Numerical Analysis of Self-Heating in Silicon Devices

Author

Faraz Kaiser Malik

00000276894

Supervisor

Dr. Tariq Talha

DEPARTMENT OF MECHANICAL ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

OCTOBER, 2020

Numerical Analysis of Self-Heating in Silicon Devices

Author

Faraz Kaiser Malik

00000276894

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Mechanical Engineering

____

Thesis Supervisor

Dr. Tariq Talha

DEPARTMENT OF MECHANICAL ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

OCTOBER, 2020

Declaration

I certify that this research work titled *"Numerical Analysis of Self-Heating in Silicon Devices"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged and cited.

Faraz Kaiser Malik

00000276894

MS-2018 Mechanical Engineering

College of Electrical & Mechanical Engineering

National University of Sciences and Technology, Islamabad

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical, and spelling mistakes. The thesis is also written according to the format provided by the university.

Faraz Kaiser Malik
00000276894
MS-2018 Mechanical Engineering
College of Electrical & Mechanical Engineering
National University of Sciences and Technology, Islamabad

Dr. Tariq Talha
Assistant Professor
Department of Mechanical Engineering
College of Electrical & Mechanical Engineering
National University of Sciences and Technology, Islamabad

Copyright Statement

Copyright © 2020 by Faraz Kaiser Malik

All rights reserved. Reproduction, distribution, or transmission of this thesis, in whole or in part in any form or by any means requires the prior written permission of the author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the author and lodged in the Library of NUST College of E&ME, details of which may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) may not be made without the permission (in writing) of the author.

The ownership of any intellectual property rights which may be described in this thesis is vested in NUST College of E&ME, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the College of E&ME, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosure and exploitation may take place is available from the Library of NUST College of E&ME, Rawalpindi.

Acknowledgements

I am immensely grateful to my thesis supervisor, Dr. Tariq Talha, for his support, continuous encouragement and constructive criticism that has enabled me to successfully complete my thesis. I am also thankful to him for his guidance and advice regarding research and career prospects during and after my MS studies.

I would also like to express my heartfelt gratitude to Dr. Faisal Ahmed for providing invaluable guidance and feedback during this work, and for being on my thesis guidance and evaluation committee. Thanks also go out to Dr. Bilal Anjum and Dr. Hasan Aftab Saeed for being a part of this committee.

I would like to extend my gratitude to my colleague, Mr. Saleem Akhter, for several enlightening discussions. The support of the staff at the Supercomputing Research and Education Center in the Research Center for Modeling and Simulation, National University of Sciences and Technology, is also gratefully acknowledged, as is the kind permission for the use of the facilities at this center.

Dedicated to my beloved parents for their invaluable support and cooperation.

Abstract

The electronics industry has pursued the aggressive miniaturization of solid-state devices to meet future technological demands, which has resulted in a significant increase in the chip transistor count and power density as device dimensions have been downscaled into the nanoscale regime. Nanoscale materials exhibit notably different characteristics from their bulk counterparts due to quantum confinement effects and changes in the mechanism of transport of the energy carriers at shorter length scales. The resistive heating of electronic devices during their operation is the single biggest hindrance to the efficient, safe, and reliable operation of these devices, since the thermal and electrical conductivities of most materials are adversely affected by a reduction in spatial dimensions and an increase in lattice temperatures, which creates a positive feedback loop that can lead to thermal runaway and breakdown as device dimensions are reduced. The study of the electrical and thermal properties of ultra-short nanoscale devices that operate in the ballistic transport regime, and the variation of these properties with device geometry and operating conditions, is therefore necessary to allow for the development of thermally-aware device designs and cooling methods in the future. This work uses the numerical Monte Carlo approach to simulate the carrier transport through nanoscale bulk silicon and one-dimensional silicon devices that serve as simplified models of nanoscale transistors. The simulations are performed over a range of operating conditions (applied voltage and initial temperature) and for different geometric specifications (channel length, electrode length, electrode doping) that are likely to be encountered in future device applications. The impact of changes in the carrier transport mechanism that are brought about by a reduction in the dimensions of characteristic device features on the device performance and the generation of localized thermal hotspots is thereby investigated systematically.

Key Words: *Joule Heating, Nanoscale Devices, Nanoscale Silicon, Monte Carlo Method, Solid-state Electronics*

Table of Contents

Declaration i
Language Correctness Certificate ii
Copyright Statement iii
Acknowledgements iv
Abstract vi
Table of Contents vii
List of Figures ix

CHAPTER 1: INTRODUCTION 1
1.1 Background 1
1.2 Electrical Conductivity and Material Classifications 3
1.3 Semiconductors in Electronics 5
1.4 Heat Conduction in Semiconductors 6
1.4.1 Phonon Modes 7
1.4.2 Phonon Scattering Processes 8
1.4.3 Energy Dissipation and Transport in Nanoscale Devices 9
1.5 Nanoscale Heat Conduction Models in Literature 11
1.5.1 Experimental Methods 11
1.5.2 Analytical Methods 12
1.6 Nanoscale Heat Generation Models in Literature 15
1.6.1 Drift-Diffusion Method 16
1.6.2 Hydrodynamic Method 17
1.6.3 Monte Carlo Simulations 18
1.7 Research Objectives 19
1.8 Thesis Outline 20

CHAPTER 2: NUMERICAL METHODOLOGY 22
2.1 General Monte Carlo Algorithm 22
2.2 Monte Carlo Models in Literature 25
2.3 Details of the Numerical Model 29
2.3.1 Electron Band Model 29
2.3.2 Scattering Mechanisms 30

2.3.3 Self-Consistent Poisson Equation Solution....32
2.3.4 Boundary Conditions....33
2.4 Summary....34

CHAPTER 3: CARRIER TRANSPORT IN BULK SILICON....35
3.1 Introduction....35
3.2 Variation of the Electron Drift Velocity....36
3.3 Variation of the Electron Drift Mobility....39
3.4 Phonon Emission Rates....41
3.5 Variation of the Volumetric Heat Generation Rate....43
3.6 Summary....46

CHAPTER 4: JOULE HEATING IN BALLISTIC SILICON DEVICES....47
4.1 Introduction....47
4.2 Preliminary Setup....48
4.3 Time Step and Grid Spacing Sensitivity....52
4.4 Validation....54
4.5 Effect of Voltage....55
4.6 Effect of Channel Length....59
4.7 Effect of Electrode Lengths....63
4.8 Effect of Temperature....64
4.9 Effect of Electrode Doping Density....68
4.10 Two-Dimensional Simulations....70
4.11 Summary....73

CHAPTER 5: CONCLUSION....75
5.1 Summary....75
5.2 Future Work....77

REFERENCES....79

List of Figures

Figure 1.1: Semi-log plot of 46 years of microprocessor transistor count data. Data collected by F. Labonte and K. Rupp [8]. 2

Figure 1.2: Semi-log plot of improvements in computations per kilowatt-hour from 1946 to 2009. Data from Koomey et al. [9]. 2

Figure 1.3: Energy band diagram for crystalline solid classifications. 4

Figure 1.4: Illustration of scattering mechanisms. 9

Figure 1.5: Illustration of transport mode classifications based on device size relative to mean free path. 10

Figure 2.1: Flowchart of the ensemble Monte Carlo method used in this study. 23

Figure 3.1: Validation of the variation of the electron drift velocity in bulk silicon with the applied electric field strength, lattice temperature and doping density. Solid lines are the results of the present work, triangles represent the experimental results of Canali et al. [42], and circles represent the MC results of Pop et al. [110], [111]. 37

Figure 3.2: Variation of the electron drift velocity in intrinsic bulk silicon with the lattice temperature and applied electric field strength. 38

Figure 3.3: Variation of the electron drift velocity in bulk silicon at 200 K with the doping density and applied electric field strength. 38

Figure 3.4: Validation of the variation of the electron drift mobility in intrinsic bulk silicon with the lattice temperature at a low electric field strength of 500 V cm^{-1}. 40

Figure 3.5: Variation of the electron drift mobility in bulk silicon with the doping density and lattice temperature at an applied electric field strength of 500 V cm^{-1} (low-field regime). 41

Figure 3.6: Net phonon generation spectrum for 10^{17} cm^{-3} doped bulk silicon at 300 K and an applied electric field of (a) 5 kV cm^{-1} and (b) 50 kV cm^{-1}. Solid lines represent results of the present work, while dots represent the results of Pop et al. [111] for the corresponding cases. 42

Figure 3.7: Validation of the variation with the applied electric field strength of the rate of energy dissipation through each phonon mode in 10^{17} cm^{-3} doped bulk silicon at 300 K. Solid lines are the results of the present work, while circles represent the results of Pop et al. [111]. 44

Figure 3.8: Effect of the doping density and applied electric field strength on the volumetric energy dissipation rate in bulk silicon at 300 K. 45

Figure 3.9: Effect of the lattice temperature and applied electric field strength on the volumetric energy dissipation rate in 10^{17} cm^{-3} doped bulk silicon. 45

Figure 4.1: Schematic diagram of the n^+nn^+ device investigated in this study.....................47

Figure 4.2: Snapshot of the geometry of a 10 nm channel 1D n^+nn^+ device created using the COMSOL Multiphysics software. The *x*-axis length scale is in nanometers. 49

Figure 4.3: Comparisons of the distributions of characteristic parameters over a section of the length of a 10 nm channel n^+nn^+ device, as determined through a MC simulation, with the preliminary drift-diffusion simulation (COMSOL Multiphysics) results for the corresponding case. The device, with 100 nm long electrodes doped to 10^{20} cm^{-3}, is subjected to an applied potential difference of 0.8 V and an initial uniform temperature of 300 K. 50

Figure 4.4: Variation of the maximum volumetric heat generation rate in a 10 nm channel n^+nn^+ device at 300 K with the simulation time step size. 53

Figure 4.5: Variation of the maximum volumetric heat generation rate in a 10 nm channel n^+nn^+ device at 300 K with the grid spacing. 53

Figure 4.6: Validation of the volumetric heat generation rates at 300 K over the lengths of (a) 20 nm and (b) 500 nm channel n^+nn^+ silicon devices. Solid lines are the results of the present work, while dotted lines are the MC results obtained from Pop et al. [73]. 55

Figure 4.7: Effect of voltage on the distributions of (a) heat generation, (b) electron energy and (c) electron velocity over the length of a 10 nm channel n^+nn^+ silicon device at 300 K. 57

Figure 4.8: Variation of the spatial distributions of the (a) heat generation rate, (b) electron energy, and (c) electron velocity with the channel length of the n^+nn^+ silicon devices. 60

Figure 4.9: Variation of the device performance parameter maxima in an n^+nn^+ device with the channel length at 300 K. In (a), the separate contributions of the optical and acoustic phonon modes to the total heat generation rate are shown. 61

Figure 4.10: Effect of channel length on the electric current flowing through the n^+nn^+ device. 62

Figure 4.11: Effect of electrode length on the (a) heat generation rate and (b) electron energy over the length of a 20 nm channel device at 300 K (legends identical). 63

Figure 4.12: Variation of the device performance parameter distributions over a section of the total device length with the initial temperature. 65

Figure 4.13: Effect of initial temperature on the maximum (a) volumetric heat generation rate, (b) electron energy, (c) electron drift velocity and (d) electric field strength in a 10 nm channel n^+nn^+ device. 66

Figure 4.14: Effect of temperature on the (a) minimum electron density in a 10 nm channel n^+nn^+ device and (b) density of the electric current flowing through the device. 67

Figure 4.15: Effect of electrode doping on the maximum (a) volumetric heat generation rate, (b) electron energy, (c) electric field strength and (d) electron drift velocity in a 10 nm channel n^+nn^+ device. 69

Figure 4.16: Effect of electrode doping on the (a) minimum electron density in a 10 nm channel n^+nn^+ device and (b) density of the electric current flowing through the device. 70

Figure 4.17: Steady-state distribution of electrons in the two-dimensional transistor geometry. The *x*- and *y*-axes represent the horizontal and vertical distance, respectively, from the bottom left corner of the transistor in nanometer units. 71

Figure 4.18: Steady-state distribution of the *x*-component of the electric field strength over the two-dimensional transistor geometry. The *x*- and *y*-axes represent the horizontal and vertical distance, respectively, from the bottom left corner of the transistor in nanometer units. 72

Figure 4.19: Steady-state distribution of the electric field strength over the length of a 1D n^+nn^+ device subjected to an applied potential difference of 0.8 V with a 20 nm long channel and 50 nm long electrodes doped to 10^{20} cm^{-3}, as determined through a Monte Carlo simulation. 72

CHAPTER 1: INTRODUCTION

In a world increasingly reliant upon microprocessor systems, the efficient, predictable, and reliable performance of electronic devices constituting the circuitry has become critically important. A significant amount of effort has been put into reducing the size of such devices over the past few decades, as well as into designing and implementing appropriate cooling systems that ensure safe operating conditions are maintained. However, with characteristic feature sizes now well into the nanoscale, the roles of localized impurities and size fluctuations have become important, making conventional manufacturing and cooling approaches inadequate. Due to the changes in device behavior and the spatial constraints that accompany downscaling, identification of the location and intensity of potential thermal hotspots and knowledge of their variation with device parameters is an essential part of system design.

1.1 Background

The process by which the passage of electric current through a conductor produces heat due to the natural resistivity of the material is called Joule heating, also known as resistive or Ohmic heating. The inelastic collisions of the charge carriers, set in motion under the influence of an electric field, with the material lattice result in the conversion of electrical energy to thermal energy [1], [2]. While this phenomenon is favorable for many applications where the generation of heat is desired—including, but not limited to, electric irons, heaters, kettles, and toasters—it presents a significant hindrance to the safe and efficient performance of electronic devices. In microprocessors on computer motherboards and control systems, maintaining low operating temperatures (below ~ 85 °C) is vital to ensuring longevity, stable operation and prevention of electrical and thermal hazards [3], [4].

