensive semi-classical model is used to investigate the electronic energy transport properties of a GAA Si NW MOSFET. Simulation results show that plasma effects and electron degeneracy elevate the energy of the electrons in the device in conjunction with other scattering mechanisms. Not all hot electrons lose their energy after crossing the drain depletion region; instead, some move quasi-ballistically in the drain. Hot-carrier reliability of the GAA transistor is investigated by comparing experimental time-to-failure data to the simulated density and average energy of the hot-carriers. We find that the average energy and density of hot electrons correlate well with the measured TTF in ranges of operation for which both drain and gate voltages are applied for significant fractions of device lifetime. It is proposed that the release of hydrogen from passivated phosphorus dopant complexes in the drain and the subsequent reaction of diffusing atomic hydrogen with passivated silicon dangling bonds at the Si/SiO_2 interface may contribute significantly to hot-carrier induced

degradation in these nanoscale devices. When devices are primarily biased statically, other failure mechanisms such as PBTI dominate, and different models must be applied.

Chapter 5

Energy Exchange Processes and Barriers to Ascending Device Performance

5.1 Introduction

Scaling of device dimensions and doping profiles yielded performance improvement for MOS devices until the early 2000s and represented a remarkable achievement of technology. In a quest to control the leakage current and improve the functional yield of the chip, the semiconductor industry has long ago shifted from planar designs to the silicon on insulator, double-gate (DG), and FinFET structures at reduced dimensions. Currently, the gate-all-around (GAA) geometry is under significant consideration for attaining better electrostatic control. However, further dimensional scaling for these devices only provided diminishing returns and the return became vanishingly small for the current sub-30 nm (gate-length) era. In these short-channel devices, the meaning is lost for average-based concepts like apparent mobility and velocity. Electrons in the channel of these devices do not go thorough enough scattering events for the mobility concept to be meaningful. However, these devices continue to perform below the ballistic limit, and the task of performance improvement is becoming exponentially harder in reduced dimensions. It has been recognized that the barrier in performance improvement may be fundamental and related to one of the most exotic phenomena of nature – Coulomb scattering, thanks to earlier MC simulation works addressing this topic.

We found that the Coulomb scattering indeed plays a role in reducing the channel current of the GAA device introduced earlier in ***chapter IV***. This finding is in agreement with the earlier predictions in refs. [14], [61], [69], [98]. Both the predictions and simulation results are profound because (i) they are obtained by a scattering rate data that is widely agreed upon worldwide ii) the employed model was verified in the past across bulk to quasi-ballistic regime with extensive experimental data often not available for materials other than silicon

iii) simulated current is verified to experiment, as presented in an earlier chapter. The question that we are addressing here is indeed not unique to silicon; however, scattering in Si is better understood than any other materials, whereby silicon would be an ideal candidate at his stage. Dynamic variables (e.g., interface roughness, halo doping, oxide defects, etc.) also influence device performance but change fast with every new generation and were not considered here.

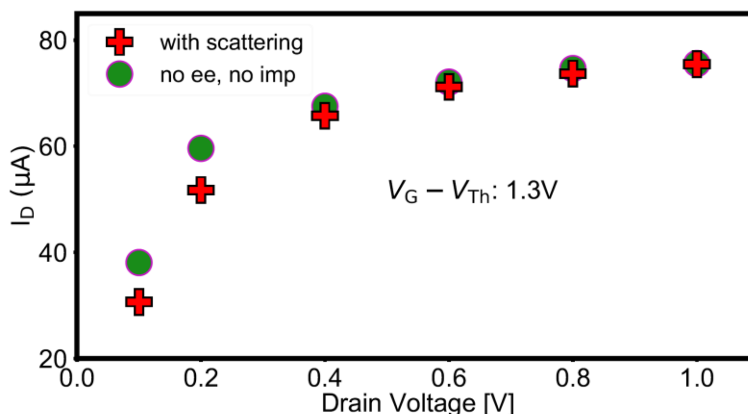

Fig. 15. Dependance of the drive current to the elastic scattering processes in the channel for several drain biases. Points with cross and circles respectively indicates the simulation results with and without accounting for Coulomb processes.

5.2 Scattering Barrier to Ballistic Transport

The dependency of the drive current on these elastic Coulomb processes is illustrated in Fig. 15. The selected overdrive voltage $(1.3V)$, used for the comparison with and without Coulomb scatterings, was intentionally chosen large to induce a large concentration $(\geq 10^{19} cm^{-3})$ of carriers in the channel as desired for the present and next-generation devices. Current is significantly reduced at or below 0.4 V drain-voltage. This substantial reduction in current for the 24 nm channel length, operating around the so-called 'ballistic' transport distance from source to drain, clarifies the barrier to the performance improvement of the GAA devices. In other words, beyond the obvious improvement in subthreshold performance, GAA devices will not be the answer to realize the '*ballistic dream*', even with

the idealistic oxide and interface properties, as considered for this simulation. The results obtained here are in agreement with earlier predictions for Si in refs. [59], [61].

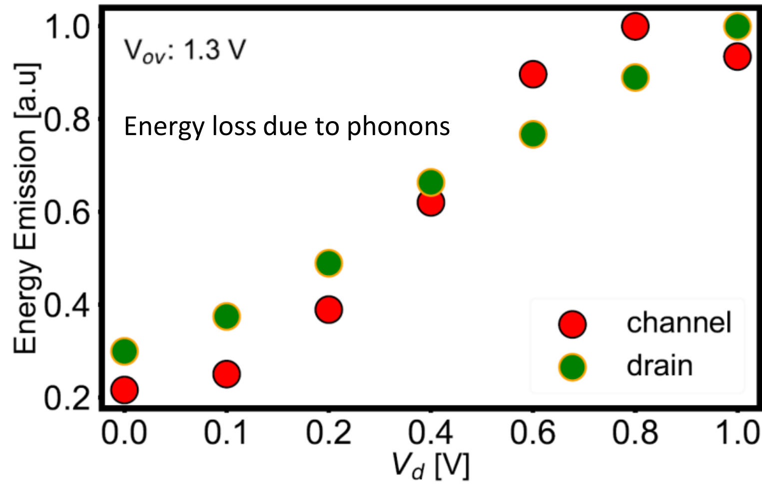

Fig. 16. Trend for phonon emission (loss) by electrons in the channel and drain region with respect to the applied drain-voltage. Phonon emissions are calculated as a sum of transverse acoustic, longitudinal acoustic, and non-polar optical phonon modes. Both curves are individually normalized by their maximum value since the number of particles in the channel and drain region is different in the simulation.

The combined effects of the electron-impurity and electron-electron scattering to drive current are implicit in the scattering rate vs. momentum relaxation dependence, provided in eqs. [2.23-2.25] and ref. [70]. Both electron-impurity and electron-electron are Coulomb processes with matrix elements that are inversely proportional to the transferred momentum. As the drain bias increases, considering the short-channel length $(\sim 24\ nm)$ of the GAA device under consideration, electrons become hotter in a larger fraction of the channel, so these Coulomb processes become less important. So, they are both strong at low energy, but become weak for carriers accelerate to higher energies due to large drain-biases. As plotted in Fig. 16, phonon processes also exist for higher drain-biases but saturate in the channel due to weakening Coulomb effects and the saturated carrier density. Instead, source starvation becomes the limiting factor for transport under elevated fields and causes the currents to saturate.

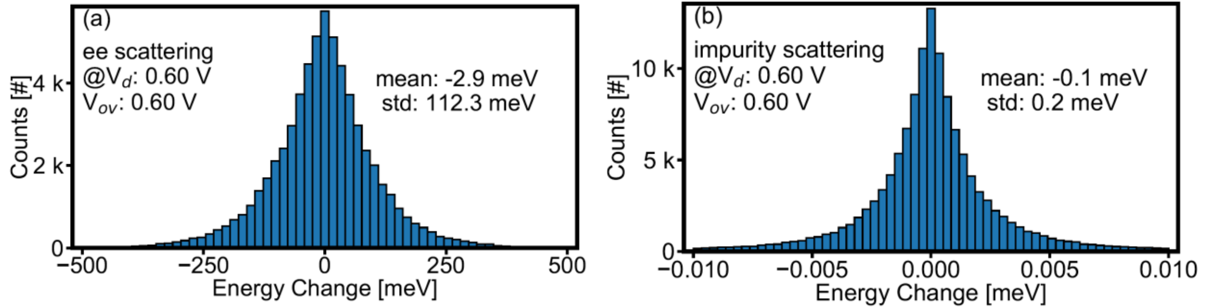

Fig. 17. Histograms showing the errors in selecting the final state in k-space for elastic (a) electro-electron and (b) electron-ionized impurity scattering. The mean and standard deviation of the error while selecting the final state of the electrons are much smaller than the energetics of the simulation.

The numerical handing of the simulation in this chapter is an improvement to those presented in ref. [68] and in ***chapter IV***. In contrast to [68], ten times smaller threshold for energy conservation and full numerical electronic energy bands in the entire spectrum (including the parabolic regime) has been used. The pedantic energy conservation ensured in the simulation is presented in Fig. 17. As can be seen, the average energy change for the elastic events approaches zero, and the standard deviation is an order of magnitude lower than the energies of the carriers of interest — non-equilibrium electrons in the channel and drain. In addition, the erroneous total nonzero energy emission by elastic means shown in Fig. 18 is indeed dwarfed by those of phonon processes, especially considering the high volume of events associated with the Coulomb interactions. This energy conservation issue (minor) is a byproduct of discretizing the k-space, which is unavoidable while including full electron energy bands since no analytical solution is available. Nevertheless, the results presented here are found to be invariant to any further improvement in accuracy.

Phonon scattering for the entire device in Fig. 18 resembles the results obtained during the thermalization of hot-carriers in bulk (large-volume) silicon in ***chapter III***. The new aspect is the inclusion of the elastic processes in the simulation within the context of a short-channel device. Elastic scattering changes the momentum of the electrons and consequently

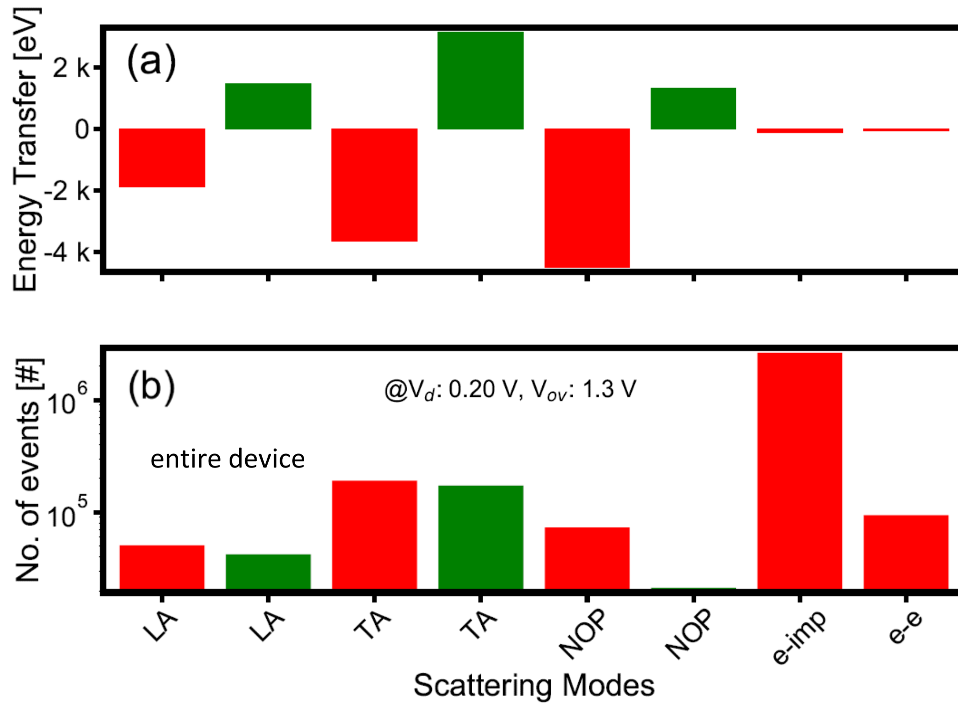

Fig. 18. (a) The number of events (b) and the energy loss or gain for electron scattering with different elastic and phonon modes. The plotted data were temporally accumulated over some period of steady state simulation in the entire device. Some of the elastic and phonon modes considered in the simulations are indicated in the x-axis of the figure, with LA indicating the longitudinal acoustic phonon, TA indicates the transverse acoustic phonon, NOP indicates the nonpolar optical phonon, e-imp indicates the electron-impurity, and e-e indicates the electron-electron scattering mode. The red and green color correspondingly indicates the emission and absorption of energy by the electrons. Note: The capital typefaces in x-axis indicate inelastic events and differentiate to those from the Coulomb processes in small typefaces. The energy loss-gain by the ee and e-imp processes are not real but only provided to illustrate the numerical aspect of energy conservation.