The study of Joule heating gains greater significance when the recent developments in micro and nanoscale electronic technologies are considered, since the trend of aggressive miniaturization pursued by the industry has exacerbated the cooling problem. Miniaturization has progressed at an exponential scale guided by Moore’s law, named after Intel co-founder Gordon Moore, who predicted in 1965 that the number of transistors on a dense integrated circuit would double every year [5], a prediction revised in 1975 to a doubling of chip density every two years [6]. The 10 nm chip lithography node has been realized, with the industry now setting its sights on 7 nm and 5 nm fabrication technologies [7]. Figure 1.1 shows how the

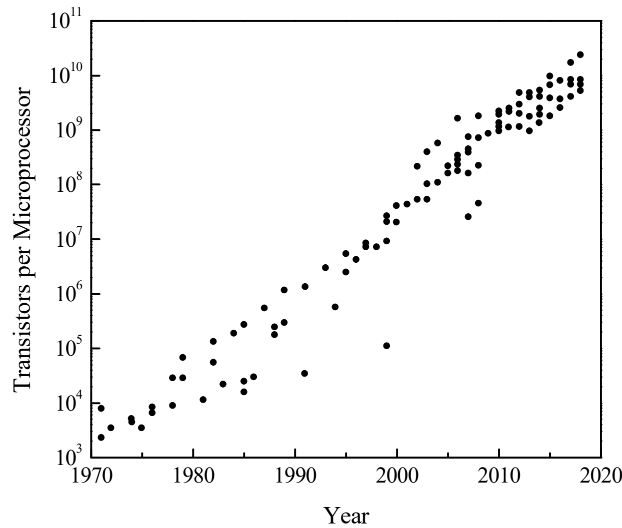

Figure 1.1: Semi-log plot of 46 years of microprocessor transistor count data. Data collected by F. Labonte and K. Rupp [8].

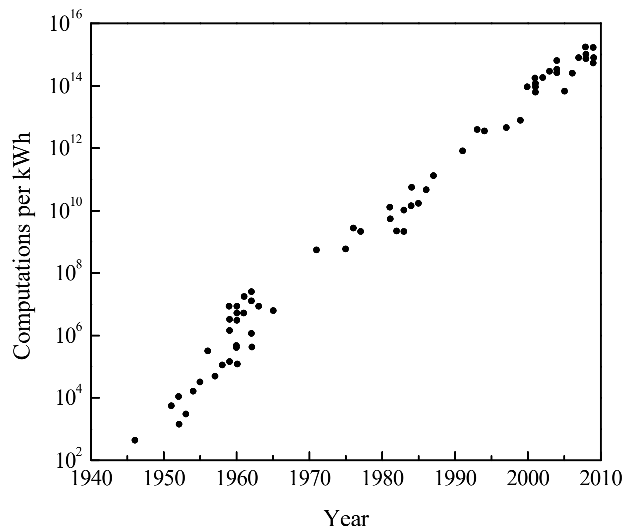

Figure 1.2: Semi-log plot of improvements in computations per kilowatt-hour from 1946 to 2009. Data from Koomey et al. [9].

transistor count per chip has increased over the past half century with the introduction of new processors by the industry's leading companies.

Dennard scaling, a transistor power law based on work published in 1974, postulates that even as transistors get smaller, their power density remains constant. A microprocessor of the same geometric size with a higher transistor count would therefore still consume the same power as one with a lower count of larger transistors [10]. This was made possible by a reduction in the voltage required to maintain a constant electric field across smaller transistors, in turn reducing the power requirements. As shown in Figure 1.2, together with Moore's law,

this allowed for a doubling of performance per watt every 18 months since the beginning of the computer age—a historical trend known as Koomey's law [9].

Even though the transistor count has continued to increase in recent years, significant changes in device behavior and material properties that accompany the transition from the continuum regime to the nanoscale resulted in the breakdown of Dennard scaling around 2006 [11], meaning that modern electronic components are unable to operate within the same power envelope. Static power losses increase rapidly as a fraction of the overall supplied power with decreasing device dimensions and voltages due to short channel effects and a reduction in the material's thermal conductivity, and are also a function of device temperature, thus being implicitly dependent upon the total power dissipated [12].

Continued miniaturization therefore presents significant challenges in the thermal management and control of electronic devices due to a very high increase in power density for high performance chips. A closer reexamination of the development of technologies since 2000 by Koomey and Naffziger [13] indicated that the performance per watt was now doubling every 31 months due to the end of Dennard scaling and the slowing of Moore's Law [14], [15]. Despite the industry shifting its focus to increasing processor cores instead of clock frequency, the number of active switching elements per unit area has continued to increase. Conventional cooling approaches are increasingly failing to deal with the high cooling demand of emerging electronic devices and the development of more effective cooling technologies requires an improved understanding of the mechanisms governing heat generation and dissipation.

1.2 Electrical Conductivity and Material Classifications

Translational symmetry in geometry is a characteristic feature of all crystalline solids, which in turn mandates a translationally symmetric atomic electrostatic potential distribution. For such a system, the electron energy levels are arranged in bands, which may be conducting or forbidden: electrons can propagate in a conduction band, but cannot be placed in a forbidden band, resulting in gaps in the allowed energy spectrum [16]. Since the flow of current requires electrons to move from one state to another, neither completely empty nor completely full bands can conduct electricity—the latter due to the Pauli exclusion principle, which prohibits two or more electrons from occupying the same quantum state [17].

The highest occupied band in a material is called the valence band, and the next available unfilled higher-energy band is called the conduction band. Generally, all crystalline solids can be classified into one of four categories based on their electrical properties: metals, semimetals,

semiconductors, and insulators (in order of decreasing electrical conductivity). Figure 1.3 is a simplified energy band diagram that presents a visual comparison of these classes. In a metal, the conduction band is only partially filled even at absolute zero temperature [16], meaning that it consists of both electrons and empty states. This allows electrons to move between states under the influence of an electric field. Since the energy levels of the conduction and valence bands overlap, very little energy is required to make metals conduct, and they thus have very high electrical conductivity.

In semiconductors, however, the conduction band is completely empty and the valence band is completely full at absolute zero temperature. At 0 K, the system attains its minimum energy configuration in which electron bands are filled progressively from lower to higher energy bands. Since temperature is a measure of the internal energy of a system, higher temperature configurations are attained by promoting electrons from lower energy states to higher energy states. At room temperature, the energy gap between the valence and conduction bands therefore becomes small enough to allow some measurable population of the conduction band. For semimetals, the energy gap partly vanishes at room temperature so that the conduction and valence bands intersect. In an insulator, by contrast, the forbidden energy gap between the two bands remains very high even at room temperature, so that very little current flows under the influence of an electric field.

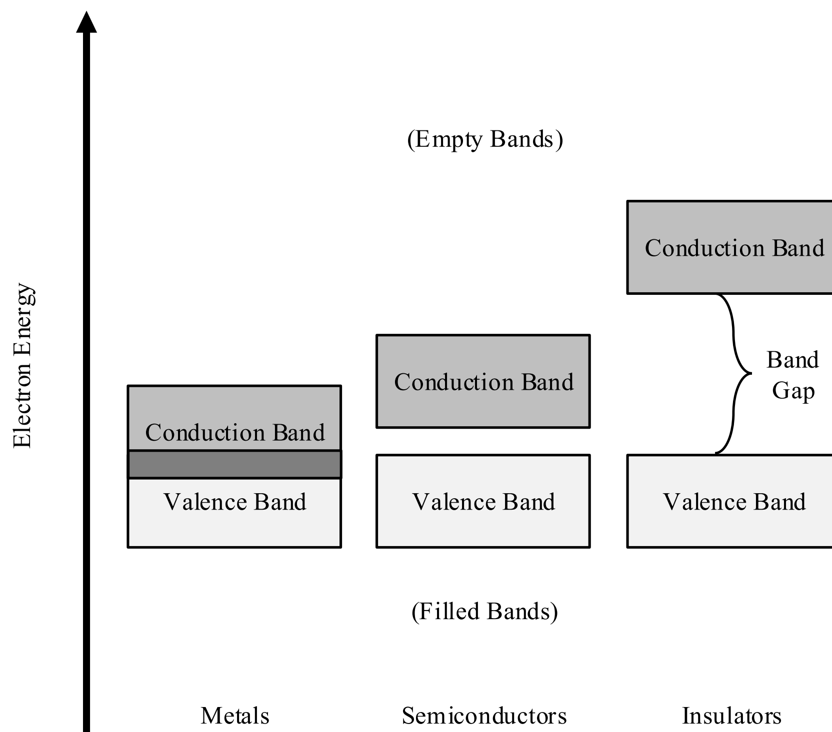

Figure 1.3: Energy band diagram for crystalline solid classifications.

1.3 Semiconductors in Electronics

Semiconductors are indispensable in the fabrication and operation of electronic devices. Unlike conductors and insulators, their small but nevertheless extant band gap can be varied through temperature and, more importantly, impurity level [17], [18], therefore allowing for their electrical and thermal properties to be readily and significantly altered to meet requirements. Their sensitivity to these external variables allows for their use in diodes, amplifiers, and transistors, which power modern electronics.

Elemental semiconductors, such as silicon (Si), germanium (Ge) and selenium (Se), consist of single species of atoms. Binary, ternary, and quaternary compound semiconductors also exist—consisting of two, three or four different elements respectively—the most prominent of which contain elements from the third and fifth groups of the periodic table of elements, such as gallium arsenide (GaAs) and indium phosphide (InP). Silicon remains the most widely used semiconductor in the electronics industry, as it is the second most readily available element on earth through silicate crystals present in the earth's crust. Si crystals have the ability to withstand high temperatures and have a lower cutoff (reverse bias) current than most other elemental semiconductors due to the four valence electrons of each atom forming perfect covalent bonds, although this comes with the disadvantage of a higher potential barrier in the forward bias mode [19].

The promotion of electrons from the valence band to the conduction band—whether by an increase in temperature or photoexcitation—allows a semiconductor to start conducting electricity. In addition, positively charged vacancies, called holes, are left behind in the valence band, which allow that band to also start conducting as it is no longer completely full. Therefore, unlike a metal, in which electric conduction occurs only due to electrons, the electrical conductivity of a semiconductor can be attributed to either electrons or holes, or both.

In a semiconductor with no intentionally added impurities (intrinsic), the hole and electron concentrations at thermal equilibrium are equal. The electrical conductivity of such a semiconductor increases with increasing temperature [18] as more electrons are promoted from the valence band to the conduction band and a greater number of charge carriers are thus made available. An alternate way to change the conductivity by several orders of magnitude at any given temperature is to add impurities to the semiconductor; a process called doping. A doped semiconductor is also called an extrinsic semiconductor, while a degenerate semiconductor is one that is so highly doped as to cause it to behave more like a metal than a semiconductor.

The general relationship for the carrier concentrations in any semiconductor at thermal equilibrium is given by the Law of Mass Action [17]:

$$n_0 \cdot p_0 = n_i^2 \tag{1.1}$$

where n_0 and p_0 are the concentrations of the conducting electrons and holes respectively, and n_i is the intrinsic carrier concentration of the material. Note that for an intrinsic semiconductor at thermal equilibrium: $n_0 = p_0 = n_i$.

A semiconductor can be doped with either electron acceptors (p-type dopants) or electron donors (n-type dopants). Boron, aluminum, and gallium serve as p-type dopants when they replace host atoms in a silicon lattice since they have fewer valence electrons than Si, which results in the creation of a hole in the valence band for each atom replaced. Phosphorous and arsenic, which have more valence electrons than Si, serve as n-type dopants due to the addition of extra, chemically-unbound electrons to the lattice that can be easily ionized into the conduction band. Holes are thus the majority charge carriers in p-type semiconductors, while electrons primarily contribute to electrical conduction in n-type semiconductors.

1.4 Heat Conduction in Semiconductors

In crystalline solids, the migration of free conduction band electrons and interatomic collisions through lattice vibrational waves together contribute to the transport of thermal energy via conduction. Applying the basic quantum mechanics concept of wave-particle duality, according to which all forms of energy and matter may be described as either a particle or a wave [16], each lattice vibration mode is quantized as a phonon, which can be treated as a harmonic oscillator with a characteristic frequency ω and a wave vector $\boldsymbol{q}$, analogous to the use of a photon as a quantization of electromagnetic radiation [18]. Crystal momentum is analogous but not equivalent to linear momentum: no net mass transport occurs in a propagating lattice vibration, so a phonon does not carry physical momentum.

At the continuum scale, Fourier's law is used as the conduction rate equation, which takes the general form [20]:

$$\boldsymbol{Q}'' = -\kappa \nabla T \tag{1.2}$$

where $\boldsymbol{Q}''$ is the heat flux vector, κ is the thermal conductivity of the material (a material property), ∇ is the three dimensional del operator and T is the scalar temperature field. The negative sign represents the flow of heat in the direction of decreasing temperature. The thermal

conductivity can be expressed as the sum of the contributions of the electrons and phonons: $\kappa = \kappa_e + \kappa_p$. For pure metals, which have an abundance of free electrons present in the lattice, $\kappa_e \gg \kappa_p$, while the opposite is true for non-metallic solids (semiconductors and insulators) [21]. Kinetic theory provides the following expression for thermal conductivity [22]:

$$\kappa = \frac{1}{3} C \bar{v} \lambda \tag{1.3}$$

where $C = C_e$ is the electron specific heat per unit volume, $\bar{v}$ is the mean electron velocity and λ is the electron mean free path for conducting materials. For non-conducting materials, $C = C_p$ is the phonon specific heat per unit volume, $\bar{v}$ is the average speed of sound and λ is the phonon mean free path. The mean free path is the average distance that the energy carrier travels before collision with discontinuities in the lattice and other electrons or phonons. A longer mean free path translates to a higher thermal conductivity, since collisions reorient the momentum of the energy carrier and can cause energy dissipation to the lattice.

1.4.1 Phonon Modes

Phonons are categorized based on the lattice vibrational modes of a material, which may be in-plane or out-of-plane. In-plane modes include longitudinal and transverse phonon modes. The former corresponds to atomic displacements along the wave propagation direction (compressive waves), while the latter refers to in-plane displacements perpendicular to the propagation direction (shear waves). In typical three-dimensional solids, transverse modes can have two equivalent polarizations, but in single or few-layered solids, out-of-plane atomic displacements, known as flexural phonons, manifest themselves.

Additionally, the displacements themselves may be in-phase or out-of-phase. Coherent displacements of atoms from their equilibrium lattice positions result in lattice waves typically travelling at the speed of sound, and are hence termed as the acoustic mode. Optical modes, on the other hand, are excited by infrared radiation and involve out-of-phase atomic displacements. They are typically observed in ionic crystals containing multiple atoms as basis in the primitive cell, in which the adjacent anions and cations move in opposite directions when energized by the electric field of the incident radiation. Phonon modes are thus identified by the phase relationship of the atomic displacements in addition to the direction of these displacements relative to the direction of wave propagation.

The relationship between the frequency and wavevector for each vibrational mode, expressed as $\omega = \omega(\boldsymbol{q})$ and called the dispersion relation, allows for the determination of the phonon group velocities. The group velocity is the velocity with which the envelope of the lattice vibration amplitudes is transported by a specific phonon classification, and is the slope of the dispersion relationship for that phonon group [23]. The group velocity is important to thermal conductivity in nonmetallic solids, with faster velocities corresponding to a higher conductivity since thermal energy can be transported across the lattice at a higher rate. In silicon, the group velocity for the acoustic modes ranges from 5000 m s^{-1} (transverse acoustic) to 9000 m s^{-1} (longitudinal acoustic), and is much higher than the velocity for the optical modes, which is on the order of 1000 m s^{-1} [24], [25]. Acoustic phonons are thus the primary contributors to heat conduction in silicon.

The energy of a phonon is calculated as the product of the reduced Planck's constant and the frequency of vibration of the mode. The collective energy E of a mode is [16], [18]:

$$E = \left(n + \frac{1}{2}\right)\hbar\omega \tag{1.4}$$

where n is the number of phonons excited in that mode (also called the phonon occupation number) and $\hbar$ is the reduced Planck's constant.

1.4.2 Phonon Scattering Processes

Phonons contribute to thermal conductivity through the vibrational interactions of adjacent atoms, which are affected by impurities, crystal anharmonicities and geometric discontinuities. Phonons, electrons, and holes can scatter through several mechanisms as they travel through a material, with each scattering mechanism having an associated relaxation time: the average time between two consecutive scattering events. These events impact the momenta, directions and energies of the energy and charge carriers in the material and therefore play a key role in the determination of the thermal and electrical conductivities of a material by limiting carrier mobilities [18]. Figure 1.4 illustrates a few different scattering types, which include scattering by stationary defects (impurities, device boundary, grain boundary etc.) and dynamic defects (electrons and phonons).

Lattice self-heating is caused by phonon-electron scattering. These inelastic collision events transfer energy from the charge carriers (electrons accelerated by an applied electric field) to the phonons, with the direction of particle motion becoming increasingly random with

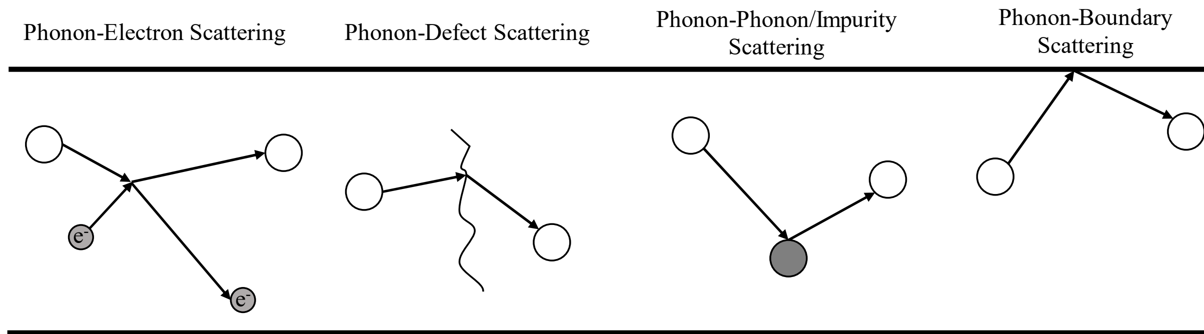

Figure 1.4: Illustration of scattering mechanisms.

consecutive collisions as opposed to being aligned with the electric field, therefore resulting in heat dissipation to the lattice. This increases the local temperature of the lattice and leads to suboptimal device performance. The likelihood of this type of scattering is higher when the material is heavily doped due to the increased abundance of charge carriers.

Other scattering events not involving charge carriers then cause further heat dissipation from the energized phonons. Boundary scattering is dependent upon surface roughness, and is particularly important in low-dimensional nanostructures. Defect scattering occurs due to the presence of grain boundaries, vacancies, or other deformities in the crystalline lattice. Phonons may also be scattered by other phonons or by differently-sized impurity atoms, especially in alloys and doped materials [16].

1.4.3 Energy Dissipation and Transport in Nanoscale Devices

The thermal conductivity of a material is strongly dependent upon the mean free path of the energy carriers in that material. For this reason, scattering mechanisms, which limit carrier mobility, are extremely important to the thermal conductivity. When limited by the crystal lattice anharmonicity, the thermal conductivity is termed intrinsic. The intrinsic limit is reached when the crystal is perfect, without defects or impurities, and phonons can only be scattered by other phonons [26]. If no anharmonicity were present, the material would have an infinite thermal conductivity.

Thermal conductivity is extrinsic when mostly limited by external effects, such as phonon-boundary or phonon-defect scattering [26]. In nanostructures, which have a higher surface-to-volume ratio than microstructures or bulk materials, the thermal conductivity is significantly reduced by scattering from boundaries, and the role of localized impurities and fluctuations in sizes also becomes important, causing extrinsic effects to dominate. For two-dimensional materials like graphene or transition metal dichalcogenide (TMDC) monolayers,

which are only a few times thicker than an atomic layer, interactions with the environment also influence the thermal conductivity [1].

When the characteristic length L of the device is significantly larger than the mean free path, phonons undergo a significant number of scattering events and a large amount of heat is dissipated to the lattice. This transport mechanism is labelled diffusive due to the randomness in energy carrier motion. If the device size is on the same order as the mean free path (typically in tens of nanometers), phonons undergo fewer scattering events, with some travelling without scattering. Called quasi-ballistic transport, the thermal conductivity is higher and local temperatures in the device are lower in this case given that the mean free path remains unchanged, since the phonons can transport energy away from the generation region more effectively without dissipating it to the lattice. In ballistic transport, $L \ll \lambda$, and no significant scattering occurs. As a result, the material appears to conduct electricity without heating up itself. Figure 1.5 illustrates the three different transport modes using a simplified 2D device representation.

Power dissipation in nanoscale circuits is determined by two components: dynamic power, used during switching for charging and discharging the load, and the static power consumption, which is the result of leakage current. The leakage component scales exponentially with the voltage overdrive, is very sensitive to changes in the threshold voltage, and is implicitly dependent on the total power dissipated as it is a strong function of temperature [12], [27], [28]. The downscaling of dimensions and voltage thus makes sub-threshold leakage increasingly important due to short channel effects like drain induced barrier lowering (where the drain is close enough to gate the channel and turn on the transistor prematurely).

The performance- and demand-driven trend of miniaturization that has led to a higher chip density has, in turn, resulted in unprecedented levels of power dissipation and temperature

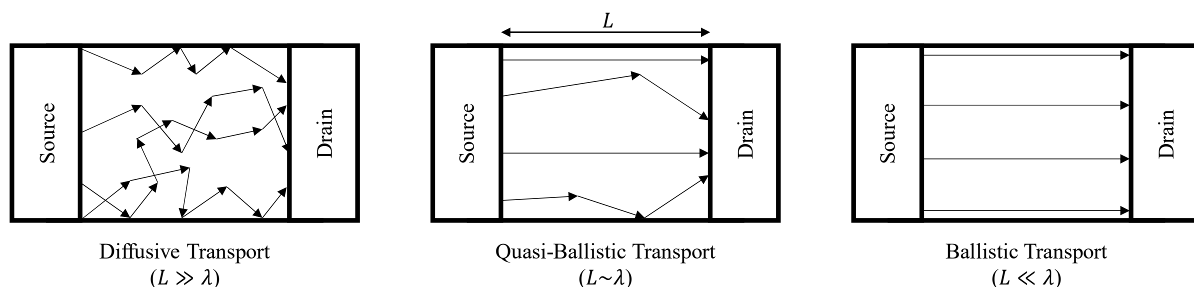

Figure 1.5: Illustration of transport mode classifications based on device size relative to mean free path.

at the chip level, raising reliability concerns for mesoscopic devices. Thermal conductivity in micro and nanoscale electronic devices is significantly lower than the material's bulk conductivity value since increased boundary confinement results in a reduction in the carrier mean free path [21]. Furthermore, the increased power density associated with a smaller device size leads to higher local temperatures, and the thermal conductivity of a semiconductor typically decreases with increasing temperatures (except at very low temperatures where an initial region of increasing thermal conductivity is observed) due to increased vibrational motion of the lattice atoms, which impedes the motion of energy carriers [20]. In the absence of an adequately designed cooling system, the reduction in thermal conductivity and the increase in local temperatures lead into one another in a cyclic chain, resulting in thermal runaway and breakdown.

1.5 Nanoscale Heat Conduction Models in Literature

Having established the significant role played by Joule heating in hindering the efficient performance of electronic devices, the basic physical mechanisms governing it, and how the problem is exacerbated by modern technological developments, it is pertinent to now review the methods established in literature that are used to study this phenomenon and help develop solutions to the challenges presented by it.

1.5.1 Experimental Methods

The complexity involved in using experimental methods to study Joule heating increases with a reduction in device dimensions. The deposition of thin films on substrates in modern electronic devices makes the accurate measurement of cross-plane film thermal conductivity challenging [29], [30] and necessitates the use of methods that operate on a picosecond timescale to perform surface-sensitive measurements and eliminate the impact of heat conduction in the substrate [31]. Several experimental techniques are capable of accurately providing information about the phonon modes and material conductivity. The upside of experimental techniques is the elimination of the need for simplifying assumptions associated with analytical and numerical models, but they are costly and time consuming, particularly due to the fine precision involved in preparing measurement instruments and high-purity samples.

Raman spectroscopy is a non-destructive optical technique that uses the inelastic scattering of photons with phonons to determine the modes of vibration of a system. The vibrational energy of the lattice atoms causes a shift in the energy of the incident photons emitted from a high intensity laser source [32]. The information about the intensities and

wavelengths of the scattered light show up as multiple peaks on a Raman spectrum, with each peak corresponding to a particular chemical bond and mode of vibration [33]. The technique can provide information about the phonon dispersion, lattice anharmonicity, phonon population, phonon lifetime and temperature distribution of the sample [34]–[40].

The time of flight technique analyses the current signal induced by the motion of charge carriers created by an ionizing radiation as a means of measuring the time taken by these charge carriers to cross a sample of known thickness under the influence of an applied electric field [41], [42]. Together, the transit time and sample thickness provide information about the carrier transport properties of the semiconductor sample, such as drift velocity, carrier mobility and carrier trapping, as well as their dependence on the electric field and temperature [43]–[45]. This, in turn, allows for the determination of the electrical and thermal conductivities of the material being investigated.

Microscale thermometry techniques are capable of performing measurements at length scales on the order of the mean free paths of energy carriers, thereby allowing for temperature mapping over the surfaces of micro and nanoscale devices at a high spatial resolution [46]. Scanning thermal microscopy uses a sharp-tipped probe to measure the local temperature on a surface, and measurement of the local temperature change subject to a known heat flux can allow for the thermal properties of electronic devices to be deduced and for localized hotspots and defects to be identified [47]–[49]. Scanning optical thermometry measures the local temperature by relying on optical reflectance, and uses lasers as sources of periodic heating [50]. It is better suited to the measurement of temperature distributions resulting from electrical heating in micro and nanoscale devices, as, unlike with scanning thermal microscopy, it does not depend on heat diffusion into a solid sensor and thereby eliminates the associated time delay [46]. Picosecond reflectance thermometry uses short pulses of laser light split into two beam paths with a mechanical delay between them [31]. Energy from the first beam causes an increase in temperature near the surface of the sample, with the decay of this temperature rise being measured by the energy of the second beam [46], [50]–[53].

1.5.2 Analytical Methods

In the analytical and numerical study of Joule heating, as well as for the theoretical interpretation of experimental results, a heat conduction model with an appropriate level of detail for the length scale and temperatures involved must be selected. Several different models to quantify Joule heating, with their own sets of assumptions, are widely used in literature and

are applicable to a specific range of conditions, with experimental results serving as the benchmark for accuracy. Given the advancements in computing technology, numerical simulations provide fast results with a range of visualization and postprocessing effects and are capable of handling complex geometrical configurations with reasonable accuracy.