elongates their path in the device, causing additional energy transfer events through phonons. Comparing the frequency of the scattering modes in the entire device shown in Fig. 18 to that of the channel region in Fig. 19 confirms that Coulomb processes play a major role in reducing conductance in the channel under moderately applied drain-field. The comparison reveals that electron-impurity scattering is the most common scattering event in both channel and access regions. A larger fraction of impurity processes takes place in the access region where the doping is extreme. However, as can be seen in these figures, for the most part, e-

e scattering occurs in the channel region. In addition to plasma oscillations and short-range forces, e-e scattering in the channel-region benefits from the abundance of free states to scatter into. Contrary to the commonly employed ballistic transport approximation, our results imply that a noticeable fraction of electrons in ~24 nm channel scatter both elastically and via phonon processes. Both e-e and e-imp are barriers to the ballistic dream, but the ionized impurity scattering can be tackled using an unintentionally doped channel instead of the extensive doping of $10^{18}\ cm^{-3}$ chosen for this work to reveal their role in reducing conductance. However, the situation is complicated for e-e scattering. With continuous dimensional scaling and increased carrier density in the device, ballistic dreams at any length scale and geometry may remain unattainable due to e-e scattering.

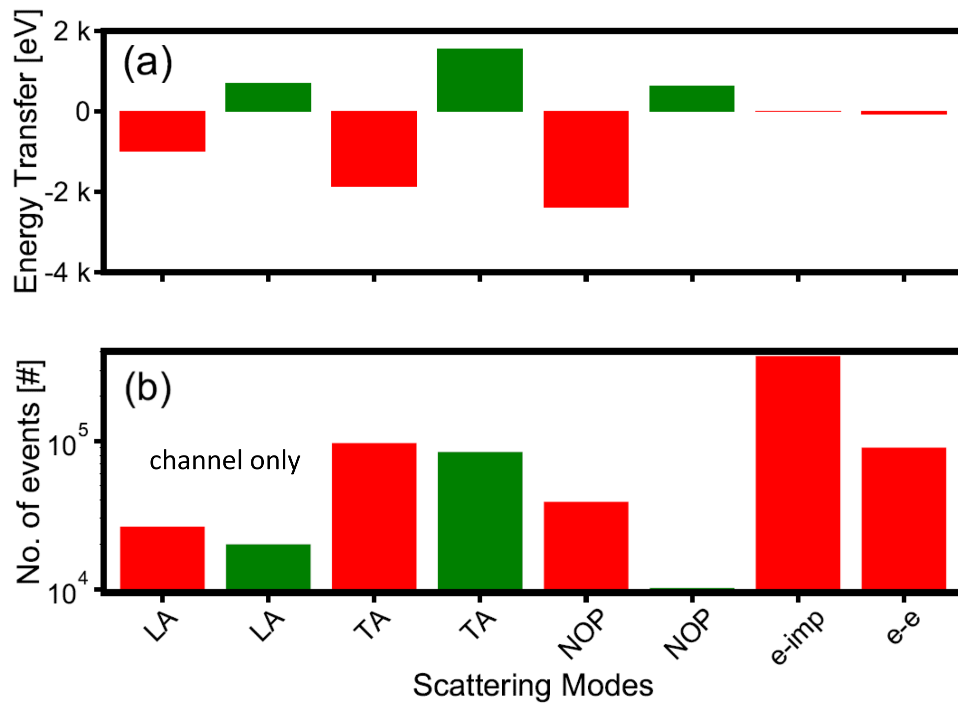

Fig. 19. (a) The number of events (b) and the energy loss or gain for electron scattering with different elastic and phonon modes in the channel region only. Similar conditions to that of Fig. 17 apply. Except e-e scattering, all other momentum transfer events are significantly reduced in the channel regions compared to those plotted including the access regions in Fig. 18.

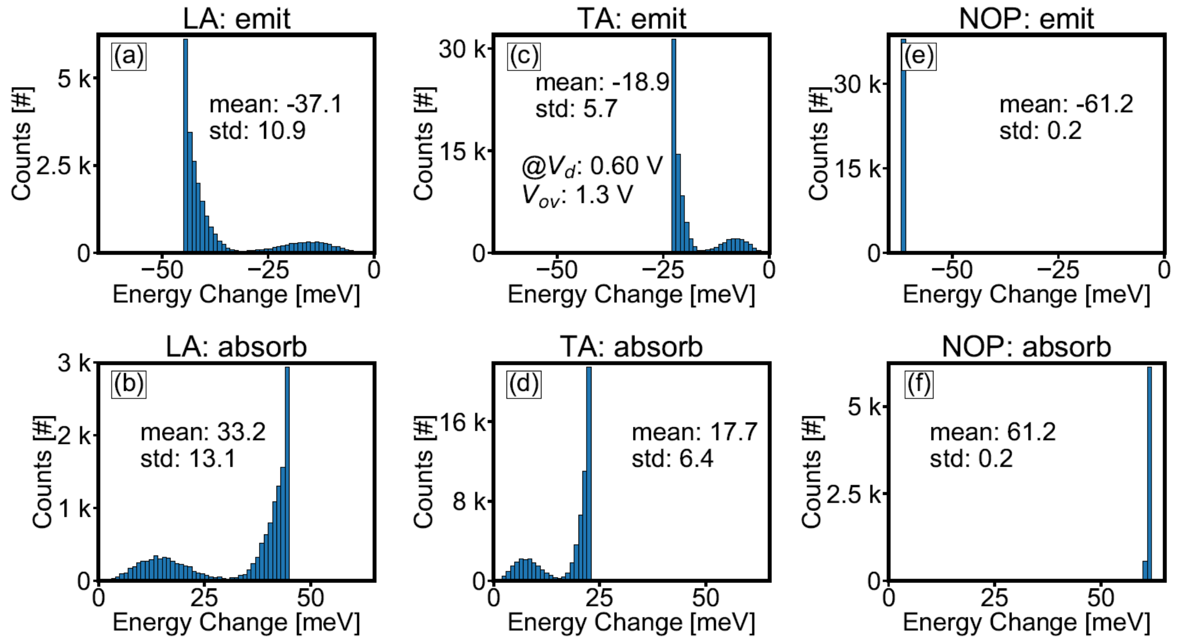

Fig. 20. The energy loss/gain spectrum of the electrons in the channels with the phonon modes considered in the simulation. The range of the interaction in LA, TA, and NOP modes follow the phonon energy bands, as considered in the simulation. For acoustic phonon modes in (a)-(d), electrons emit almost twice energy compared to the energy absorbed. Thus, emission spectrum for the acoustic phonons skewed to higher energies resulting in net loss of energy for the electrons. Energy loss by the NOP phonons in (e) are five times higher than the absorption spectra in (f). NOP interactions occur at a single energy due to the choice a flat energy band for the optical phonons.

Electrons lose energy to the lattice by several phonon processes, namely longitudinal acoustic (LA), transverse acoustic (TA), and non-polar optical modes (TA). We find that the electrons in the access region lose 70% of their energy to NOP photons and rest to the acoustic phonons. This result is consistent with the previous finding for bulk-silicon devices. However, in the channel region, NOP phonon modes account for ~64.5% of the entire energy loss, which is in line with the previous findings (~66%) for n-channel silicon devices. The minor discrepancy in the estimations is due to the differences in the initial distribution function for the electrons.

The energy spectra of the channel-electrons' interaction with six phonon modes are plotted in Fig. 20. Energy loss/gain to both longitudinal and acoustic modes are spread across the entire branch. The shape of the energy-transfer spectrum appears bimodal with a sharp peak at the edge of the zone boundary. The peak at the boundary is associated with the highest DOS value at the zone edge since the slope of the dispersion approaches zero. There are only a few available states for small wavevector values at low energies, and thus the interaction is weaker. In between the zone and center and boundary, the shape of the interaction is associated with the joint density of states. The TA modes are most frequent in the device, whereas the energy transfer is highest with the NOP phonons since the energy of their interaction is the highest. However, the acoustic energy losses are comparable to NOP events and cannot be neglected. The bimodal shape of the acoustic modes results from the chosen deformation potential and the overlap between Bloch states.

5.3 Summary

Coulomb interactions are found to depreciate channel-current under moderate bias conditions. Acoustic and optical phonon processes in the channel are compared to that of the entire device. Acoustic phonons contribute to one-third of all energy losses for the electrons in the channel and need to be included in simulations. In low-field operation, electron-electron scattering restricts channel currents, whereas the availability of the electrons in the source region with the velocity component in the transport direction may dominate the transport in the high-field regime.

Chapter 6

Confinement Barriers for Ultimate Scaling

6.1 Introduction

Quantum effects have been extensively studied in the Si inversion layers where the high-density carriers are enclosed in a triangular potential well between the $Si - SiO_2$ interface and bulk region of the semiconductor [63], [133], [164]. Similar confinement occurs in GAA geometry as well; only the effects are much more profound due to the additional geometric confinements. The transport of the carriers in such an ultra-thin body (UTB) of silicon non-linearly diverges from bulk as a function of the reduced channel-width and geometry [47], [69], [165]. The resultant change in the channel-current and threshold characteristic would be substantial, but the extent of the mobility variation is not agreed upon for a channel width down to 2nm.

The mobility of the carriers in a THB is a complex function of dispersion spectra, carrier density, volume inversion, the interplay among the inversion layers, surface roughness from surrounding sides, ionized-impurity scattering, confinement of the phonon bands, form-factors (overlap between the wavefunctions), many-body effects, and Coulomb scattering [63], [133], [164], [166]. Bulk-like dispersion remains no longer valid since the subband separation turns out to be comparable or higher than kT for nanowire diameters smaller than 4 nm. That is to say that the substantial threshold voltage shift is a direct consequence of the shifted ground-level energy and discretization of the high-energy ladders. The shift in the threshold voltage can be accounted for with an appropriate gate-metal but the reduction of the drive-current is the eventual barrier towards ultimate scaling [4], [55].

Gamiz et. al have treated the mobility variation in the channel as a function of shrinking channel-width and the strength of the applied field [69]. The simulation study is important for comprehensively treating phonon, surface/interface, and Coulomb scattering using a

coupled Poisson-Schrödinger framework. Their simulation study shows that the mobility of the electrons in the inversion layers of a planar double-gate (DG) structure peaks in successions in response to the reduction of channel-width. The mobility first peaks for 10-15 nm channel-width depending on the applied field in the transverse (confined) direction. This increase is found to be associated with volume inversion, which is desirable in these dimensions because it drives the carriers away from the interface and reduces the scattering rate [52], [69]. The second but larger mobility peak occurred between 4 $\longmapsto$ 3nm channel-width and was due to the decreased conductivity effective mass. However, a dramatic decrease in the mobility was simulated at or below 3nm channel-width once the geometric confinements became stronger. This sharper reduction in mobility is due to momentum confinement and the associated paramount increase in Coulomb, phonon, and surface roughness scattering [52], [59]. These results contrast our hope for scaling silicon to 2 nm and represent a significant concern since considerable efforts are ongoing to axially scale the GAA NWs [18], [60], [167].

Since dimensional scaling in the lateral (transport) direction is well understood to hit a brick-wall for silicon due to diminishing return in scaling effective current and tunneling issues at sub-10 nm gate-lengths, it has become essential to investigate the degrees of freedom available in the axial direction [40], [46], [59], [109]. In this chapter, an extension to the methods employed in ***chapter III*** and ***IV*** is discussed to enable semiclassical treating carrier transport of a GAA device with a 2 nm-wide channel. To distinguish the treatment of quantum confinement in 2D to that of bulk 3D simulations, a simplified approach is presented. A GAA devices with a channel-diameter of 2 nm but otherwise similar to that in ref. [68], and ***chapter IV*** is approximated in two dimensions as a DG MOSFET, thanks to the circular symmetry of the nanowire devices throughout.

The discussion in the next chapter should be viewed more as a guideline to get started in solving the electron transport problems whenever quantum confinement effects cannot be

ignored. It is not a literature review on the topic but simply provides the knowledge of the author about semiclassically treating quantum effects. In addition, it provides a route for a possible extension of the works presented in earlier chapters.

Note: *The chosen theoretical and numerical approximations (or flexibility) for presenting or calculating electronic energy bands, phonon energy bands, deformation potentials, discretization of DOS, screening in two-dimensions, Fermi's golden rule, pseudopotential method,* ***k.p*** *method, Coulomb scatterings (e.g., surface roughness, interface traps, oxide defects, impurity, etc.), band-gap narrowing, many body effects, overlap-integral, Bloch-wave representation in nano dimensions, and anisotropy in phonon dispersion may affect the accuracy of the results obtained by semiclassical simulations. Nevertheless, the sensitivity of these factors is well-known for silicon since most employed data, approximations, and numerical methods are already fitted or verified to experimental findings, thanks to decades of research in this. The literature in each of them are huge and discussing their treatment suitable for reduced dimensionality is outside the scope of this chapter. Interested readers may read the book in ref. [52] and contents in refs.* [58], [59], [67], [128], [133], [137], [166], [168] *to get an overall perspective.*

Only the factors already identified to be the most important for semiclassically simulating Si NW FETs are discussed here and associated references are mentioned for treating them in a way the author deems accurate. Also, the author acknowledges that other popular methods such as NEGF are an alternative to semiclassical MC method for similar transport questions and they can sometime provide complementary results to the questions asked here. Again, the literature in both is huge and comparing them is outside the scope of this chapter.

6.2 Discussions

Quantum confinement effects can be treated within the semiclassical framework by solving the one dimensional Schrödinger equation and coupling it to the self-consistent Poisson-

Boltzmann formulation [55], [133]. The electronic transport for the carriers in a quantum well differs from that of the bulk in three major areas: i) carriers become subjected to a discretized dispersion in k-space, ii) particle density in real space depends on the squared wavefunction associated with the Schrödinger formulation iii) scattering rates become considerably different as the density of states gets substantially modified. The eigen solution of the problem is typically obtained by solving the effective mass Schrödinger equation with approximations [52]. Resultant computed subbands are simply a consequence of defining appropriate boundary conditions on the oxide-semiconductor interfaces [69], [133]. For the rest of the chapter, simulations with and without treating quantum confinement are respectively referred as 2D and 3D transport, in line with the adapted convention in the literature.

In theory, the quantized dispersion can be derived by solving an effective-mass Schrodinger equation with appropriate boundary conditions. But the precise solution for the realistic band-structure is a function of both wavevector and energy and thus becomes computationally prohibitive to treat [52]. Moreover, even after ignoring the wave-vector dependence, calculating the DOS and scattering rates asks for unrealistic computational resources [70]. The scenario differs from the 3D transport, where both the dispersion and phonon scattering rates are precomputed and available in a tabulated form. However, both DOS and scattering rates in 2D simulation depend on dynamic variables (e.g., applied bias), and cannot be used precomputed [133]. Ultimately, a cruder analytical dispersion with a non-parabolicity parameter is shown to be acceptable till 0.5 eV and is suitable for computing scattering rates and screening effects in a 2D regime [133]. A complete numerical dispersion is still required for addressing electron transport in energies higher than 0.5 eV [68], [83]. In the past, tens of subbands were used in simulating transport till 0.5 eV for treating quantization in an inversion layer of a planar DG MOSFET. However, for UTB devices, only a few subbands are needed to fill up the energy ladder till 0.5 eV in view of the stronger

geometric confinement at sub-4 nm diameters. Consequently, nanowires with a smaller radius are computationally favorable for treating with semiclassical means.

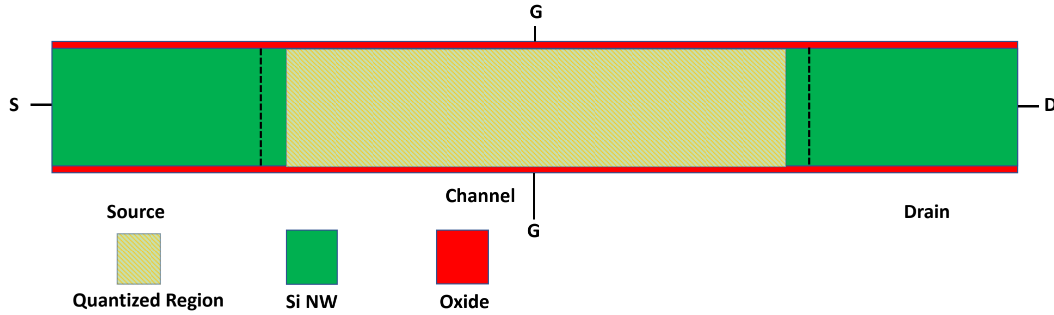

Fig. 21. Two-dimensional approximation of an idealistic circular NW nMOSFET. Effective mass Schrodinger equation is solved only on the region shown in beige (light-brown). The metallurgical junction boundaries in the source and drain side are indicated with dotted lines. The illustration is not drawn to scale.

Mobility in UTB Si nanowires is strongly influenced by phonon and surface roughness (SR) scattering. In theory, phonon scattering increases inversely to the squared power of the channel-diameter. However, this dependence is dwarfed by SR scattering, which overshoots to the sixth-power of thinning the channel [61], [63], [69], [79]. Jin et al. have simulated the effects of SR, phonon, and Coulomb scattering to the size-reduction of the channel diameter down to 3nm [47], [63]. Their results show that electron mobility for the low-field regime decreases monotonically, whereas the high-field mobility peaks around 5 nm channel-diameter. This peak at high-field is due to the fallen scattering rates in consideration of the shrunk density of states at reduced diameter-size. However, increased SR interaction with further dimensional-scaling soon takes over, and mobility falls drastically at 2 nm diameters. Log-linear extrapolation from their data shows that SR scattering would increase ten-fold if the channel diameter is reduced from 8 to 2 nm. In addition, screening effects were found to elevate the effective-field and but increase overall mobility, whereas non-parabolicity parameters were found to affect the accuracy by 20%. In addition, phonon-limited mobility was observed to decrease four times with diameter shrinking from 5 to 3nm. These results

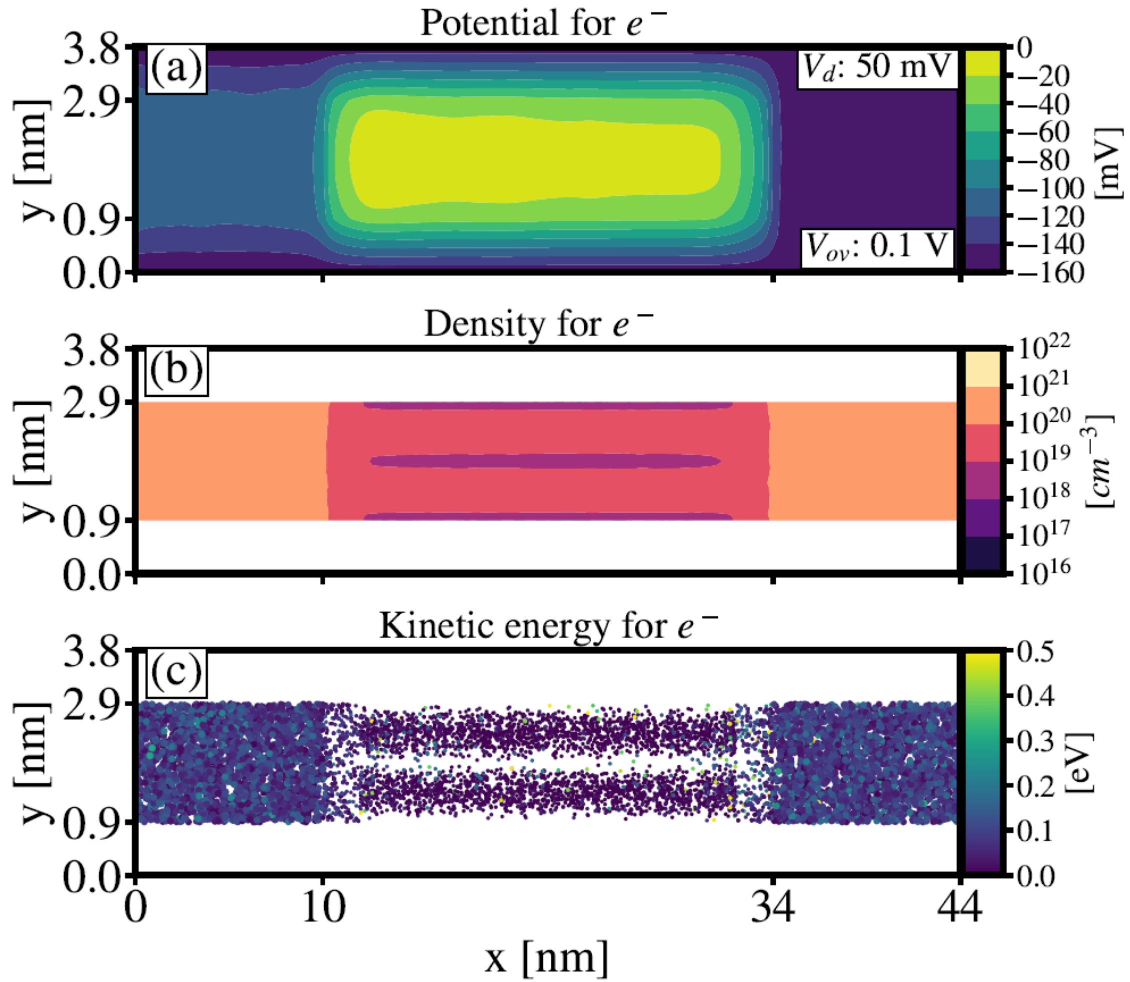

Fig. 22. Simulated (a) electrostatic potential (b) charge density and (c) distribution of the charge carriers in the device. The maximum of the potential in (a) is arbitrary chosen as zero. Source-junction barriers is highest along the mid-channel for electron conduction from source (left) to drain (right). This barrier drops from the center towards the interface. The charge density for the electrons in (b) show two separate inversion layers associated to the second subband. The electric field in the transverse direction is less than needed for volume inversion. Illustrated in (c), electrons in the quantized region $(12 < x < 32)$ are distributed in real space accordingly to the squared wavefunction and in k-space accordingly to their energy. Electrons populate the first and second subband up to 0.15 eV and are treated as bulk at higher energies. The different sizes of the particles reflect their differences in charge-weight on logscale.

are consistent with the trend observed during semiclassically treating the quantization in the inversion layer of planar silicon FETs, except the effects are pronounced for extreme geometrical scaling. The strong effect of SR scattering at 2 nm can be explained in one

sentence — surface charges and potential in the interface do influence all channel carriers within a radius of 2 nm due to close proximity [169]. The stronger phonon scattering at 2 nm diameter is due to the strong overlap of the wavefunctions in the confined geometry [52]. Overall, it is necessary to treat the aforementioned scatterings for UTB devices, especially at sub-4nm diameters for silicon NWs.