The typical heat diffusion equation, expressed as [20]:

$$C\frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + Q''' \tag{1.5}$$

in which Q''' is the volumetric heat generation rate, is based on Fourier's law. It is thus only applicable at the continuum scale due to its inability to account for carrier scattering events and to capture heat transfer phenomena at short time and length scales. Similarly, the calculation of the power dissipated in a mesoscopic electronic device as:

$$Q = IV = I^2 R \tag{1.6}$$

(where Q is the dissipated power, I is the current flowing through the device, V is the potential difference across its terminals and R is the electrical resistance of the device) will result in an overestimation, since unlike at larger length scales, the energized carriers do not undergo a sufficient number of scattering events to fully dissipate their energy to the lattice before exiting modern devices, which operate in the quasi-ballistic regime [54].

In small devices, the low numbers of carriers present in total due to the nanoscale dimensions involved make the position of each carrier, dopant and impurity important to device performance, and the use of averaged densities and statistics leads to inaccuracies in calculations [55]. Keyes [56] was one of the first to predict that the variations in dopant atom concentrations in the channel region of transistors could lead to fluctuations in the device threshold voltage. In addition, the different scattering probabilities and group velocities associated with the each of the phonon modes necessitate the evaluation of the contribution of each mode separately at small length scales.

Important characteristic length scales that determine the applicability of various models to the study of Joule heating include the phonon mean free path, the phonon wavelength and the lattice spacing, as these help identify whether the phonons may be modelled as particles or waves [4]. The Debye temperature θ_D, which is the temperature at which a crystal's highest mode of vibration is excited, serves as a similar characteristic measure on the temperature

scale [57]. For silicon, the Debye temperature is 645 K [58], the acoustic phonon mean free path is around 300 nm [59] and the phonon wavelength is 1 nm at room temperature [60].

For $T < \theta_D$ and length scales shorter than the acoustic phonon mean free path, the Fourier equation is no longer applicable, and the phonon Boltzmann transport equation (BTE) must be solved instead, which treats phonons as semi-classical particles. To account for the granularity, it is necessary to solve the BTE for each phonon mode. At even shorter length scales, below the phonon wavelength, and/or for $T \ll \theta_D$, the quantum mechanical behavior of phonons becomes important, and lattice dynamics equations are used [60]–[62]. However, in typical electronic devices, the dimensions and temperatures involved are not so small as to require quantum mechanical modelling, and semi-classical methods suffice.

The Boltzmann transport equation incorporates the effects of external electric fields, electron-phonon interactions, phonon decay and various scattering events in its self-consistent formulation of phonon and electron transport [61]. Based on the fundamental assumption that a statistical distribution function $f(\boldsymbol{r}, \omega, t)$ of an ensemble of carriers exists—which depends on time t, position vector $\boldsymbol{r}$ and phonon frequency ω, and is a measure of the average particle (electron or phonon) occupation in the neighborhood of the location defined by $\boldsymbol{r}$ at t—and that each carrier occupies a spatial and momentum coordinate, the BTE applies the principle of charge conservation by equating the total rate of change in time of the distribution function to the change due to various scattering mechanisms in order to model carrier transport to a desired degree of accuracy [4], [63].

Unlike electrons, phonons have no charge and thus do not interact with any applied potentials, only being subject to the possibility of anharmonic decay via interactions through the atomic potential [64]. Therefore, no term associated with the influence of external forces appears in the phonon BTE, which can be expressed as [61], [63], [65]:

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \nabla f = \left(\frac{\partial f}{\partial t}\right)_s + \left(\frac{\partial f}{\partial t}\right)_g \tag{1.7}$$

in which $\boldsymbol{v}$ is the phonon velocity vector and the only change in the distribution function is due to the motion of phonons in real space. The first term on the right-hand side represents the rate of change in phonon occupation due to scattering events (which include anharmonic phonon decay). Moreover, every time an electron undergoes a transition between two states, a phonon is emitted or absorbed depending on the energy transfer involved, which is accounted for by the second term on the right-hand side.

Solution of the BTE requires an evaluation of the scattering term, and while approaches to do so by applying an integral in wave vector space over all scattering processes exist, the solution of this integral-differential form is a rather complicated task to accomplish [66], [67]. For this reason, the relaxation time approximation is instead commonly used in literature to model the scattering term, according to which the rate of change of phonon occupation due to scattering can be expressed as [68], [69]:

$$\left(\frac{\partial f}{\partial t}\right)_s = \frac{f_0 - f}{\tau_{ph}} \tag{1.8}$$

in which f_0 is the equilibrium phonon distribution and τ_{ph} is the relaxation time: the time taken by the phonons to relax to equilibrium due to all scattering processes combined, which is a function of position and momentum. The equilibrium phonon distribution can be calculated using Bose-Einstein statistics as [70]:

$$f_0 = \left[\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1\right]^{-1} \tag{1.9}$$

in which k_B is the Boltzmann constant, while the relaxation time is related to the phonon mean free path as $\lambda = v\tau_{ph}$. Substitution of Eq. 1.8 in Eq. 1.7 and integration over the phonon frequency and density of states allows for the BTE to be expressed in terms of the phonon energy density u as [70]:

$$\frac{\partial u}{\partial t} + \boldsymbol{v} \cdot \nabla u = \frac{u_0 - u}{\tau_{ph}} + Q''' \tag{1.10}$$

As evident from Eq. 1.10, solution of the BTE for the conduction problem also requires modelling of the phonon dispersion, electron energy bands, phonon transport and volumetric heat generation in the device. The following section reviews established methods for the evaluation of the heat generated by electron-phonon scattering in an electronic device.

1.6 Nanoscale Heat Generation Models in Literature

The inaccuracy involved in the computation of heat generation in a micro or nanoscale electronic device using the expression of Eq. 1.6 has been briefly discussed previously. Other shortcomings of that model are the inability to predict the nonuniform spatial distribution of heat generation and thereby identify critical design shortcomings, the lack of information about

the contribution of individual phonon modes to heat generation and the inability to account for hot energy carriers travelling into the device contacts and dissipating heat there [54]. Meanwhile, the direct solution of the BTE is a cumbersome process and analytical solutions exist only for simple heat conduction problems [68]. Alternate approaches to quantify heat generation in electronic devices that are well established in literature include the drift-diffusion, hydrodynamic and Monte Carlo simulation methods, which are based on the numerical solution of the BTE.

1.6.1 Drift-Diffusion Method

The drift-diffusion method has been one of the most used approaches for obtaining a solution to the BTE with the aim of evaluating heat generation in an electronic device. The volumetric heat generation term that appears in Eq. 1.10 is calculated as [71], [72]:

$$Q''' = \boldsymbol{J} \cdot \boldsymbol{E} + (R - G)(E_g + 3k_B T) \tag{1.11}$$

where the first time on the right-hand side is the dot product of the current density $\boldsymbol{J}$ and the electric field $\boldsymbol{E}$, and represents the volumetric rate of Joule heating. The second term, in which E_g is the band gap energy, represents the heat generation rate associated with the non-radiative electron and hole generation (G) and recombination (R) rates, which result in heat dissipation to the lattice either directly or via phonon emission from an excited charge carrier [16]. The current density can be calculated as:

$$\boldsymbol{J} = en\overline{\boldsymbol{v}} \tag{1.12}$$

where e is the elementary charge of an electron, n is the electron density and $\overline{\boldsymbol{v}}$ is the average electron velocity.

The advantage offered by this approach over the expression of Eq. 1.6 is the availability of spatial heat generation information through the involvement of the electric field and current density vectors in the mathematical formulation. However, this method still does not provide information regarding the contribution of the separate phonon modes, nor does it account for the fact that heat generation in a device is slightly offset from the electric field distribution, since energized electrons travel a few mean free paths before they are able to fully dissipate their energy to the lattice through phonon emission [73]. This method is therefore suited to device lengths that lie within the continuum regime, but has insufficient detail to accurately capture the heat generation phenomena in modern nanoscale devices.

1.6.2 Hydrodynamic Method

The hydrodynamic model, first used by Stratton [74] and Blotekjaer [75], is derived from the BTE using the method of moments. The zeroth, first and second moments, which respectively represent the electron charge, momentum and energy conservation, serve as the basic equations of this model, and are as follows [76]:

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{v_d}) = \left(\frac{\partial n}{\partial t}\right)_c \tag{1.13}$$

$$\frac{\partial \boldsymbol{p}}{\partial t} + \nabla \cdot (\boldsymbol{v_d}\boldsymbol{p}) = -en\boldsymbol{E} - \nabla(nk_BT_e) + \left(\frac{\partial \boldsymbol{p}}{\partial t}\right)_c \tag{1.14}$$

$$\frac{\partial W_e}{\partial t} + \nabla \cdot (\boldsymbol{v_d}W_e) = -en\boldsymbol{v_d} \cdot \boldsymbol{E} - \nabla \cdot (\boldsymbol{v_d}nk_BT_e) - \nabla \cdot \boldsymbol{Q} + \left(\frac{\partial W_e}{\partial t}\right)_c \tag{1.15}$$

Here, T_e is the electron temperature, $\boldsymbol{p}$ is the electron momentum density, W_e is the electron energy density, $\boldsymbol{v_d}$ is the electron drift velocity, and the subscript c indicates terms that change due to collisions, which can be rewritten using the relaxation time approximation. Furthermore, in the case where electrons are the majority charge carriers, the volumetric heat generation is evaluated as:

$$Q''' = \frac{3}{2}k_B\frac{n(T_e - T)}{\tau_e} + (R - G)\left[E_g + \frac{3}{2}k_B(T_e + T)\right] \tag{1.16}$$

where T is the lattice temperature, and τ_e is the electron energy relaxation time. In this case, the holes are assumed to be in thermal equilibrium with the lattice.

While the mathematical formulation of the drift-diffusion model is similar in structure to the hydrodynamic model since both are derived by applying the method of moments to the BTE, the latter is more detailed as it uses three moments as opposed to two in the former. The advantage offered by the hydrodynamic method arises from the fact that the majority carrier temperature is not assumed to be equivalent to the lattice temperature in this model, which allows for a more accurate prediction of carrier concentration, velocity and energy, particularly in the vicinity of highly peaked electric fields [77]. This improvement in accuracy makes up for the increased computational cost associated with the additional equation.

Despite the improved handling of carrier transport, this method still suffers from an inability to identify the contribution of the individual phonon modes to heat generation, and

the use of a single averaged carrier temperature T_e and relaxation time τ_e means that the energy-dependent scattering rates are not calculated to a desirable level of accuracy in mesoscopic devices [78]. Furthermore, the truncation of the model to a finite number of moment equations during derivation from the BTE results in the appearance of spurious peaks in carrier velocity profiles in the vicinity of peaked electric fields during device simulation, which are reduced when a greater number of equations are used [79]–[81].

1.6.3 Monte Carlo Simulations

The Monte Carlo (MC) method is a statistical approach that is used to predict the likelihood of various possible outcomes of a problem involving random variables and to perform risk assessment. Outcomes of events are drawn independently and at random from a collection weighted by pre-defined probabilities, while accounting for the interdependence of certain parameters [82]. Due to the stochastic nature of the process, the results obtained through a MC simulation are never exact, but lie within specified intervals with given probabilities, and the use of a larger sample size increases confidence in results. In addition to being applied to a problem that is purely statistical in nature, the MC method can also be used to seek solutions to a well-defined system of mathematical equations, or—as is the case with the simulation of semiconductor devices—a mixture of the two [83].

When applied to the study of carrier transport in semiconductor devices, the MC method uses a number of super-particles to represent groups of mobile charge carriers inside the device [83]. After attributing energy and momentum information to each particle based on the initial conditions, the motion of these particles under the influence of externally applied electric and magnetic fields is investigated. The pre-defined probabilities of the various scattering processes are used to randomly determine the duration of initial free flight that each particle undergoes subject to the external field, after which it is subjected to a randomly selected scattering phenomenon that reorients the particle momentum [84], [85].

Due to the individual particle-tracking nature of the simulation and the intentional incorporation of scattering phenomena, the MC method provides detailed information regarding phonon generation and the exchange of energy between electrons and phonons, which is important because different phonon modes with different group velocities have varying contributions to heat generation and current confinement in an electronic device. The volumetric heat generation rate at steady state can be computed as the sum of the energies of all phonons emitted minus the energies of all phonons absorbed per unit volume over a certain

simulation time t_{sim}, as follows [80]. This approach therefore accounts for every phonon generation/absorption event in its calculation, and consequently, is much more comprehensive than the drift-diffusion and hydrodynamic methods.

$$Q''' = \frac{1}{t_{sim}} \sum (\hbar\omega_{emitted} - \hbar\omega_{absorbed}) \tag{1.17}$$

The accuracy of the simulation improves with a more complete description of the various possible scattering mechanisms, and a significant amount of effort has been put into the development of MC models that contain a greater detail of the energy bands, phonon dispersion and scattering mechanisms while optimizing computational costs [86], [87]. The ability to incorporate accurate quantum mechanical treatment of the various scattering mechanisms and to simulate without making assumptions regarding the carrier distribution in energy space make MC simulations the most comprehensive approach for simulating charge transport in semiconductors [88].

1.7 Research Objectives

Having established how the demand for improved processing power and the trend of miniaturization necessitate the study of Joule heating in electronic devices, and having identified the different methods used to investigate the impact of this phenomenon, this work aims to:

1. Use Monte Carlo simulations to examine heat dissipation in nanoscale electronic devices. This work does not aim to develop a MC model from scratch; a combination of well-established models developed by past researchers will be compared to each other and evaluated for fitness of purpose.

2. Perform validation against published results for nanoscale device performance under various operating conditions in order to establish a level of confidence in the adopted numerical methodology, as well as to tune simulation parameters like the number of particles simulated, the time after which transients are assumed to die out and the total simulation time.

3. Investigate the effects of critical parameters (e.g. electric field strength, channel length etc.) on the performance of nanoscale devices, with the intention of identifying important transition points where device behavior changes as a result of crossover from the quasi-ballistic to ballistic transport regime as device dimensions are scaled down.

1.8 Thesis Outline

The discussion thus far has described, in detail, the physical phenomena at the atomic scale that cause Joule heating in electronic devices. The known impact of this self-heating on device performance has then been placed in the context of future device design and development goals. A cross-over in the nature of the transport of energy carriers through electronic devices has been identified as an inevitable outcome of further device miniaturization, and an explanation of the potential impact of this cross-over on the rate of Joule heating in modern devices has been presented. Based on a review of the existing knowledge of the field—which was found to be lacking on the subject of the performance of ultra-short devices operating in the ballistic transport regime—together with a review of the methods used to study Joule heating and device performance, appropriate research objectives have been defined for the present study in the previous section.

Chapter 2 of this document reviews and compares the models for the phonon dispersion, electron energy bands, and scattering mechanisms that have been implemented by past researchers in their Monte Carlo simulations to describe the carrier transport in electronic devices with varying degrees of accuracy, and which are applicable over differing ranges of operating conditions and length scales. The numerical methodology employed in the present work is then described in detail, and the mathematical relationships used to model the carrier transport phenomena are presented together with simplifying assumptions. The discretization schemes used to numerically solve the partial differential equations involved in the model are also discussed, in addition to the boundary conditions used for the simulation of realistic device geometries in later chapters.

Chapter 3 presents the results of the numerical simulations performed on bulk silicon, modelled as an infinite resistor, using the methodology described in Chapter 2. Following a validation of the employed methodology against data reported in published literature, the results of investigations into the impact of parameters such as the applied electric field strength, initial lattice temperature, and doping density on the heat generation and carrier transport characteristics of the bulk semiconductor are presented and discussed. The results of this chapter are particularly important for developing an understanding of the reasons behind the trends observed for realistic silicon devices in Chapter 4.

In Chapter 4, several parameters that could potentially have a significant influence on the performance of electronic devices, particularly those operating in the ballistic transport regime,

are identified based on the review of literature conducted in Chapters 1 and 2 and the results of Chapter 3. A one-dimensional geometry representing a simplified but realistic model of nanoscale electronic devices is then presented as a suitable specimen for numerical simulations, by means of which the gap in the existing knowledge identified in Chapter 1 could be addressed. The chapter describes the limitations and benefits of the assumptions made, and covers, in detail, the preliminary process of creating the geometry and mesh, assigning the appropriate initial conditions, and choosing suitable values for the grid spacing and time step size to be used for the numerical simulations. A comparison of the results obtained through the employed Monte Carlo technique and the alternative drift-diffusion approach is also presented, and the advantage offered by the former over the latter is highlighted. An extensive discussion of the impact of the aforementioned critical parameters then follows, which is based on the accompanying results of the simulations performed at various operating conditions on variations of the generic 1D device model. The chapter also touches on the use of two-dimensional simulations to achieve the objectives of this study, and presents reasons behind the adequacy of one-dimensional modelling for the intended purposes.

Chapter 5 summarizes the findings of this work, while laying out its limitations and identifying potential areas for improvement and further research in the future. Based on the findings of this study, recommendations for the future design of electronic devices and related cooling systems that could potentially improve device performance are presented, and avenues for further developments in this field through the numerical simulation of alternate materials and/or device geometries are identified.

CHAPTER 2: NUMERICAL METHODOLOGY

Upon investigating the various methods available to study Joule heating in electronic devices, as described in the previous chapter, the numerical Monte Carlo method was chosen for the present study due to the better accuracy it offers over other numerical techniques. Despite the higher computational costs involved, the MC method allows for flexibility in the descriptions of the phonon dispersion, electron energy bands, and scattering mechanisms through simplifying assumptions that are applicable to a certain range of operating conditions. This makes for significantly less computationally-intensive simulations than if full analytical descriptions were to be used. Furthermore, the method allows for simulations to be performed at conditions that cannot be maintained in physical experiments, for the investigation of the theoretical performance of materials that do not exist, and for the identification of the effects of material properties on the observed phenomena. This chapter describes the methodology adopted by the present study in detail, and includes a comparative discussion of the different MC models that have been used by previous researchers.

2.1 General Monte Carlo Algorithm

The basics of the Monte Carlo method as applied to various problems of interest have been covered in extensive detail in past studies [82], [83], [89], and form the basis for the development of Joule heating models with varying assumptions. The MC method is applicable to both steady-state and transient transport processes, although the former are less computationally demanding [83]. Unlike the basic MC method that is applied to steady state problems, in which the simulation of a single energy carrier for a sufficiently long period of time is reflective of the behavior of all carriers in the material [90], the simulation of time- or space-dependent phenomena requires the explicit simulation of an ensemble of carriers.

The present study uses the ensemble Monte Carlo approach to perform transient simulations of the carrier transport in semiconducting devices. Figure 2.1 presents a flowchart of the basic steps of the MC model implemented in the current study. Several thousand super-particles, each in turn representing billions of real electrons, are used in these simulations to model the mobile charge inside the semiconductor. Information about the applied electric field, device geometry, and mesh setup is obtained from the output of preliminary simulations performed on a commercial device simulator based on the drift-diffusion model (COMSOL Multiphysics [91]), as are the initial conditions for the carrier concentration and doping

profiles. This eliminates the potential convergence issues associated with randomly initialized distributions, and makes for faster MC computations [92].

In the MC method, the generation of a sequence of random numbers, subject to previously specified event probabilities, is integral to the simulation of the motion of charge carriers in an electronic device under the influence of external forces and scattering events. Moreover, for transient simulations, a set of numbers containing velocity and momentum information is associated with each individual carrier [82]. Scattering probabilities are computed at the beginning of the simulation as a function of the carrier energy [93], and are used to stochastically determine the initial free flight time interval τ during which the super-particles are allowed to drift unimpeded under the influence of the external electric field. To incorporate the influence of the lattice in the computation of the carrier acceleration due to the electric field, an effective electron mass is used instead of the free electron mass [83].

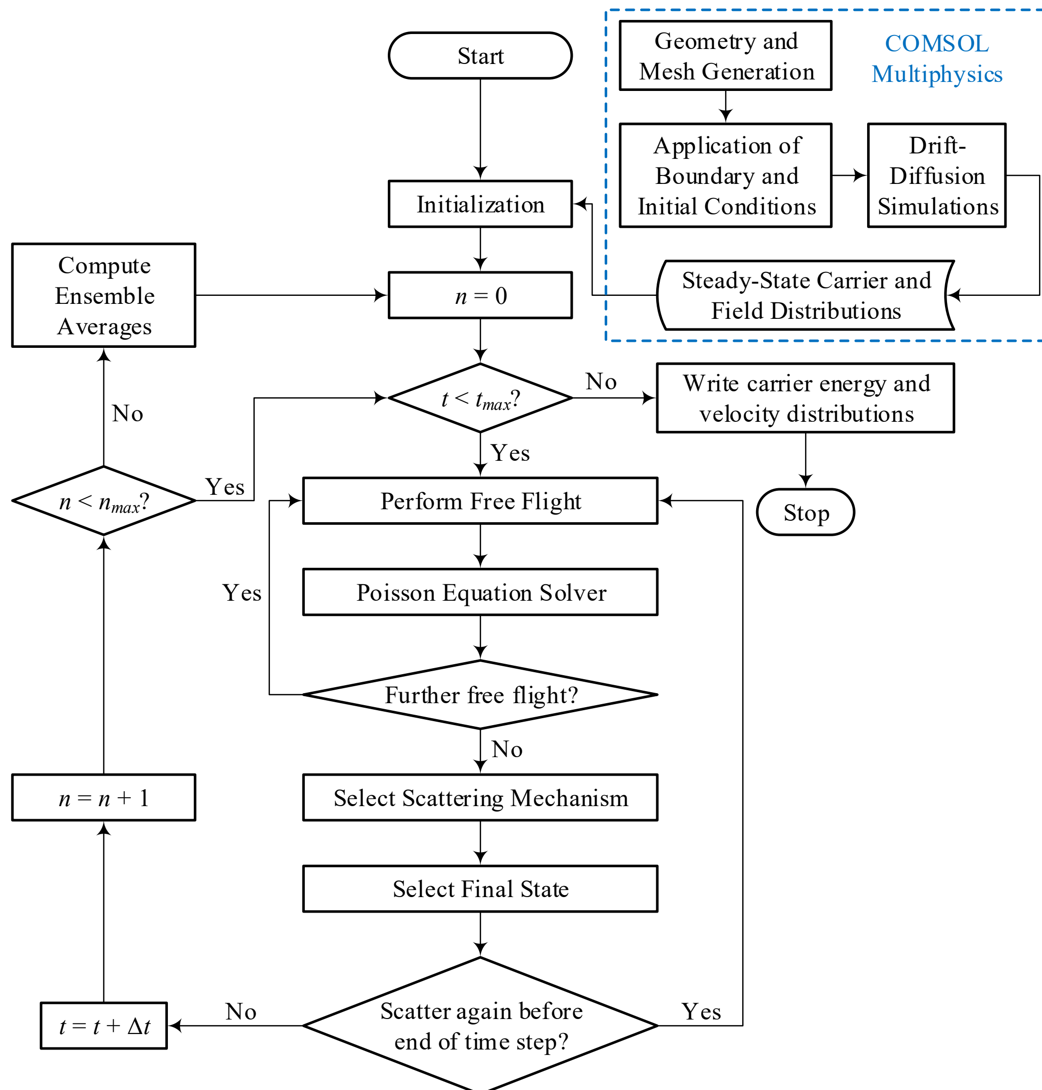

Figure 2.1: Flowchart of the ensemble Monte Carlo method used in this study.

When the free flight ends, a scattering mechanism is randomly applied in keeping with the predetermined probabilities. To maintain a constant total scattering rate Γ that is independent of the carrier energy and thus simplify computation, a fictitious self-scattering rate is introduced, as first described by Rees [94], [95], which does not alter the calculation of the distribution function. The free flight interval is then related to the total scattering rate as [83]:

$$\tau = -\frac{\ln(r)}{\Gamma} \tag{2.1}$$

where $0 < r < 1$ is a uniformly distributed random number. Note that Γ represents the maximum probability that the charge carrier will be subject to a scattering event.

During the free flight interval, the carrier momentum increases over time due to the external electric field. This momentum is altered by the collisions that end free flight, which are treated as instantaneous events because of the much shorter duration of time that they involve in comparison to τ [93]. The final state of the charge carrier after collision is also chosen stochastically while accounting for energy and momentum conservation; models with varying levels of complexity have been developed by past researchers to accurately capture the exchange of energy that occurs during these collision events. If self-scattering is selected, the carrier is assigned the same state after the event that it had before scattering.

For accurate and realistic device simulations, it is preferable to self-consistently update the electric field information as the charge carriers move through the device. Although some past works have used a fixed potential distribution determined at the beginning of the simulation [96], [97], it has been shown that self-consistent simulations provide more accurate results [83], [92], [98]. The Poisson equation, in which ϵ_s is the semiconductor dielectric constant, correlates the electrostatic potential ϕ to the net charge density ρ as [83]:

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_s} \tag{2.2}$$

and is used to update the electric field information periodically after intervals of free flight. The ensemble of carriers is simulated up to the field-adjusting time (which is less than the total free-flight time), after which the carrier distribution is frozen. Once the electric field has been updated, the simulation resumes using the carrier information stored when it was paused. The super-particles are treated as individual charge carriers for free-flight and scattering, but as clouds of charge during solution of the Poisson equation.

Due to the stochastic nature of the process, a significant amount of randomness is expected in MC results, which is reduced by averaging the results over a large number of time steps. The total simulation time t_{max} is divided into multiple sub-histories over which intermediate ensemble averages are calculated, and an initial portion of the simulation is not included in the computation of the final ensemble averages to exclude the transient effects of the initial condition selection. In the flowchart of Figure 2.1, n_{max} is the number of time steps over which intermediate ensemble averages are to be computed (a single sub-history), while n represents the number of time steps that have been simulated since the last ensemble averages were computed. The carrier information is updated after every sub-history and can be monitored to check convergence of the solution as the simulation progresses. A final output file containing the carrier information is written at the end of the simulation, in which the values have been averaged over all sub-histories for the steady portion of the simulation.

2.2 Monte Carlo Models in Literature

The behavior of electrons near the band edges (the highest energy state for the valence band and the lowest energy state for the conduction band) determines most electronic properties of a material, and an accurate modelling of the electron band structure is thus necessary to ensure realistic simulations. In semiconductors, the dependence of the electron energy E_k on the wave vector $\boldsymbol{k}$ is more complex than for an electron in free space: near the conduction band edge, $E_k(\boldsymbol{k})$ can be approximated by a parabolic function similar to that for free space, but the curvature of this function is significantly different and may also be dependent on the crystallographic direction [99]. This complexity of semiconductor band structures has prompted different groups of researchers to develop their own MC models with varying assumptions to describe the transport processes involved in semiconductor devices. Two different analytical approximations—the parabolic and the non-parabolic models—have been widely used in literature as alternatives to the full band model that simplify the electron band description and reduce computational costs.