The primary challenge in a semiclassical simulation is to account for these overlap integrals and consequently calculate the wave-vector dependent strength of scattering, namely matrix element [52]. The scattering rate is calculated using Fermi's golden rule, similar to bulk [68], [71], [83], [133]. Mathematical methods for treating these scattering events are detailed in refs. [61], [65], [66], [137]. Interested readers may look at these references before designing a semiclassical simulation including quantization effects.

Next, quantum confinement effects are illustrated using a semiclassical simulation with minimal rigor. To bypass the computational strain for solving the Schrödinger equation in two dimensions, the GAA device is represented in a two-dimensional mesh. For the idealistic circular symmetry considered for the NW MOSFETs in ***chapter IV***, an equivalent DG representation is a valid approximation. In doing so, the Schrödinger equation would need to be solved only in one dimension, and the Poisson equation would be treated in two dimensions. MC particles are represented as line charges in the other direction not included in the Poisson mesh.

The dimension and geometry for the simulated device is shown in Fig. 21. The effective-mass Schrödinger equation (SE) is solved for the light brown regions within the channel under applied low-field conditions. In the selected region, potential variation is minimal and thus numerically favorable for treating the transport by SE. To reduce the computational burden further, a parabolic energy dispersion is assumed for deriving the two subbands (ground and first excited state) employed in the calculation. Electrons with energies above the second subband are treated with the numerical electronic band without quantization. Such

an approximation is gross and far from what is ideal for attaining optimum accuracy in the results. But these approximations are used here only to qualitatively show the quantization effects without quantitatively measuring the extent of it or comparing it to the existing literature. Larger energy separation of the quantized subbands for a UTB restricts the use of many subbands, as the parabolic approximation is considered valid only up to 50 meV [61], [137]. Acoustic and non-polar phonon modes are accounted for, but SR and carrier-carrier scattering, wavefunction penetration in the oxide were not included in the simulations. A similar 2D treatment including these mechanisms is detailed in ref. [133] while treating the confined carriers in the inversion layer of an nMOSFET and must be included to simulate UTB devices.

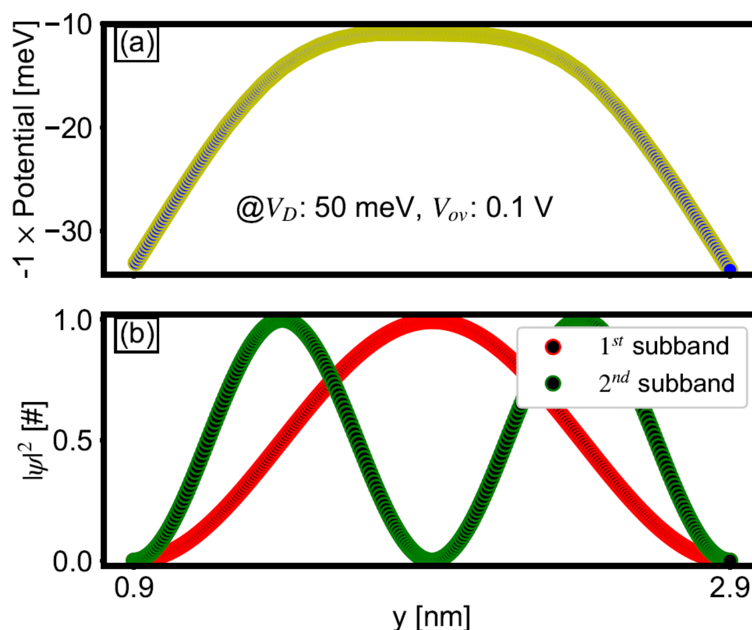

Fig. 23. (a) Potential for the electrons in the quantized region along the transverse direction at mid-channel. Corresponding wavefunctions for the ground (red) and first excited state (green) unprimed subbands.

Under low-field conditions, simulated electrostatic potential, charge-density, and electron distribution in real and k space are presented in Fig. 22. The applied electric field is weak to induce a volume inversion, but the effect of employing boundary conditions in SE is visible

in the charge density and electron distribution plot, especially near the oxide-semiconductor interface. Only electrons with energy up to the 2nd subband (1st excited state) were treated with quantum solutions. Wavefunctions in the quantized region for the first and second subbands are shown in Fig. 23. As can be realized by comparing Fig. 23(a) to Fig. 23(b), the first subband is mostly empty while the second subband is highly populated. Such a scenario exists because the access regions considered in the simulation are not large enough for electrons to thermalize to equilibrium distributions. Thus, the carriers' average energy entering the quantized region is higher than appropriate and induces errors while populating the subbands. Of course, in real devices, the access regions are larger and conically expands towards the contacts. Using a similar approach, volume inversion was observed (not shown here) at a higher transverse field [0.5-5 MeV].

Including the impact of surface roughness, interface traps, and oxide traps within the simulation would respectively require the knowledge of the associated (i) scatterer potential and screening effects ii) density of interface traps iii) location and density of oxide traps. In the same order, (1) can be derived from simulations similar to refs. [47], [133] and (ii-iii) can be experimentally measured and can be incorporated in the calculations with a method similar to [133]. In addition, form-factors associated with the calculation of phonon scattering rates for scaled nanowire structure can be found in experimental studies similar to ref. [170].

6.3 Summary

The means for simulating an ultimately scaled GAA transistors are discussed. Mechanisms that would be important to enable extreme scaling are highlighted. More importantly, techniques for treating quantum confinement effects are displayed separately as an addition to the regular simulation framework for bulk devices.

Chapter 7

Conclusion

The Monte Carlo technique is employed for self-consistently solving the coupled Boltzmann-Poisson transport equations to study electron transport in materials and devices. Simulation results match well with the experimentally derived average ionization energy for silicon and germanium, drive current, and device lifetime for GAA Si MOSFETs. In summary, it appears that computing (semiclassically) carrier kinetics directly from fundamental physical quantities such as the mass, charge, electric field, etc., provides realistic estimates of electronic properties at least down to 8 nm length scale (for Silicon), given that realistic electronic and phonon density of states are integrated into the simulation.

Simulated carrier kinetics in the depletion region of a reverse-biased PN junction show that the energy deposition process (excluding collisions with the nuclei) in the semiconducting materials are independent of the incident particle (e.g., photon, electron) properties, which is in line with several experimental studies beforehand. The major contribution of this work is to show that the average energy required for electron-hole pair creation appears from optical/electron absorption spectra, threshold energy of the impact ionization events, and scattering loss due to phonons and impact ionization events. Absorption spectra and impact ionization cross-section for silicon being well known by experimental means, an excellent match to the experiments validates the employed methods for treating energy kinetics in large area and low-density silicon devices.

Simulated electron transport properties for GAA Si NW MOSFETs show that Coulomb processes, which are often ignored as a higher-order phenomenon, reduce mobility significantly at low-field conditions. Plasma effects, electron degeneracy, and many-body interactions of the fermions are shown to elevate a population of electrons in the access region to much higher energies than the energy of the equilibrated carriers at room

temperature without considering these scattering events. Not all electrons in the channel scatter due to phonon processes such that the transport is quasi-ballistic, but a fraction of those that do significantly extend the tail of the energy distribution. The average energy and density of the electrons populating this tail correlate well to the device degradation due to applied drain-gate (coupled) and drain field, respectively. A significant population of electrons in the drain rises above the energy required to release hydrogens from the passivated phosphorus dopant complexes. Whether this is a competing process to the direct release of hydrogens by the hot carriers in the interface would require a first principle investigation.

Semiclassical and first principle simulations have become very important ever since the process technologies for silicon have matured in fabricating devices at atomic resolution. Considering that the semiconductor performance is becoming a much stronger function of the physics intervention than shrinking effects, physics-based methods of accounting geometric and potential confinements are discussed in view of their appropriateness in simulating cylindrical silicon devices with an ultra-thin channel. Without such simulations, performance scaling may fall behind the Moore's curve, a phenomenon long predicted but always proven untrue.

References

[1] J. Ryckaert, M. H. Na, P. Weckx, D. Jang, P. Schuddinck, B. Chehab, S. Patli, S. Sarkar, O. Zografos, R. Baert, and D. Verkest, "Enabling sub-5nm CMOS technology scaling thinner and taller!," *Tech. Digest - Int. Electron Devices Meet. (IEDM)*, Dec. 2019.

[2] R. Ritzenthaler, H. Mertens, V. Pena, G. Santoro, A. Chasin, K. Kenis, K. Devriendt, G. Mannaert, H. Dekkers, A. Dangol, Y. Lin, S. Sun, Z. Chen, M. Kim, J. Machillot, J. Mitard, N. Yoshida, N. Kim, D. Mocuta, and N. Horiguchi, "Vertically stacked gate-all-around Si nanowire CMOS transistors with reduced vertical nanowires separation, new work function metal gate solutions, and DC/AC performance optimization," *Tech. Digest - Int. Electron Devices Meet. (IEDM),* pp. 21.5.1-21.5.4, Dec. 2018.

[3] E. Mohapatra, T. P. Dash, J. Jena, S. Das, and C. K. Maiti, "Strain induced variability study in Gate-All-Around vertically-stacked horizontal nanosheet transistors," Phys. Scr., vol. 95, no. 6, pp. 65808-65814, Jun. 2020.

[4] H. Mertens, R. Ritzenthaler, V. Pena, G. Santoro, K. Kenis, A. Schulze, E. D. Litta, S. A. Chew, K. Devriendt, R. Chiarella, S. Demuynck, D. Yakimets, D. Jang, A. Spessot, G. Eneman, A. Dangol, P. Lagrain, H. Bender, S. Sun, M. Korolik, D. Kioussis, M. Kim, K. H. Bu, S. C. Chen, M. Cogorno, J. Devrajan, J. Machillot, N. Yoshida, N. Kim, K. Barla, D. Mocuta, and N. Horiguchi, "Vertically stacked gate-all-around Si nanowire transistors: Key process optimizations and ring oscillator demonstration," *Tech. Digest - Int. Electron Devices Meet. (IEDM),* pp. 37.4.1-37.4.4, Dec. 2018.