In the study of Joule heating and carrier transport in electronic devices, consideration of the phonon dispersion is also important, since interactions with phonons can cause electrons to transition between states in different conduction band valleys. As made evident by Eq. 1.17, lattice self-heating occurs through electron-phonon scattering, which allows the electrons accelerated by externally applied electric fields to dissipate their energy and return to equilibrium as they leave the device [61]. MC models developed in the past differ in the level

of detail with which the phonon dispersion is incorporated. Furthermore, most MC models use selection rules that are heavily dependent on specified parameter values to determine important carrier properties. Despite the possibility that all such parameters may vary with energy, constant values assumed to be independent of direction and carrier energy are conventionally used in MC simulations [100]. The values of some of these parameters can be difficult to determine by theoretical means, and are instead tuned to fit benchmark experimental results [101]. Deformation potentials, which are representative of the expected degree of electron-phonon interactions, are one such example; significantly different sets of deformation potentials have been proposed by various researchers for the same semiconducting material.

Canali et al. [42] used the experimental time-of-flight technique to investigate the electron transport properties of high purity n-type silicon surface barrier diodes along different crystallographic directions for a range of temperatures and electric field strengths. These results were compared with those obtained through MC simulations that employed a many-valley model with parabolic electron bands and incorporated acoustic intravalley phonon scattering, in addition to intervalley scattering with six fixed-energy phonons and ionized impurity scattering. To achieve a reasonable degree of agreement between the experimental and MC simulation results, two separate sets of deformation potentials were compared, one of which incorporated the influence of low-energy intervalley phonons, while the other did not. It was found that the inclusion of these phonons in that model was necessary to accurately predict the electron drift velocity as a function of temperature, field strength and field orientation.

Brunetti et al. [102] performed a similar comparison between results obtained through time-of-flight experiments and MC simulations performed on $n^+ - n - n^+$ phosphor-doped silicon diodes at varying electric field strengths. Their multi-valley MC model modelled the electron bands non-parabolically and incorporated acoustic intravalley scattering in addition to intervalley scattering by six fixed-energy phonons, but did not consider impurity scattering. They presented a revised set of deformation potentials, which was also used by Jacoboni and Reggiani [85], that readjusted the weights associated with the intervalley scattering mechanisms to better fit the results of their own experiments.

Tang and Hess [103] were among the first to use a complete full band description for the band structure of silicon in MC simulations by employing the empirical pseudopotential method, which they coupled with the scattering mechanisms used by Canali et al. [42] and the revised parameter values proposed by Brunetti et al. [102]. Fischetti and Laux [86] compared experimental and simulation results for sub-micrometer silicon field-effect transistors (FETs)

at different temperatures. Their MC solver also used empirical pseudopotentials for a full band model, but improved on previous models by accounting for the separate contributions of the longitudinal and transverse acoustic intravalley scattering mechanisms. Additionally, they were the first to couple the MC solver with a two-dimensional Poisson equation solver to self-consistently update the electric field information during simulation.

Sangiorgi et al. [104] developed a MC device simulator capable of handling 1D and 2D metal-oxide-semiconductor (MOS) device geometries. The program was, however, developed to serve as the postprocessor of a simulator based on the drift-diffusion model, only considered one non-parabolic conduction band, and ignored intervalley scattering. Venturi et al. [92] demonstrated the feasibility of using a self-consistent MC solver to investigate the carrier transport characteristics, and subsequently improve the design, of sub-micron MOSFETs. Sano et al. [105] built on previous works that used full band models [86], [103] by introducing wavevector-dependent impact ionization rates that accounted for the anisotropy resulting from the realistic band structure to improve the modelling of high-field carrier transport in crystalline silicon; apart from the revised impact ionization process, which was treated as an additional scattering mechanism, the scattering mechanisms and deformation potentials presented by Canali et al. [42] were used.

Yoder and Hess [106] developed one of the most comprehensive MC simulators for carrier transport in bulk silicon by using a full-band description based on the empirical pseudopotentials method, accompanied by wavevector-dependent phonon scattering rates. Kunikiyo et al. [107] developed a similarly detailed model which left no room for adjustable parameters like deformation potentials, since the calculated scattering rates were consistent with the full band structure and the full anisotropic phonon dispersion curve of silicon. However, both these models could not be applied to the study of carrier properties in realistic device geometries due to the high computational requirements associated with the level of detail incorporated.

Duncan et al. [108] performed self-consistent ensemble MC simulations on different types of MOSFET structures to study the effects of device scaling on hot carriers by employing the full-band model, but limited the scattering model to energy-dependent scattering rates and ignored anisotropy in order to reduce the computational requirements of the simulation. Kosina [109] presented a new model to incorporate ionized impurity scattering in MC solvers which reduced small-angle scattering events, and showed that—except at very low

temperatures—the use of his isotropic scattering model for the heavily anisotropic process did not significantly affect the solution accuracy despite greatly reducing the computational time.

Pop et al. [110], [111] developed a MC model for electron transport in silicon which used an analytical non-parabolic band model as well as an analytical description of the phonon dispersion. This allowed for an increased computational efficiency in comparison to the more demanding full-band and full-dispersion models, while maintaining a mutually comparable level of accuracy in the description of the electron and phonon behaviors, unlike other models that described the electron band structure with high accuracy but used simplistic phonon dispersion descriptions, or vice versa. Their approach—suitable only for low-voltage applications—differentiated between heat generation through the transverse and longitudinal phonon modes in addition to the optical and acoustic modes, and was used to simulate carrier transport in bulk and strained silicon for a range of temperatures and electric field strengths. Furthermore, the work presented a revised set of deformation potentials that was found to facilitate more accurate simulations in comparison to previous models. This same model, which treated all scattering events inelastically in order to correctly compute the transfer of energy and, in turn, allow for the determination of localized heat generation rates, was used in later works [2], [73] to study Joule heating in ultra-short silicon devices subject to quasi-ballistic transport conditions.

Efforts to couple the charge and heat transport processes in MC simulations have also been undertaken more recently. Aksamija and Ravaioli [61], [112], [113] coupled phonon transport with a full-band and full-dispersion MC simulation to trace the propagation of phonons after their emission and generate local temperature distribution maps over time for a silicon transistor. Raleva et al. [114] developed an electrothermal MC device simulator that self-consistently solved energy balance equations for both optical and acoustic phonons and fed back the phonon temperatures to the electron transport solver for temperature-dependent scattering rates. This simulator was used in a subsequent work by Vasileska et al. [60] to investigate the effects of device dimensions and boundary conditions on the temperature distribution and current degradation in nanoscale silicon devices.

Additionally, MC models originally developed for the study of silicon devices have since been modified to investigate the electrical and thermal properties of novel materials. Ashok et al. [115] were the first to consider the effects of electron-electron interactions in the MC simulations of nitride devices, and found that these cause the charge carriers to lose energy more rapidly near the drain end of the device than if not incorporated. Harada et al. [116]

compared the differences in the electron transport properties of silicon and monolayer graphene diodes through self-consistent MC simulations by modeling the former with a single-valley parabolic band model and the latter with a linear band model. Nofeli and Arabshahi [87] used an ensemble MC approach to study the electron transport properties of bulk ZnO and $Zn_{1-x}Mg_xO$, while Adeleh and Reza [88] investigated the properties of sub-micrometer ZnO diodes due to interest in the material as a possible substitute for silicon.

Despite concerns about the scaling limit of silicon being approached, it remains the most widely used semiconductor in electronic applications, primarily due to its natural abundance and favorable properties. Shomali et al. [117] performed energy carrier transport simulations using the MC method to investigate the effects of changing boundary conditions on the temperature distribution profiles and localized hotspots in silicon nanodevices. Fang et al. [118] demonstrated the contribution of both acoustic and optical phonon emission to the energy loss in silicon and germanium devices through full band MC simulations. Nghiem et al. [119] studied the temperature distribution and heat confinement resulting from ballistic transport in a nanoscale silicon sample heated on the top by a localized heat source. With the industry now aiming for the 5 nm lithography node, it is important that the behavior of electronic devices like silicon diodes is studied to understand the impact of ballistic transport becoming increasingly dominant on device performance characteristics.

2.3 Details of the Numerical Model

As aforementioned, the objective of this work was not to develop a new or revised Monte Carlo methodology for the simulation of carrier transport in semiconductor devices, but instead to apply a well-established approach to investigate the effects of ballistic transport on device behavior. For this purpose, features of the various methodologies developed over time were considered—as discussed in the previous section—and the model presented by Pop et al. [110] was chosen due to its computational efficiency, correct treatment of the energy transfer through scattering events, and detailed modeling of the band structure and phonon dispersion. The propriety of this model for the simulation of simple but realistic silicon device geometries has been demonstrated by subsequent works [73], [80], [111]. The salient features of the implemented model are discussed in this section.

2.3.1 Electron Band Model

As discussed previously, the MC solver developed by Pop et al. [110] used non-parabolic analytical descriptions for the electron band structure. This approximation, which

significantly reduces computation time in comparison to full-band methods, is justified for sub-bandgap device voltages and low energy studies, since impact ionization and high-energy transport are insignificant at these conditions. This model built upon the parabolic band model used by Canali et al. [42] by adding a temperature-dependent non-parabolicity parameter α to the mathematical relationship between the electron energy E_k and the wavevectors k_i. This relation can be expressed as follows for a non-parabolic conduction band with ellipsoidal equi-energetic surfaces (of which there are six in silicon), by using the many-valley model [83]:

$$E_k(1 + \alpha E_k) = \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y} + \frac{k_z^2}{m_z}\right) \tag{2.3}$$

where m_i is the reduced component of the effective mass tensor along the ith direction. The temperature dependence of the non-parabolicity parameter for silicon was incorporated as:

$$\alpha = 0.5\frac{E_{g_0}}{E_g(T)} \tag{2.4}$$

where E_{g_0} is the band gap energy at room temperature, and $E_g(T)$ is the temperature-dependent band gap, which, for silicon, can be expressed in terms of the absolute temperature T as [120]:

$$E_g(T) = 1.1756 - (8.8131 \times 10^{-5})T - (2.6814 \times 10^{-7})T^2 \tag{2.5}$$

Therefore, at room temperature, $\alpha = 0.5$ eV^{-1}, while using $\alpha = 0$ results in the original parabolic model of Canali et al. [42]. Due to the limitation on the maximum electron energy imposed by the non-parabolic band approximation, the implemented model ignores inter-band scattering with higher energy bands, since the second conduction band valley of silicon lies approximately 1 eV above the bottom of the band edge.

2.3.2 Scattering Mechanisms

To be able to calculate the rate of heat generation in the devices investigated in this study, it is essential that the exchange of energy between the carriers is accurately captured. The model proposed by Pop et al. [110] is suitable in this regard, since it models all scattering events inelastically. Furthermore, scattering with the longitudinal acoustic (LA), transverse acoustic (TA), longitudinal optical (LO) and transverse optical (TO) phonon modes is evaluated separately in order to determine the balance of emission and absorption events for each phonon type and to thereby quantify the individual contribution of each mode to the heat

generation rate. The phonon dispersion for each mode is modeled with the quadratic analytical approximation expressed as follows, in which the set of values presented by Pop et al. [111] for the dispersion coefficients ω_0, v_s and c is used. These values are chosen to replicate the experimentally determined dispersion relationship, similar to the approach employed by Fischer and Hofmann for only acoustic phonons [121].

$$\omega = \omega_0 + v_s q + c q^2 \tag{2.6}$$

Intervalley electron-phonon scattering is incorporated in the present methodology using six fixed-energy phonons in a manner similar to that adopted by Canali et al. [42] and Jacoboni and Reggiani [85]. Three of these intervalley phonons are associated with intervalley scattering on the same wavevector axis (g-type), while the other three are associated with scattering between valleys on mutually perpendicular axes (f-type). The intervalley scattering rate Γ_{if}, which is a function of the initial electron energy, can thus be calculated as [85], [110]:

$$\Gamma_{if} = \frac{\pi \Delta_{if}^2 Z_f}{2\rho_m \omega}\left(N_q + \frac{1}{2} \mp \frac{1}{2}\right) g_{df}(E_k \pm \hbar\omega) \tag{2.7}$$

Here, Δ_{if} is the constant deformation potential associated with each phonon mode, which helps determine the relative strength of the f- and g-type scattering processes. The revised set of deformation potentials presented by Pop et al. [110] is used in the present investigation. Z_f is the number of final valleys available ($Z_f = 1$ for g-type scattering and 4 for f-type scattering), ρ_m is the mass density of silicon, N_q is the phonon occupation, as determined through Bose-Einstein statistics, and g_{df} is the energy-dependent density of states in the final valley.

Intravalley electron-phonon scattering, which occurs within the same conduction band valley, is typically caused only by acoustic phonons [122]. The total intravalley scattering rate Γ_i, also dependent on E_k, is therefore computed by using the following expression separately for the LA and TA modes, as also done by Pop et al. [110]. The optical modes are not considered.

$$\Gamma_i = \frac{D_a^2 m_d}{4\pi \rho_m \hbar^2 k_s} \int_q \frac{1}{\omega}\left(N_q + \frac{1}{2} \mp \frac{1}{2}\right) I_q^2 q^3 dq \tag{2.8}$$

Here, D_a is the isotropically averaged deformation potential corresponding to each phonon mode, m_d is the electron density of states effective mass, I_q is the wave function overlap

integral modelled after the rigid ion approximation and k_s is the electron wave vector transformed to the spherical Herring-Vogt space [85], [123]. In both Eq. 2.7 and 2.8, the upper and lower signs correspond to phonon absorption and emission, respectively.