[5] H. Mertens, R. Ritzenthaler, A. Chasin, T. Schram, E. Kunnen, A. Hikavyy, L. Å. Ragnarsson, H. Dekkers, T. Hopf, K. Wostyn, K. Devriendt, S. A. Chew, M. S. Kim, Y. Kikuchi, E. Rosseel, G. Mannaert, S. Kubicek, S. Demuynck, A. Dangol, N. Bosman, J. Geypen, P. Carolan, H. Bender, K. Barla, N. Horiguchi, and D. Mocuta, "Vertically stacked gate-all-around Si nanowire CMOS transistors with dual work function metal gates," *Tech. Digest - Int. Electron Devices Meet. (IEDM),* pp. 19.7.1-19.7.4, Dec. 2016.

[6] A. Chasin, J. Franco, B. Kaczer, V. Putcha, P. Weckx, R. Ritzenthaler, H. Mertens, N. Horiguchi, D. Linten, and G. Rzepa, "BTI reliability and time-dependent variability of stacked gate-all-around Si nanowire transistors," 2017 IEEE Intl. Reliability Physics Symposium, Monterey, CA, pp. 5C-4.1-5C-4.7, Apr. 2017.

[7] A. Chasin, E. Bury, B. Kaczer, J. Franco, P. Roussel, R. Ritzenthaler, H. Mertens, N. Horiguchi, D. Linten, and A. Mocuta, "Complete degradation mapping of stacked gate-all-around Si nanowire transistors considering both intrinsic and extrinsic effects," *IEEE Int. Electron Devices Meet.*, San Francisco, CA, USA, pp. 7.1.1-7.1.4, Dec. 2017.

[8] F. M. Bufler, R. Ritzenthaler, H. Mertens, G. Eneman, A. Mocuta, and N. Horiguchi, "Performance comparison of n-type Si nanowires, nanosheets, and FinFETs by MC device simulation," *IEEE Electron Device Lett.*, vol. 39, no. 11, pp. 1628–1631, Nov. 2018.

[9] A. Chasin, E. Bury, J. Franco, B. Kaczer, G. Rzepa, V. Putcha, P. Weckx, R. Ritzenthaler, H. Mertens, N. Horiguchi, and D. Linten, "Reliability in stacked gate-all-around Si nanowire devices: Time-dependent variability and full degradation mapping," *IEEE Int. Integrated Reliability Workshop*, South Lake Tahoe, CA, USA, pp. 39-46, Oct. 2019.

[10] M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov, "Modeling of nanoscale devices," *Proc. IEEE*, vol. 96, no. 9, pp. 1511–1550, Sep. 2008.

[11] K.C. Wang, R. Grassi, Y. Chu, S. H. Sureshbabu, J. Geng, P. Sarangapani, X. Guo, M. Townsend, and T. Kubis, "Introduction of multi-particle Büttiker probes - Bridging the gap between drift diffusion and quantum transport," *J. Appl. Phys.*, vol. 128, no. 1, pp. 014302, Jul. 2020.

[12] N. Seoane, A. Martinez, A. R. Brown, J. R. Barker, and A. Asenov, “Current variability in Si nanowire MOSFETs due to random dopants in the source/drain regions: A fully 3D NEGF simulation study,” *IEEE Trans. Electron Devices*, vol. 56, no. 7, pp. 1388–1395, Jul. 2009.

[13] M. V. Fischetti and S. E. Laux, “Monte Carlo study of sub-band-gap impact ionization in small silicon field-effect transistors,” *Techn. Dig. - Int. Electron Devices Meet.*, pp. 305–308, Dec. 1995.

[14] M. V. Fischetti, “Long-range Coulomb interactions in small Si devices. Part II. Effective electron mobility in thin-oxide structures,” *J. Appl. Phys*., vol. 89, no. 2, pp. 1232–1250, Jan. 2001.

[15] M. S. Lundstrom and D. A. Antoniadis, “Compact models and the physics of nanoscale FETs,” *IEEE Trans. Electron Devices*, vol. 61, no. 2, pp. 225–233, Feb. 2014.

[16] L. Tsetseris, X. J. Zhou, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, “Physical mechanisms of negative-bias temperature instability,” *Appl. Phys. Lett.*, vol. 86, no. 14, Mar. 2005, Art. no. 142103.

[17] S. E. Rauch and G. La Rosa, “The energy-driven paradigm of NMOSFET hot-carrier effects,” *IEEE Trans. Device Mater. Reliab*., vol. 5, no. 4, pp. 701–705, Dec. 2005.

[18] U. K. Das and T. K. Bhattacharyya, “Opportunities in device scaling for 3-nm node and beyond: FinFET versus GAA-FET versus UFET,” *IEEE Trans. Electron Devices*, vol. 67, no. 6, pp. 2633–2638, Jun. 2020.

[19] A. Suzuki, T. Kamioka, Y. Kamakura, K. Ohmori, K. Yamada, and T. Watanabe, “Source-induced RDF overwhelms RTN in nanowire transistor: statistical analysis with full device EMC/MD simulation accelerated by GPU computing,” *Tech Digest - Int. Electron Devices Meet. (IEDM)*, pp. 30.1.1-30.1.4, Feb. 2015.

[20] K. Nayak, S. Agarwal, M. Bajaj, K. V. R. M. Murali, and V. R. Rao, “Random dopant fluctuation induced variability in undoped channel si gate all around nanowire n-MOSFET,” *IEEE Trans. Electron Devices*, vol. 62, no. 2, pp. 685–688, Feb. 2015.

[21] R. Wang, J. Zhuge, R. Huang, T. Yu, J. Zou, D. Kim, D. Park, and Y. Wang, “Investigation on variability in metal-gate Si nanowire MOSFETs: Analysis of variation sources and experimental characterization,” *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2317–2325, Aug. 2011.

[22] S. Dey, J. Jena, E. Mohapatra, T. P. Dash, S. Das, and C. K. Maiti, “Design and simulation of vertically-stacked nanowire transistors at 3nm technology nodes,” *Phys. Scr*., vol. 95, no. 1, p. 014001, Dec. 2020.

[23] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, “Metal Grain Granularity Study on a Gate-All-Around Nanowire FET,” *IEEE Trans. Electron Devices*, vol. 64, no. 12, pp. 5263–5269, Dec. 2017.

[24] K. Nayak, S. Agarwal, M. Bajaj, P. J. Oldiges, K. V. R. M. Murali, and V. R. Rao, “Metal-gate granularity-induced threshold voltage variability and mismatch in Si gate-all-around nanowire n-MOSFETs,” *IEEE Trans. Electron Devices*, vol. 61, no. 11, pp. 3892–3895, Nov. 2014.

[23] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, “Metal grain granularity study on a gate-all-around nanowire FET,” *IEEE Trans. Electron Devices*, vol. 64, no. 12, pp. 5263–5269, Dec. 2017.

[24] K. Nayak, S. Agarwal, M. Bajaj, P. J. Oldiges, K. V. R. M. Murali, and V. R. Rao, “Metal-gate granularity-induced threshold voltage variability and mismatch in Si gate-all-around nanowire n-MOSFETs,” *IEEE Trans. Electron Devices*, vol. 61, no. 11, pp. 3892–3895, Nov. 2014.

[25] S. M. Nawaz, S. Dutta, A. Chattopadhyay, and A. Mallik, “Comparison of random dopant and gate-metal workfunction variability between junctionless and conventional FinFETs,” *IEEE Electron Device Lett.*, vol. 35, no. 6, pp. 663–665, 2014.

[26] T. C. Ong, P. K. Ko, and C. Hu, “Hot-Carrier Current Modeling and Device Degradation in Surface-Channel p-MOSFET’s,” *IEEE Trans. Electron Devices*, vol. 37, no. 7, pp. 1658–1666,

1990.
[27] D. Rossi, F. Santoni, M. Auf Der Maur, and A. Di Carlo, “A multiparticle drift-diffusion model and its application to organic and inorganic electronic device simulation,” *IEEE Trans. Electron Devices*, vol. 66, no. 6, pp. 2715–2722, Jun. 2019.
[28] P. Sharma, M. Jech, S. Tyaginov, F. Rudolf, K. Rupp, H. Enichlmair, J. Park, and T. Grasser, “Modeling of hot-carrier degradation in LDMOS devices using a drift-diffusion based approach,” Intl. Conference on Simulation of Semiconductor Processes and Devices, Oct. 2015, pp. 60–63.
[29] R. Kotlyar, R. Rios, C. E. Weber, T. D. Linton, M. Armstrong, and K. Kuhn, “Distributive quasi-ballistic drift diffusion model including effects of stress and high driving field,” *IEEE Trans. Electron Devices*, vol. 62, no. 3, pp. 743–750, 2015.
[30] M. Lundstrom, “Drift-diffusion and computational electronics - Still going strong after 40 years!,” in International Conference on Simulation of Semiconductor Processes and Devices, SISPAD, 2015, vol. 2015-October, pp. 1–3.
[31] S. Rakheja, M. S. Lundstrom, and D. A. Antoniadis, “An improved virtual-source-based transport model for quasi-ballistic transistors - Part II: Experimental verification,” *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2794–2801, Sep. 2015.
[32] S. Rakheja, M. S. Lundstrom, and D. A. Antoniadis, “An improved virtual-source-based transport model for quasi-ballistic transistors-Part I: Capturing effects of carrier degeneracy, Drain-Bias Dependence of Gate Capacitance, and Nonlinear Channel-Access Resistance,” *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2786–2793, Sep. 2015.
[33] K. Hess, A. Haggag, W. McMahon, K. Cheng, J. Lee, and J. Lyding, “The physics of determining chip reliability,” *IEEE Circuits Devices Mag.*, vol. 17, no. 3, pp. 33–38, May 2001.
[34] K. Hess, B. Tuttle, F. Register, and D. K. Ferry, “Magnitude of the threshold energy for hot electron damage in metal-oxide-semiconductor field effect transistors by hydrogen desorption,” *Appl. Phys. Lett.*, 1999.
[35] D. K. Ferry, S. M. Goodnick, and K. Hess, “Energy exchange in single-particle electron-electron scattering,” *Phys. B Condens. Matter*, vol. 272, no. 1–4, pp. 538–541, Dec. 1999.
[36] K. Hess et al., “Theory of channel hot-carrier degradation in MOSFETs,” Phys. B Condens. Matter, vol. 272, no. 1–4, pp. 527–531, Dec. 1999.
[37] M. Vandemaele, B. Kaczer, S. Tyaginov, Z. Stanojević, A. Makarov, A. Chasin, E. Bury, H. Mertens, D. Linten, and G. Groeseneken, “Full (V_g, V_d) bias space modeling of hot-carrier degradation in nanowire FETs,” *Intl. Reliability Physics Symp.*, Monterey, CA, USA, pp. 6C.3.1-6C.3.7, Apr. 2019.
[38] M. V. Fischetti, L. Wang, B. Yu, C. Sachs, P. M. Asbeck, Y. Taur and M. Rodwell, "Simulation of electron transport in high-mobility MOSFETs: Density of states bottleneck and source starvation," *Tech. Dig. - Int. Electron Devices Meet. (IEDM*), Washington, DC, pp. 109-112, Dec 2007.
[39] M. V. Fischetti, S. Jin, T. W. Tang, P. Asbeck, Y. Taur, S. E. Laux, M. Rodwell and N. Sano, “Scaling MOSFETs to 10 nm: Coulomb effects, source starvation, and virtual source model,” *J. Comput. Electron.*, vol. 8, no. 2, pp. 60–77, Jul. 2009.
[40] D. J. Frank, S. E. Laux, and M. V. Fischetti, “Monte Carlo simulation of a 30 nm dual-gate MOSFET: How short can Si go?,” *Tech. Digest - Int. Electron Devices Meet. (IEDM)*, pp. 553–556, Dec. 1992.
[41] Z. Liana, L. Khiangte, S. Dash, and R. S. Dhar, “Drift diffusion and advance hydrodynamic simulation for the design of double-gate SOI MOSFET in nano-scale regime,” *2nd Int. Conf. on Adv. Comp. and Communication Paradigms (ICACCP)*, Feb. 2019.
[42] T. Grasser, T. W. Tang, H. Kosina, and S. Selberherr, “A review of hydrodynamic and energy-

transport models for semiconductor device simulation," *Proc. IEEE*, vol. 91, no. 2, pp. 251–273, Feb. 2003.

[43] T. W. Tang, S. Ramaswamy, and J. Nam, "An improved hydrodynamic transport model for silicon," *IEEE Trans. Electron Devices*, vol. 40, no. 8, pp. 1469–1477, Aug. 1993.

[44] M. V. Fischetti and S. E. Laux, "Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures - part II: Submicrometer MOSFETs," *IEEE Trans. Electron Devices*, vol. 38, no. 3, pp. 650–660, Mar. 1991.

[45] M. V. Fischetti, S. E. Laux, and E. Crabbé, "Understanding hot-electron transport in silicon devices: Is there a shortcut?," *J. Appl. Phys.*, vol. 78, no. 2, pp. 1058–1087, Jul. 1995.

[46] M. V. Fischetti and S. E. Laux, "Long-range Coulomb interactions in small Si devices. Part I: Performance and reliability," *J. Appl. Phys.*, vol. 89, no. 2, pp. 1205–1231, Jan. 2001.

[47] S. Jin, M. V. Fischetti, and T. W. Tang, "S. Jin, M. V. Fischetti, and T. W. Tang, "Modeling of surface-roughness scattering in ultrathin-body SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2191–2203, Sep. 2007.

[48] H. Ness, L. K. Dash, and R. W. Godby, "Generalization and applicability of the Landauer formula for nonequilibrium current in the presence of interactions," *Phys. Rev. B* - Condensed Matter & Mater. Phys., vol. 82, no. 8, p. 085426, Aug. 2010.

[49] S. Datta, "Non-equilibrium green's function (NEGF) method: A different perspective," *18th Int. Workshop on Computational Electronics (IWCE)*, pp. 1-6, Sep. 2015.

[50] H. H. Park, W. Choi, M. A. Pourghaderi, J. Kim, U. Kwon, and D. S. Kim, "NEGF simulations of stacked silicon nanosheet FETs for performance optimization," *Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sep. 2019.

[51] W. Goes, Y. Wimmer, Å. M. El-Sayeda, G. Rzepaa, M. Jecha, A. L. Shluger, and T. Grasser, "Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence," *Microelectron. Reliab.*, vol. 87, pp. 286–320, Aug. 2018.

[52] M. V. Fischetti and W. G. Vandenberghe, Advanced Physics of Electron Transport in Semiconductors and Nanostructures, Cham, Switzerland: Springer Int., 2016.

[53] M. Aldegunde, A. Martinez, and A. Asenov, "Non-equilibrium Green's function analysis of cross section and channel length dependence of phonon scattering and its impact on the performance of Si nanowire field effect transistors," *J. Appl. Phys.*, vol. 110, no. 9, pp. 94518, Nov. 2011.

[54] A. Svizhenko and M. P. Anantram, "Role of scattering in nanotransistors," *IEEE Trans. Electron Devices*, vol. 50, no. 6, pp. 1459–1466, Jun. 2003.

[55] L. Donetti, C. Sampedro, F. Gámiz, A. Godoy, F. J. García-Ruíz, E. Towiez, V. P. Georgiev, S. M. Amoroso, C. Riddet, and A. Asenov, "Multi-subband ensemble Monte Carlo simulation of Si nanowire MOSFETs," *Intl. Conf. on Simulation of Semiconductor Proc. and Dev.*, Washington, DC, pp. 353-356, Sep. 2015.

[56] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H. S. P. Wong, "*Device scaling limits of Si MOSFETs and their application dependencies*," Proc. IEEE, vol. 89, no. 3, pp. 259–287, Mar. 2001.

[57] F. G. Pikus and K. K. Likharev, "Nanoscale field-effect transistors: An ultimate size analysis," *Appl. Phys. Lett.*, vol. 71, no. 25, pp. 3661–3663, Dec. 1997.

[58] J. Fang, S. Chen, W. G. Vandenberghe, and M. V. Fischetti, "Theoretical study of ballistic transport in silicon nanowire and graphene nanoribbon field-effect transistors using empirical pseudopotentials," *IEEE Trans. Electron Devices*, vol. 64, no. 6, pp. 2758-2764, Jun. 2017.

[59] M. V Fischetti, S. E. Laux, P. M. Solomon, and A. Kumar, "Thirty Years of Monte Carlo Simulations of Electronic Transport in Semiconductors: Their Relevance to Science and

Mainstream VLSI Technology," *J. Comput. Electron.*, vol. 3, no. 3–4, pp. 287–293, Oct. 2004.

[60] T. P. Dash, S. Dey, S. Das, E. Mohapatra, J. Jena, and C. K. Maiti, "Strain-engineering in nanowire field-effect transistors at 3 nm technology node," *Phys. E Low-Dimensional Syst. Nanostructures*, vol. 118, pp. 113964, Apr. 2020.

[61] M. V. Fischetti, T. P. O'Regan, S. Narayanan, C. Sachs, S. Jin, J. Kimand, and Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nMOSFETs at the 10-nm gate-length," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2116-2136, Sep. 2007.

[62] S. Sinha, E. Pop, R. W. Dutton, and K. E. Goodson, "Non-equilibrium phonon distributions in sub-100 nm silicon transistors," *J. Heat Transfer*, vol. 128, no. 7, pp. 638–647, Jul. 2006.

[63] S. Jin, Y. J. Park, and H. S. Min, "A three-dimensional simulation of quantum transport in silicon nanowire transistor in the presence of electron-phonon interactions," *J. Appl. Phys.*, vol. 99, no. 12, p. 123719, Jun. 2006.

[64] G. Gaddemane, "Theoretical studies of electronic transport in two-dimensional materials for transistor applications," Ph.D. dissertation, Dept. of Mat. Sci. & Eng., Univ. of Texas at Dallas, Dallas, TX, Dec. 2018, Accessed on: June 19, 2021. [Online]. Available: https://utd-ir.tdl.org/handle/10735.1/6372

[65] M. V. Fischetti, "Monte Carlo solution to the problem of high-field electron heating in SiO2," *Phys. Rev. Lett.*, vol. 53, no. 18, pp. 1755–1758, Oct. 1984.

[66] M. V. Fischetti, "Erratum: 'Theory of electron transport in small semiconductor devices using the Pauli master equation' [J. Appl. Phys. 83, 270 (1998)]," *J. Appl. Phys.*, vol. 83, no. 11, p. 6202, Jun. 1998.

[67] A. Martinez, M. Bescond, J. R. Barker, A. Svizhenko, M. P. Anantram, C. Millar, and A. Asenov, "A self-consistent full 3-D real-space NEGF simulator for studying nonperturbative effects in nano-MOSFETs," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2213–2222, Sep. 2007.

[68] M. Reaz, A. M. Tonigan, K. Li, M. B. Smith, M. W. Rony, M. Gorchichko, A. O'Hara, D. Linten, J. Mitard, J. Fang, E. X. Zhang, M. L. Alles, R. A. Weller, D. M. Fleetwood, R. A. Reed, M. V. Fischetti, S. T. Pantelides, S. L. Weeden-Wright, and R. D. Schrimpf, "3-D Full-band Monte Carlo simulation of hot-electron energy distributions in gate-all-around Si nanowire MOSFETs," *IEEE Trans. Electron Devices*, vol. 68, no. 5, pp. 2556-2563, May 2021, doi: 10.1109/TED.2021.3068328.

[69] F. Gamiz and M. V. Fischetti, "Monte Carlo simulation of double-gate silicon-on-insulator inversion layers: The role of volume inversion," *J. Appl. Phys.*, vol. 89, no. 10, pp. 5478–5487, May 2001.

[70] M. Fischetti and S. E. Laux, "Damocles Theoretical Manual." IBM Corporation, Yorktown Heights, NY, 1995.

[71] M. V. Fischetti and S. E. Laux, "Monte carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev. B*, vol. 38, no. 14, pp. 9721–9745, Nov. 1988.

[72] V. Heine, "Electronic structure from the point of view of the local atomic environment," *Solid State Phys. - Adv. Res.* Appl., vol. 35, no. C, pp. 1–127, Jan. 1980.

[73] M. V. Fischetti and J. M. Higman, Monte Carlo Device Simulation: Full Band and Beyond. Kluwer Academic, Norwell, MA, 1991.

[74] J.-L. van der Steen, D. Esseni, P. Palestri, L. Selmi, and R. J. E. Hueting, "Validity of the parabolic effective mass approximation in silicon and germanium n-MOSFETs with different crystal orientations," *IEEE Trans. Electron Devices*, vol. 54, no. 8, pp. 1843–1851, Aug. 2007.

[75] S. E. Laux, M V Fischetti, and D. J. Frank, "Monte Carlo analysis of semiconductor devices: The DAMOCLES program," *IBM J. Res. Dev.*, vol. 34, no. 4, pp. 466–494, Jul. 1990.

[76] S. Katayama, D. L. Mills, and R. Sirko, “Dynamical electrical conductivity of n-type PbTe,” *Phys. Rev. B*, vol. 28, no. 10, p. 6079, Nov. 1983.

[77] T. Uechi, T. Fukui, and N. Sano, “3D Monte Carlo analysis of potential fluctuations under high electron concentrations,” *J. Comput. Electronics*., vol. 7, no. 3, pp. 240–243, Feb. 2008.

[78] T. Uechi, T. Fukui, and N. Sano, “3D Monte Carlo simulation including full Coulomb interaction under high electron concentration regimes,” *Physica Status Solidi (c)*, vol. 5, no. 1, pp. 102–106, Jan. 2008.

[79] K. Throngnumchai, K. Asada, and T. Sugano, “Modeling of 0.1-μm MOSFET on SOI structure using Monte Carlo simulation technique,” *IEEE Trans. Electron Devices*, vol. 33, no. 7, pp. 1005–1011, Jul. 1986.

[80] G. D. Mahan, “Energy gap in Si and Ge: Impurity dependence,” *J. Appl. Phys*., vol. 51, no. 5, p. 2634, Jul. 1980.

[81] C. Kittel, Quantum theory of solids. New York: John Wiley and Sons, 1963.

[82] W. Shockley, “Problems related to p-n junctions in silicon,” Czechoslov. *J. Phys*., vol. 11, no. 2, pp. 81–121, Feb. 1961.

[83] J. Fang, M. Reaz, S. L. Weeden-Wright, R. D. Schrimpf, R. A. Reed, R. A. Weller, M. V. Fischetti, and S. T. Pantelides, “Understanding the average electron–hole pair-creation energy in Si and Ge based on full-band Monte Carlo simulations,” *IEEE Trans. Nucl. Sci.*, vol. 66, no. 1, pp. 444–451, Jan. 2019.

[84] A. Rothwarf, “Plasmon theory of electron-hole pair production: Efficiency of cathode ray phosphors,” *J. Appl. Phys.*, vol. 44, no. 2, pp. 752–756, Feb. 1973.