Ionized impurity scattering, which is an elastic but heavily anisotropic process, occurs due to the influence on electrons of the locally unbalanced charge distributions that result from the presence of ionized impurities in the lattice. The effect of this scattering mechanism on carrier mobility becomes more pronounced at higher doping densities due to the increase in the impurity concentration. The aforementioned method developed by Kosina [109] (also used by Pop et al. [73]), which employs an isotropic process with an equivalent momentum relaxation time to reduce the number of small angle scattering events in comparison to the anisotropic approach adopted by earlier researchers, has been used to model ionized impurity scattering in this study.

2.3.3 Self-Consistent Poisson Equation Solution

The significance of solving the Poisson equation at regular time intervals to obtain updated information regarding the spatial electric field distribution in a simulated electronic device has been discussed previously. The net charge density ρ, used in the Poisson equation as expressed in Eq. 2.2, can be calculated as [83]:

$$\rho = -e(n - p - N_D^+ + N_A^-) \tag{2.9}$$

where n and p are the electron and hole densities, respectively, and N_D^+ and N_A^- are the concentrations of the ionized donors and acceptors. All four of these variables are functions of position. Because the adopted methodology only simulates electron transport, and not that of holes, p and N_A^+ are zero. The Poisson equation can thus be expressed as:

$$\nabla^2 \phi = \frac{e}{\epsilon_s}(n - N_D^+) \tag{2.10}$$

To solve the Poisson equation, it is necessary to first map the charge carrier distribution obtained from the transport simulation onto the structured grid generated over the device geometry. This is achieved through the commonly-used cloud-in-cell (CIC) particle-mesh scheme. In comparison to other particle-mesh algorithms like the nearest-grid-point scheme, CIC results in continuous force values that are much smoother and have less fluctuations due to its use of piecewise-linear interpolation. When applied to one-dimensional simulations, the

CIC scheme uses a simple weighted linear interpolation to assign the charge Q of a super-particle at position x to its two nearest grid nodes located at x_i and x_{i+1} as $w_1 Q$ and $w_2 Q$ respectively, where $x_i < x < x_{i+1}$ and the interpolation weights w_1 and w_2 are calculated as:

$$w_1 = \frac{x_{i+1} - x}{x_{i+1} - x_i} \tag{2.11}$$

$$w_2 = 1 - w_1 = \frac{x - x_i}{x_{i+1} - x_i} \tag{2.12}$$

A numerical solution of the Poisson equation is achieved by applying a central finite difference approximation and solving the resulting matrix equation (in which the coefficient matrix is tridiagonal for one-dimensional simulations and pentadiagonal for two-dimensional simulations [83]). The discretized form of the Poisson equation for one-dimensional simulations on a uniform grid (which simplifies the charge assignment process) with spacing Δx, as performed in this study, can be expressed as:

$$-\phi_{i-1} + 2\phi_i - \phi_{i+1} \approx \frac{e\Delta x^2}{\epsilon_s}\left(N_{D\,i}^{+} - n_i\right) \tag{2.13}$$

Once the equation is solved, the components of the electric field can be calculated as the negative derivative of the electrostatic potential with respect to the direction. For the one-dimensional uniform grid case, this can be expressed as:

$$E_i = -\left.\frac{d\phi}{dx}\right|_i \approx -\frac{\phi_{i+1} - \phi_{i-1}}{2\Delta x} \tag{2.14}$$

Both the discretized expressions of Eq. 2.13 and Eq. 2.14 are second-order accurate due to the use of the central differencing scheme. The approximate equality arises due to the truncation of higher order terms in the Taylor series expansion. Once the electric field information has been updated, the carrier transport simulation resumes subject to the new electric field until the next field-adjustment time, when the process is repeated.

2.3.4 Boundary Conditions

In the one-dimensional simulations performed in this work, two boundaries on either end of the device are present, which represent the device contacts. The contacts serve as an interface between the device domain and an ideal source or sink of charge carriers. A periodic boundary condition is applied for particle transport, such that charge carriers leaving the device

from one contact are reinjected with thermal energy and inwards momentum at the other contact, thus ensuring current continuity. The current flowing through the device is then calculated as the difference between the rates of flow of charge carriers leaving the device in the direction aligned with the electric field and those leaving in the opposite direction.

The potentials ϕ_1 at the first and ϕ_n at the last grid node of the device are taken to be fixed and are determined by the applied voltage so that $\phi_n - \phi_1 = V_D$ (the drain voltage). Furthermore, to maintain a consistent level of accuracy in the calculation of the electric field from the electrostatic potentials using the finite difference method, second-order accurate off-centered differencing is used for the boundary nodes. The electric field at the boundaries is thus calculated for a uniform grid as:

$$E_1 = -\frac{-3\phi_1 + 4\phi_2 - \phi_3}{2\Delta x} \tag{2.15}$$

$$E_n = -\frac{3\phi_n - 4\phi_{n-1} + \phi_{n-2}}{2\Delta x} \tag{2.16}$$

2.4 Summary

This study investigates the carrier transport properties and heat generation in simple electronic device geometries using the numerical Monte Carlo technique. Analytical descriptions are used for the electron band structure and phonon dispersion to balance computational efficiency and accuracy. All electron-phonon scattering events are treated inelastically to allow for the contribution of each phonon mode to heat generation to be determined. Furthermore, the Poisson equation is solved at regular intervals during the carrier transport simulation to self-consistently update the electric field information.

CHAPTER 3: CARRIER TRANSPORT IN BULK SILICON

3.1 Introduction

The miniaturization of electronic devices has been facilitated by significant advancements in device fabrication and cooling technologies. The results of experiments and simulations performed through various techniques have been used to make informed decisions regarding modifications to device geometries so that critical locations of localized heat generation can be better addressed. The development of numerical models to simulate device behavior has also allowed for the investigation of novel geometries and materials without the costly fabrication of physical specimens. Before a numerical model can be reliably applied to the simulation of realistic devices, however, it is essential to verify that it captures the carrier transport characteristics of the material accurately.

Numerical models can be validated against the experimental results available for existing device geometries, as well as against previously reported numerical results performed at the same operating conditions. This allows for the values of the variable selection parameters, like the deformation potentials, to be adjusted to match empirical relationships, which is especially important since these parameters determine the relative strengths of the scattering mechanisms in the material being investigated. Once confidence in the adopted methodology has been established by verifying that the carrier transport phenomena are accurately modelled, the same model can then be extended to more complex problems.

Detailed investigations of the carrier transport properties of bulk silicon under a wide range of conditions have been carried out by past researchers through both experimental and numerical means, and the results of the former have been frequently used as a benchmark against which to compare the results obtained through Monte Carlo models incorporating various assumptions [42], [101], [110], [111]. To validate the methodology adopted in this work, the electron transport properties of bulk silicon, modelled as an infinite resistor subject to a constant applied electric field, have been investigated. The effects of various important parameters, such as the initial lattice temperature and electric field strength, on the steady state carrier mobility and heat generation rate are studied one at a time, and are compared against the findings of previous researchers for corresponding cases, where possible. The latter activity

helps to ensure that the present model correctly incorporates the effects of such parameters and thereby simulates the carrier transport process accurately, while placing new findings in the context of previously understood phenomena. The net rates of emission of each phonon mode at various operating conditions are also compared to study the effects of the variable parameters on the spectral composition of heat generation in the lattice, which allows for the impact of the differing phonon group velocities on heat generation and carrier transport to be understood.

3.2 Variation of the Electron Drift Velocity

The influence of both the lattice temperature and the electric field strength on the drift velocity of charge carriers in various materials is well understood through the results of extensive investigations performed by past researchers. An attempt is therefore made to validate the methodology implemented in the present study by testing whether the simulation results accurately replicate the reported dependence of the steady state electron drift velocity in silicon on the applied electric field strength and the lattice temperature. For this purpose, the semiconductor lattice is modelled as an infinite resistor subject to a constant applied electric field at a uniform specified initial temperature. Simulations are performed for a total duration of 30 ps, with time steps of 2 fs. The initial 10 ps of carrier transport statistics calculated during the simulation are not included in the final ensemble averages used to compile results, as this is taken to be the length of time needed to allow transients to settle, determined after observation of the convergence of carrier statistics over time from preliminary test simulations. Intermediate ensemble averages are calculated after every 75 time steps (every 150 fs), so that a total of 134 intermediate ensemble averages contribute to the final ensemble averages.

The effect of the applied electric field is investigated by increasing its strength in logarithmic steps, while keeping all other parameters constant. The entire MC simulation process is repeated at each field strength value and the calculated electron drift velocities are stored. The influence of lattice temperature on the drift velocity is then investigated by changing the specified temperature while repeating the entire set of simulations for the whole range of electric field strengths investigated previously. Furthermore, simulations over the same range of field strengths are also performed on samples with varying dopant concentration levels to verify that the ionized impurity scattering model works as intended, as well as to study the influence of doping density on the carrier transport.

Figure 3.1 presents a selection of the results of the simulations performed, and compares them with the experimental time-of-flight results obtained by Canali et al. [42], as well as with

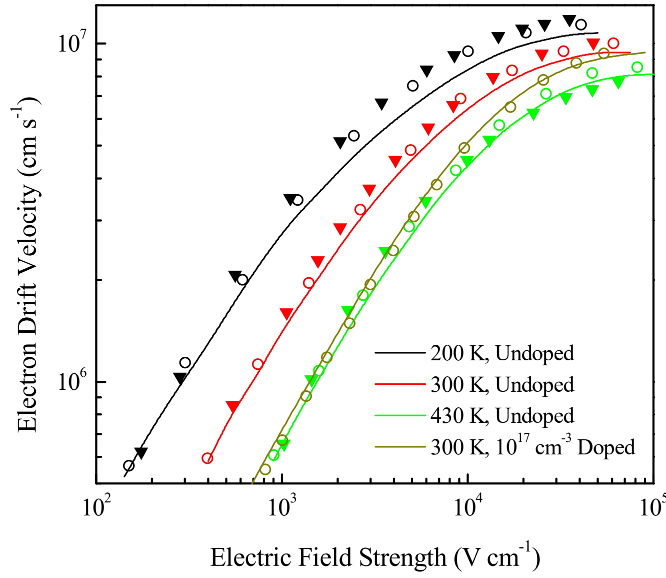

Figure 3.1: Validation of the variation of the electron drift velocity in bulk silicon with the applied electric field strength, lattice temperature and doping density. Solid lines are the results of the present work, triangles represent the experimental results of Canali et al. [42], and circles represent the MC results of Pop et al. [110], [111].

the MC simulation results obtained by Pop et al. [110], [111] through their revised set of deformation potentials. The results are found to be in good agreement with reported values, with a maximum error of less than 15% (observed at $T = 200$ K) and an average error less than 10% at all temperatures. Figure 3.2 presents the variation of the electron drift velocity with the applied electric field strength for a wider range of temperatures, as computed in the present investigation, while Figure 3.3 presents the effects of changing the semiconductor doping density on the carrier velocity while the lattice temperature is kept constant.

One trend observed to be common to all three of these figures is that, irrespective of the lattice temperature or the doping density, the electron drift velocity increases with increasing electric field strength up to a certain point, after which the trend levels out. Additionally, the drift velocity is found to be proportional to the field strength at low electric fields. This is because a stronger electric field accelerates the charge carriers to a higher velocity, up to a maximum possible saturation velocity v_{sat}, which is dependent on the material and the lattice temperature. Velocity saturation is caused primarily by electron-phonon scattering, and is of significant concern in the design of modern short-channel electronic devices [124].

Figure 3.2 indicates that the saturation velocity, as well as the drift velocity at any given electric field strength, decreases as the lattice temperature increases. This finding is consistent

with physical expectations, since the increased atomic vibrations at higher temperatures impede electron drift by increasing the frequency of collisions. It is also observed that a stronger electric field is required to achieve velocity saturation at higher temperatures.

Moreover, it is found that, at a constant temperature and field strength, an increased doping density contributes to a reduction in the carrier drift velocity, as indicated by the results shown in Figure 3.3. This is explained by the additional hindrance of carrier free flight by the

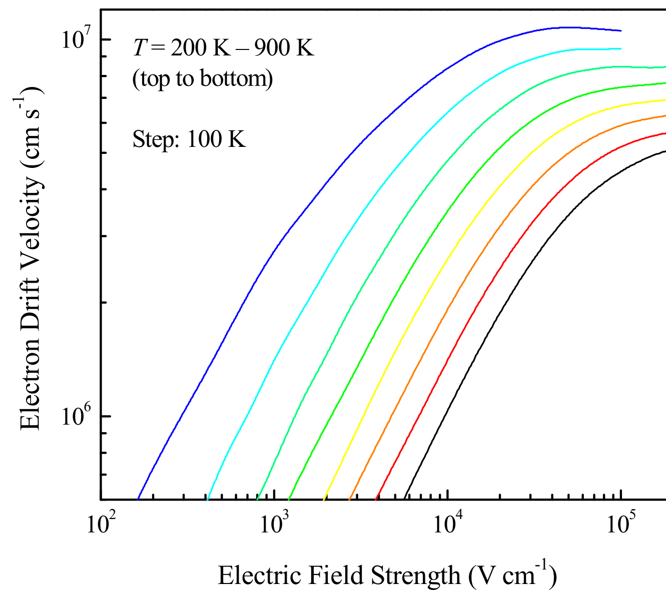

Figure 3.2: Variation of the electron drift velocity in intrinsic bulk silicon with the lattice temperature and applied electric field strength.