[85] P. E. Dodd and L. W. Massengill, “Basic mechanisms and modeling of single-event upset in digital microelectronics,” *IEEE Trans. Nucl. Sci.*, vol. 50, no. 3, pp. 583–602, Jun. 2003.

[86] S. E. Laux and M. V Fischetti, “Monte-Carlo simulation of submicrometer Si n-MOSFET’s at 77 and 300 K,” *IEEE Electron Device Lett.*, vol. 9, no. 9, pp. 467–469, Sep. 1988.

[87] R. C. Alig, S. Bloom, and C. W. Struck, “Scattering by ionization and phonon emission in semiconductors,” *Phys. Rev. B*, vol. 22, no. 12, pp. 5565–5582, Dec. 1980.

[88] C. A. Klein, “Bandgap dependence and related features of radiation ionization energies in semiconductors,” *J. Appl. Phys.*, vol. 39, no. 4, pp. 2029–2038, Mar. 1968.

[89] F. Scholze, H. Rabus, and G. Ulm, “Mean energy required to produce an electron-hole pair in silicon for photons of energies between 50 and 1500 eV,” *J. Appl. Phys.*, vol. 84, no. 5, pp. 2926–2939, Sep. 1998.

[90] B. G. Lowe and R. A. Sareen, “A measurement of the electron-hole pair creation energy and the Fano factor in silicon for 5.9 keV X-rays and their temperature dependence in the range 80-270 K,” *Nucl. Instrum. Methods Phys. Res. A*, vol. 576, no. 2–3, pp. 367–370, Jun. 2007.

[91] S. O. W. Antman, D. A. Landis, and R. H. Pehl, “Measurements of the Fano factor and the energy per hole-electron pair in germanium,” *Nucl. Instruments Methods*, vol. 40, no. 2, pp. 272–276, Mar. 1966.

[92] C. A. Klein, “Simple explanation of the electron-hole pair creation energy puzzle in germanium,” *Phys. Lett. A*, vol. 24, no. 10, pp. 513–514, May 1967.

[93] R. A. Reed, R. A. Weller, A. Akkerman, J. Barak, W. Culpepper, S. Duzellier, C. Foster, M. Gaillardin, G. Hubert, T. Jordan, I. Jun, S. Koontz, F. Lei, P. McNulty, M. H. Mendenhall, M. Murat, P. Nieminen, P. O'Neill, M. Raine, B. Reddell, F. Saigné, G. Santin, L. Sihver, H. H. K. Tang, P. R. Truscott, and F. Wrobel, “Anthology of the development of radiation transport tools as applied to single event effects,” *IEEE Trans. Nucl. Sci.*, vol. 60, no. 3, pp. 1876–1911, 2013.

[94] A. Akkerman, M. Murat, and J. Barak, “Monte Carlo calculations of electron transport in silicon and related effects for energies of 0.02-200 keV,” *J. Appl. Phys.*, vol. 106, no. 11, pp. 113703,

Dec. 2009.
[95] A. Valentin, M. Raine, M. Gaillardin, and P. Paillet, “Geant4 physics processes for microdosimetry simulation: Very low energy electromagnetic models for protons and heavy ions in silicon,” *Nucl. Instrum. Methods Phys. Res. B*, vol. 287, pp. 124–129, Sep. 2012.
[96] G. Gilat and L. J. Raubenheimer, “Accurate numerical method for calculating frequency-distribution functions in solids,” *Phys. Rev.*, vol. 144, no. 2, pp. 390–395, Apr. 1966.
[97] E. Cartier, M. V. Fischetti, E. A. Eklund, and F. R. McFeely, “Impact ionization in silicon,” *Appl. Phys. Lett.*, vol. 62, no. 25, pp. 3339–3341, Jun. 1993.
[98] M. V. Fischetti, N. Sano, S. E. Laux, and K. Natori, “Full-band-structure theory of high-field transport and impact ionization of electrons and holes in Ge, Si, and GaAs,” *J. Technol. Comput. Aided Des. (TCAD)*, pp. 1–50, Dec. 2013.
[99] E. O. Kane, “Electron scattering by pair production in silicon,” *Phys. Rev.*, vol. 159, no. 3, pp. 624–631, Jul. 1967.
[100] A. Phillips and P. J. Price, “Monte Carlo calculations on hot electron energy tails,” *Appl. Phys. Lett.*, vol. 30, no. 10, pp. 528–530, May 1977.
[101] R. A. Sinton, “Efforts to develop excess-carrier recombination lifetime measurement standards for silicon PV,” *IEEE Photovoltaic Specialists Conf.*, pp. 837–841, Jun. 2010
[102] R. Häcker and A. Hangleiter, “Intrinsic upper limits of the carrier lifetime in silicon,” *J. Appl. Phys.*, vol. 75, no. 11, pp. 7570–7572, Jun. 1994.
[103] M. N. Mazziotta, “Electron-hole pair creation energy and Fano factor temperature dependence in silicon,” *Nucl. Instrum. Methods Phys. Res. Sect. A*, vol. 584, no. 2–3, pp. 436–439, Jan. 2008.
[104] M. V. Fischetti, P. D. Yoder, M. M. Khatami, G. Gaddemane, and M. L. Van De Put, “‘Hot electrons in Si lose energy mostly to optical phonons’: Truth or myth?,” *Appl. Phys. Lett.*, vol. 114, no. 22, pp. 222104, Jun. 2019.
[105] C. L. Anderson and C. R. Crowell, “Threshold energies for electron-hole pair production by impact ionization in semiconductors,” *Phys. Rev. B*, vol. 5, no. 6, pp. 2267–2272, Mar. 1972.
[106] D. M. Fleetwood, “Radiation effects in a post-Moore world,” *IEEE Trans. Nucl. Sci.*, vol. 68, no. 5, pp. 509-545, May 2021.
[107] D. Yakimets, G. Eneman, P. Schuddinck, T. H. Bao, M. G. Bardon, P. Raghavan, A. Veloso, N. Collaert, A. Mercha, D. Verkest, A. V. -Y. Thean, and K. De Meyer, “Vertical GAAFETs for the ultimate CMOS scaling,” *IEEE Trans. Electron Devices*, vol. 62, no. 5, pp. 1433-1437, May 2015.
[108] S. Kim, J. Kim, D. Jang, R. Ritzenthaler, B. Parvais, J. Mitard, H. Mertens, T. Chiarella, N. Horiguchi, and J. W. Lee, “Comparison of temperature dependent carrier transport in FinFET and gate-all-around nanowire FET,” *Applied Sciences*, vol. 10, no. 8, p. 2979, Apr. 2020.
[109] V. Moroz, J. Huang, and R. Arghavani, “Transistor design for 5nm and beyond: Slowing down electrons to speed up transistors,” *Proc. 17th Int. Symp. Qual. Electron. Design (ISQED)*, pp. 278-283, May 2016.
[110] G. D. Barmparis, Y. S. Puzyrev, X. G. Zhang, and S. T. Pantelides, “Theory of inelastic multiphonon scattering and carrier capture by defects in semiconductors: Application to capture cross sections,” *Phys. Rev. B*, vol. 92, no. 21, article no. 214111, Dec. 2015.
[111] D. J. DiMaria, “Dependence on gate work function of oxide charging, defect generation, and hole currents in MOS structures,” *J. Appl. Phys*., vol. 81, no. 7, pp. 3220-3226, Apr. 1997.
[112] W. McMahon, A. Haggag, and K. Hess, “Reliability scaling issues for nanoscale devices,” *IEEE Trans. Nanotech.*, vol. 2, no. 1, pp. 33–38, Mar. 2003.
[113] J. Fang, M. V. Fischetti, R. D. Schrimpf, R. A. Reed, E. Bellotti, and S. T. Pantelides, “Electron transport properties of AlxGa1−xN/GaN transistors based on first-principles calculations and Boltzmann-equation Monte Carlo simulations,” *Phys. Rev. Appl.*, vol. 11, no. 4, pp. 4445-4458,

Apr. 2019.

[114] C. Guerin, V. Huard, and A. Bravaix, "C. Guerin, V. Huard, and A. Bravaix, "General framework about defect creation at the Si-SiO2 interface," *J. Appl. Phys.*, vol. 105, no. 11, article no. 114513, Jun. 2009.

[115] R. Arora, Z. E. Fleetwood, E. X. Zhang, N. E. Lourenco, J. D. Cressler, D. M. Fleetwood, R. D. Schrimpf, A. K. Sutton, G. Freeman, and B. Greene., "Impact of technology scaling in sub-100 nm nMOSFETs on total-dose radiation response and hot-carrier reliability," *IEEE Trans. Nucl. Sci.*, vol. 61, no. 3, pp. 1426–1432, Jun. 2014.

[116] C. Guérin, V. Huard, and A. Bravaix, "The energy-driven hot-carrier degradation modes of nMOSFETs," *IEEE Trans. Dev. Mater. Reliab.*, vol. 7, no. 2, pp. 225–234, Jun. 2007.

[117] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C. . -H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neirynck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, and K. Zawadzki, "A 45nm logic technology with high-k+ metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging," *Tech. Digest – Int. Electron Devices Meet. (IEDM)*, pp. 247–250, Dec. 2007.

[118] D. J. DiMaria, E. Cartier, and D. Arnold, "Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon," *J. Appl. Phys.*, vol. 73, no. 7, pp. 3367–3384, Apr. 1993.

[119] D. J. DiMaria, E. Cartier, and D. A. Buchanan, "Anode hole injection and trapping in silicon dioxide," *J. Appl. Phys*., vol. 80, no. 1, pp. 304–317, Jul. 1996.

[120] D. J. DiMaria and J. W. Stasiak, "Trap creation in silicon dioxide produced by hot electrons," *J. Appl. Phys*., vol. 65, no. 6, pp. 2342–2356, Mar. 1989.

[121] D. M. Fleetwood, "Border traps and bias-temperature instabilities in MOS devices," *Microelectron. Reliab.*, vol. 80, pp. 266–277, Jan. 2018.

[122] G. C. Ribes, P. Mora, and D. Roy, "Multi-vibrational hydrogen release: a dielectric breakdown model for ultra-thin SiO2/SiON and high-k stack," *ECS Trans.*, vol. 19, no. 2, pp. 195–225, Dec. 2019.

[123] A. Khakifirooz and D. A. Antoniadis, "MOSFET performance scaling - Part II: Future directions," *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1401–1408, Jun. 2008.

[124] D. Mamaluy and X. Gao, "The fundamental downscaling limit of field effect transistors", *Appl. Phys. Lett.*, vol. 106, no. 19, pp. 193503, May 2015.

[125] L. Wei, O. Mysore, and D. Antoniadis, "Virtual-source-based self-consistent current and charge FET models: From ballistic to drift-diffusion velocity-saturation operation," *IEEE Trans. Electron Devices*, vol. 59, no. 5, pp. 1263–1271, May 2012.

[126] E. Pop, "Energy dissipation and transport in nanoscale devices," *Nano Research*, vol. 3, no. 3, pp. 147-169, Mar. 2010.

[127] K. Banoo and M. S. Lundstrom, "Electron transport in a model Si transistor," *Solid. State. Electron.*, vol. 44, no. 9, pp. 1689–1695, Sep. 2000.

[128] M. S. Lundstrom and J. Guo, Nanoscale transistors: Device physics, modeling and simulation. 129 US, 2006.

[129] D. A. Antoniadis, I. Aberg, C. Ní Chléirigh, O. M. Nayfeh, A. Khakifirooz, and J. L. Hoyt, "Continuous MOSFET performance increase with device scaling: The role of strain and channel material innovations," *IBM J. Res. Dev.*, vol. 50, no. 4–5, pp. 363–376, Jul. 2006.

[130] V. Barral, T. Poiroux, J. Saint-Martin, D. Munteanu, J. L. Autran, and S. Deleonibus, “Experimental investigation on the quasi-ballistic transport: Part I - Determination of a new backscattering coefficient extraction methodology,” *IEEE Trans. Electron Devices*, vol. 56, no. 3, pp. 408–419, Mar. 2009.

[131] V. Barral, T. Poiroux, D. Munteanu, J. L. Autran, and S. Deleonibus, “Experimental investigation on the quasi-ballistic transport: Part II - Backscattering coefficient extraction and link with the mobility,” *IEEE Trans. Electron Devices*, vol. 56, no. 3, pp. 420–430, Mar. 2009.

[132] A. Cros, K. Romanjek, D. Fleury, S. Harrison, R. Cerutti, P. Coronel, B. Dumont, A. Pouydebasque, R. Wacquez, B. Duriez, R. Gwoziecki, F. Boeuf, H. Brut, G. Ghibaudo, and T. Skotnicki, “Unexpected mobility degradation for very short devices: A new challenge for CMOS scaling,” *Tech. Digest – Int. Electron Device Meet. (IEDM)*, Dec. 2006.

[133] M. V Fischetti and S. E. Laux, “Monte Carlo study of electron transport in silicon inversion layers,” *Phys. Rev. B*, vol. 48, no. 4, pp. 2244–2274, Jul. 1993.

[134] A. P. Kirk and M. V Fischetti, “Fundamental limitations of hot-carrier solar cells,” *Phys. Rev. B*, vol. 86, no. 16, p. 165206, Oct. 2012.

[135] M. V. Fischetti, “Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures, Part I: Homogeneous transport,” *IEEE Trans. Electron Devices*, vol. 38, no. 3, pp. 634–649, Mar. 1991.

[136] J. Y. Tang and K. Hess, “Impact ionization of electrons in silicon (steady state),” *J. Appl. Phys*., vol. 54, no. 9, pp. 5139–5144, Sep. 1983.

[137] S. E. Laux and M. V. Fischetti, “Numerical Aspects and Implementation of theDamoclesMonte Carlo Device Simulation Program,” Monte Carlo Device Simulation, Springer US, 1991, pp. 1–26.

[138] E. Conwell and V. F. Weisskopf, “Theory of impurity scattering in semiconductors,” *Phys. Rev.*, vol. 77, no. 3, pp. 388–390, Feb. 1950.

[139] J. Y. Tang and K. Hess, “Theory of hot electron emission from silicon into silicon dioxide,” *J. Appl. Phys*., vol. 54, no. 9, pp. 5145–5151, Jun. 1983.

[140] S. Tyaginov, Å. M. El-Sayed, A. Makarov, A. Chasin, H. Arimura, M. Vandemaele, M. Jech, E. Capogreco, L. Witters, A. Grill, A. De Keersgieter, G. Eneman, D. Linten, and B. Kaczer, “Understanding and physical modeling superior hot-carrier reliability of Ge pNWFETs,” 2019 *IEEE Intl. Electron Devices Meet.*, San Francisco, CA, USA, pp. 21.3.1-21.3.4, Dec. 2019.

[141] A. Lacaita, “Why the effective temperature of the hot electron tail approaches the lattice temperature,” *Appl. Phys. Lett.*, vol. 59, no. 13, pp. 1623–1625, Sep. 1991.

[142] S. Tyaginov, T. U. Wien, I. A. Starkov, H. Enichlmair, J. M. Park, C. Jungemann, and T. Grasser, “Physics-based hot-carrier degradation models,” *ECS Trans.*, vol. 35, no. 4, pp. 321-352, Apr. 2011.

[143] G. D. Mahan, “Hot electrons in one dimension,” *J. Appl. Phys.*, vol. 58, no. 6, pp. 2242–2251, Sep. 1985.

[144] J. Franco, B. Kaczer, A. Chasin, E. Bury, and D. Linten, “Hot electron and hot hole induced degradation of SiGe p-FinFETs studied by degradation maps in the entire bias space,” *Proc. IEEE Intl. Rel. Phys. Symp. (IRPS)*, pp. 5A.11-5A.17, Mar. 2018.

[145] V. H. Chan and J. E. Chung, “Two-stage hot-carrier degradation and its impact on submicrometer LDD NMOSFET lifetime prediction,” *IEEE Trans. Electron Devices*, vol. 42, no. 5, pp. 957–962, May 1995.

[146] M. Cho, P. Roussel, B. Kaczer, R. Degraeve, J. Franco, M. Aoulaiche, T. Chiarella, T. Kauerauf, N. Horiguchi, and G. Groeseneken, “Channel hot carrier degradation mechanism in long/short channel n-FinFETs,” *IEEE Trans. Electron Devices*, vol. 60, no. 12, pp. 4002–4007, Dec. 2013.

[147] A. Bravaix,V. Huard, F. Cacho, X. Federspiel, and D. Roy, “Hot-carrier degradation in decananometer CMOS nodes: From an energy-driven to a unified current degradation modeling by a multiple-carrier degradation process”, Hot Carrier Degradation in Semiconductor Devices. New York: Springer, ch. 2, pp. 57-103, 2015.

[148] S. Mukherjee, Y. Puzyrev, J. Chen, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, “Hot-carrier degradation in GaN HEMTs due to substitutional iron and its complexes,” *IEEE Trans. Electron Devices*, vol. 63, no. 4, pp. 1486–1494, Apr. 2016.

[149] J. H. Stathis, “Dissociation kinetics of hydrogen-passivated (100) Si/SiO2 interface defects,” *J. Appl. Phys*, vol. 77, no. 12, pp. 6205-6207, Jun. 1995.

[150] H. Mertens et al., “(Invited) Gate-All-Around Transistors Based on Vertically Stacked Si Nanowires,” *ECS Trans.*, vol. 77, no. 5, pp. 19-30, Apr. 2017.

[151] K. Bergman, M. Stavola, S. J. Pearton, and J. Lopata, “Donor-hydrogen complexes in passivated silicon,” *Phys. Rev. B*, vol. 37, no. 5, pp. 2770-2773, Feb. 1988.

[152] P. J. H. Denteneer, C. G. van de Walle, and S. T. Pantelides, “Microscopic nature of the hydrogen-phosphorus complex in crystalline silicon,” *Phys. Rev. B*, vol. 41, no. 6, pp. 3885-3888, Feb. 1990.

[153] D. M. Fleetwood, “Effects of hydrogen transport and reactions on microelectronics radiation response and reliability,” *Microel. Reliab*., vol. 42, no. 4-5, pp. 523-541, Apr-May 2002.

[154] S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, “Defect generation by hydrogen at the Si-SiO_2 interface,” *Phys. Rev. Lett.*, vol. 87, Art. no. 165506, Oct. 2001.

[155] F. B. McLean, “A framework for understanding radiation-induced interface states in SiO_2 MOS structures,” *IEEE Trans. Nucl. Sci.*, vol. 27, no. 6, pp. 1651-1657, Dec. 1980.

[156] E. Cartier, J. H. Stathis, and D. A. Buchanan, “Passivation and depassivation of Si dangling bonds at the Si/SiO_2 interface by atomic hydrogen,” *Appl. Phys. Lett.*, vol. 63, no. 11, pp. 1510-1512, Sep. 1993.

[157] J. Stathis and E. Cartier, “Atomic hydrogen reactions with P_b centers at the (100) Si/SiO_2 interface,” *Phys. Rev. Lett.*, vol. 72, no. 17, pp. 2745-2748, Apr. 1994.

[158] S. T. Pantelides, S. N. Rashkeev, R. Buczko, D. M. Fleetwood, and R. D. Schrimpf, “Reactions of hydrogen with Si/SiO_2 interfaces,” *IEEE Trans. Nucl. Sci.*, vol. 47, no. 6, pp. 2262–2268, Dec. 2000.

[159] L. Tsetseris, D. M. Fleetwood, R. D. Schrimpf, X. J. Zhou, I. G. Batyrev, and S. T. Pantelides, “Hydrogen effects in MOS devices,” *Microelectron. Engrg*., vol. 84, no. 9-10. pp. 2344-2349, Sep.-Oct., 2007.

[160] M. Cho, J. Lee, M. Aoulaiche, B. Kaczer, P. Roussel, T. Kauerauf, R. Degraeve, J. Franco, L. Ragnarsson, and G. Groeseneken, “Insight into *n*/*p*BTI mechanisms in sub-1-nm-EOT devices,” *IEEE Trans. Electron Devices*, vol. 59, no. 8, pp. 2042-2048, Aug. 2012.

[161] S. Tyaginov and T. Grasser, “Modeling of hot-carrier degradation: Physics and controversial issues,” *IEEE International Integrated Reliability Workshop Final Report,* South Lake Tahoe, CA, pp. 206-215, 2012.

[162] D. A. Buchanan, A. D. Marwick, and D. J. DiMaria, “Hot-electron-induced hydrogen redistribution and defect generation in MOS capacitors,” *J. Appl. Phys.*, vol. 76, no. 6, pp. 3595-3608, Sep. 1994.

[163] D. J. DiMaria, “Defect generation in field-effect transistors under channel-hot-electron stress,” *J. Appl. Phys.*, vol. 87, no. 12, pp. 8707-8715, Jun. 2000.

[164] E. X. Wang, P. Matagne, L. Shifren, B. Obradovic, R. Kotlyar, S. Cea, M. Stettler, and M. D. Giles, “Physics of hole transport in strained silicon MOSFET inversion layers,” *IEEE Trans. Electron Devices*, vol. 53, no. 8, pp. 1840–1851, Aug. 2006.

[165] P. B. Vyas, M. L. Van De Putt, and M. V. Fischetti, “Quantum mechanical study of impact of surface roughness on electron transport in ultra-thin body silicon FETs,” *IEEE 13th Nanotech. Materials and Devices Conf. (NMDC)*, Oct. 2018.

[166] S. Jin, M. V. Fischetti, and T. W. Tang, “Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity,” *J. Appl. Phys.*, vol. 102, no. 8, p. 83715, Oct. 2007.

[167] P. Weckx, J. Ryckaert, E. D. Litta, D. Yakimets, P. Matagne, P. Schuddinck, D. Jang, B. Chehab, R. Baert, M. Gupta, Y. Oniki, L. . -A. Ragnarsson, N. Horiguchi, A. Spessot, and D. Verkest, "Novel forksheet device architecture as ultimate logic scaling device towards 2nm," *IEEE Int. Electron Devices Meet. (IEDM)*, pp. 36.5.1-36.5.4, Dec. 2019.

[168] R. Kippenhahn and A. Weigert, “The Degenerate Electron Gas,” *Astronomy and Astrophysics Library*, Springer, Berlin, Heidelberg., pp. 118–128, 1990. https://doi.org/10.1007/978-3-642-61523-8_15.

[169] S. Jin, M. V. Fischetti, and T. W. Tang, “Surface roughness scattering in ultrathin-body SOI MOSFETs,” *Int. Conf. Simulation of Semiconductor Processes and Devices (SISPAD)* pp. 61–64, 2007.

[170] J. Hattori, S. Uno, N. Mori, and K. Nakazato, “Reduction of acoustic phonon limited electron mobility due to phonon confinement in silicon nanowire MOSFETs,” in *65th DRC Device Research Conference*, 2007, pp. 177–178.

[171] M. Fischetti and S. E. Laux, “DAMOCLES UserGuide” IBM Corporation, Yorktown Heights, NY, Dec. 2004.