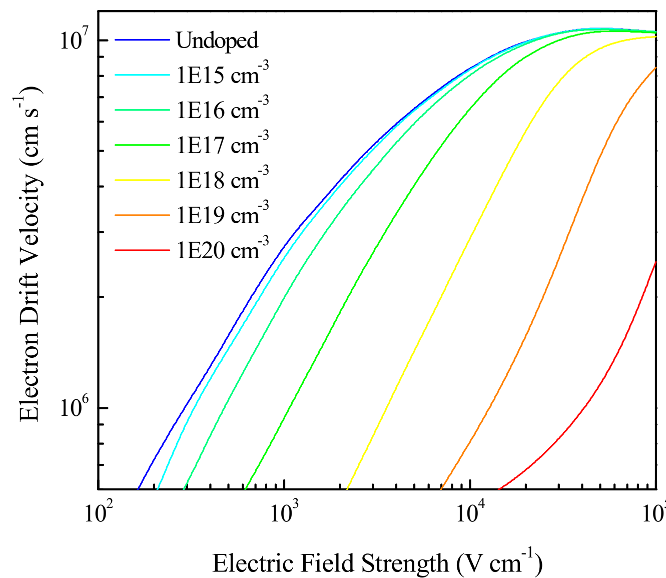

Figure 3.3: Variation of the electron drift velocity in bulk silicon at 200 K with the doping density and applied electric field strength.

added impurity atoms. The influence of doping on the carrier drift velocity is significantly more pronounced at lower electric field strengths, which is to be expected since ionized impurity scattering plays a major role in the determination of the carrier transport properties in this regime [125]. The saturation velocity is found to remain unaffected by the doping density, although a stronger electric field is required to overcome the increased resistance to carrier drift and achieve velocity saturation.

3.3 Variation of the Electron Drift Mobility

The electron drift mobility μ, defined as the ratio of the drift velocity to the electric field strength: $\mu = {}^{v_d}/_{E}$, is a measure of the rate of flow of electrons under the influence of an applied electric field. It is an important parameter in the design of electronic devices since, in conjunction with the carrier concentration, it determines the electrical conductivity of a material. The drift mobility in semiconductors is known to be dependent on both the lattice temperature and the doping density, as well as on the electric field strength at high fields [126]. Simulations have therefore been performed on bulk silicon to investigate the effects of these parameters on the carrier drift mobility, in a manner similar to how they were performed to study the variation of the drift velocity as discussed previously.

As observed from Figure 3.2 and Figure 3.3, the electron drift velocity varies linearly with field strength in the low-field regime, so the low-field drift mobility is independent of the field strength. However, as the field strength is increased, the average electron energy in the lattice also increases, and when the optical phonon energy (51 meV for silicon [2]) is exceeded, the probability of optical phonon emission increases drastically. As aforementioned, optical phonons have much slower group velocities than their acoustic counterparts, and although they eventually decay into the faster acoustic phonon modes, the decay process takes a much longer time (~5 ps) in comparison to the electron-phonon scattering time (~0.1 ps) [2]. This, therefore, leads to an increase in the optical phonon concentration over time at high electric fields, which impedes carrier drift, results in a breakdown of field strength-drift velocity proportionality, causes the carrier mobility to start decreasing with increasing field strength, and eventually leads to the velocity saturation phenomenon mentioned previously.

Figure 3.4 presents the results of the simulations performed on undoped silicon to investigate the effects of the lattice temperature on the low-field drift mobility. Comparisons are also made with the results of Canali et al. [42], Pop et al. [110], and Green [120]. The results are found to be in good agreement with the mobility values reported in literature, with

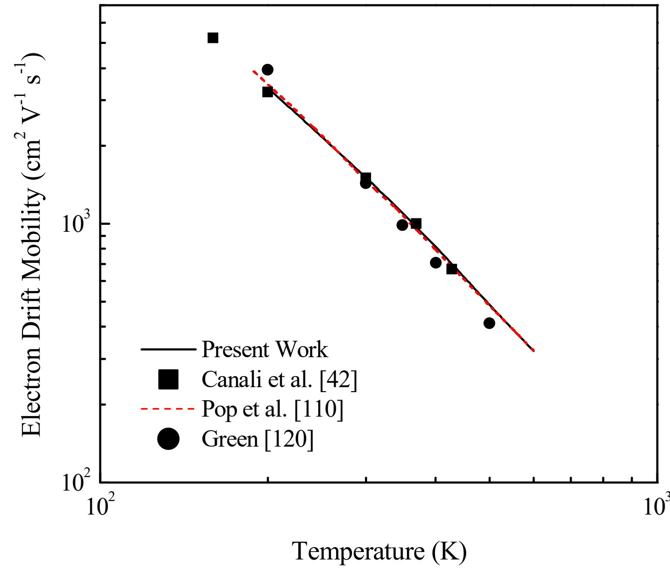

Figure 3.4: Validation of the variation of the electron drift mobility in intrinsic bulk silicon with the lattice temperature at a low electric field strength of 500 V cm^{-1}.

a maximum error of less than 15% in comparison to the results of Green [120], and an average error of less than 2% in comparison to the MC simulation results of Pop et al. [110]. The observed reduction in drift mobility with increasing temperature is expected, since the carrier motion in the direction of the electric field is hindered by increased atomic vibrations at higher temperatures.

Figure 3.5 illustrates how the low-field drift mobility in bulk silicon varies with the doping density and lattice temperature. It is observed that, at a constant temperature, the drift mobility initially decreases at an increasing rate with increasing dopant concentration. At low doping levels, the mobility is near-constant and is limited primarily by electron-phonon scattering, but the influence of ionized impurity scattering increases with increasing doping density. As the doping density is increased beyond 10^{18} cm^{-3}, however, an inflection point is encountered, after which the drift mobility starts to decrease at a decreasing rate, and even increases at low temperatures. This behavior is observed because the high concentration of ionized impurity atoms at such near-degenerate doping levels results in the material attaining electrical conductivity characteristics between those of a metal and a semiconductor. At low temperatures, the increase in the number of charge carriers in the lattice outweighs the negative impact of the increased electron-impurity scattering at high doping levels, resulting in the observed improvement in mobility. At higher temperatures, by contrast, the higher vibrational energy of the lattice atoms means that the net effect of the increased doping density is still to reduce drift mobility through increased ionized impurity scattering, albeit at a slower rate.

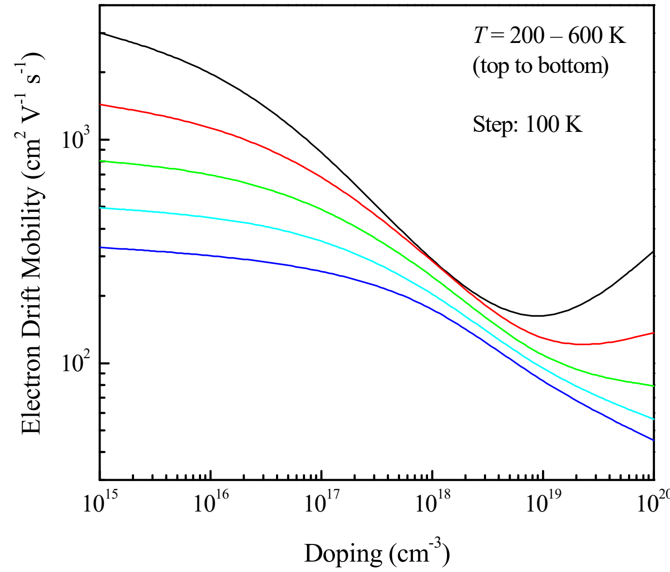

Figure 3.5: Variation of the electron drift mobility in bulk silicon with the doping density and lattice temperature at an applied electric field strength of 500 V cm^{-1} (low-field regime).

3.4 Phonon Emission Rates

The determination of the net rate of emission of each individual phonon mode during device operation is essential to the quantification of the localized rates of heat generation in the silicon devices investigated in this study. This is due to the varying group velocities and cutoff energy thresholds associated with the different phonon modes, which cause their respective contributions to heat generation and charge carrier transport to differ, as discussed previously. The implemented methodology therefore models each phonon dispersion branch separately, and treats scattering events with all phonon modes inelastically to ensure that the exchange of energy between the electrons and phonons is captured accurately.

In MC simulations, the choice of deformation potentials, dispersion coefficients, and/or other variable parameters can significantly affect the relative strengths of the scattering mechanisms, and in turn, the phonon generation rates. Several different sets of deformation potentials have been proposed by past researchers for the same scattering model. As aforementioned, the set of parameters used by Pop et al. [110] has been selected for the present investigation, as it has been shown to better reflect the empirical dispersion relationship of silicon. Simulations have been performed on doped bulk silicon to investigate the net rates of generation of the longitudinal and transverse phonon modes at different field strengths, and to determine the individual contribution of each phonon mode to the volumetric heat generation rate. To verify that the dispersion relationship has been described correctly and that scattering

events with the different phonon modes are simulated accurately, comparisons are also made with the values reported in literature.

Figure 3.6 illustrates the net rates of phonon generation in bulk silicon by phonon energy at two different applied electric field strengths. The phonon energy E is related to the phonon frequency ω as $E = \hbar\omega$, so the rates of phonon emission can be mapped to each of the four phonon types included in this study through understanding of the phonon dispersion relationship of silicon (as described by the quadratic approximation of Eq. 2.6), which provides the respective cutoff energies of each mode. Optical phonons (both LO and TO) correspond to the higher frequency (or equivalently, the higher energy) end of the spectrum, by contrast to acoustic phonons, which correspond to low- and mid-range frequencies. This is because the changes in interatomic distances are typically larger in the optical mode due to the out-of-phase atomic movements, whereas the coherent displacements that characterize the acoustic mode result in lower phonon frequencies for that mode.

A large difference in the cutoff energy levels for the TA and TO modes, as mandated by the phonon dispersion relationship, results in the region of no transverse phonon emission at mid-range energies observed in both Figure 3.6 (a) and Figure 3.6 (b). For the longitudinal branch, however, the permitted energy levels are continuous between the minimum and maximum thresholds, with $E \approx 50.9$ meV serving as the point of transition between the LA and LO modes [4]. This allows for the continuous longitudinal phonon generation spectrum

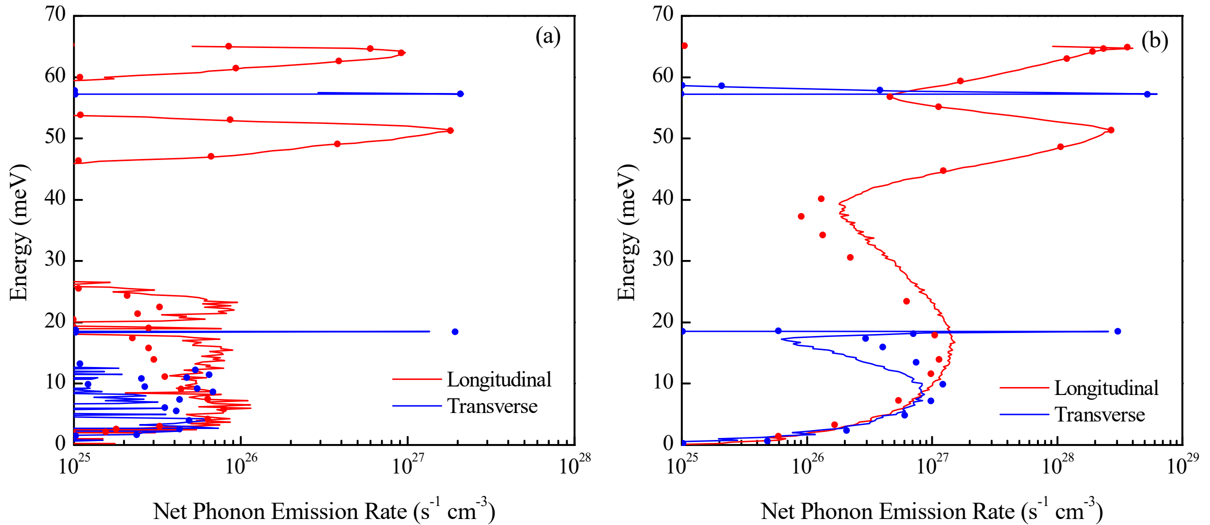

Figure 3.6: Net phonon generation spectrum for 10^{17} cm^{-3} doped bulk silicon at 300 K and an applied electric field of (a) 5 kV cm^{-1} and (b) 50 kV cm^{-1}. Solid lines represent results of the present work, while dots represent the results of Pop et al. [111] for the corresponding cases.

observed in Figure 3.6 (b). Furthermore, by comparing Figure 3.6 (a) and Figure 3.6 (b), it is evident that the rates of phonon emission increase significantly with an increase in the electric field strength, regardless of the phonon mode. At the weaker field strength, the lower rate of emission of phonons is due to the lower rate of intravalley scattering that results from a smaller number of electrons with sufficiently high energy levels being available in the lattice.

Intervalley scattering with the six fixed-energy phonons contributes to the appearance of the sharp peaks in the phonon generation plots. The relative magnitude of these peaks is therefore dependent on the values of the deformation potentials used in the simulations. A high degree of agreement between the peak values, and a reasonable match between the phonon generation spectrum results in general, is observed upon comparison of the results of the present simulations with those of Pop et al. [111] (whose set of deformation potentials has been used in this study). This indicates that the phonon dispersion and scattering events have been modelled accurately.

3.5 Variation of the Volumetric Heat Generation Rate

The total heat generation rate in the lattice can be computed through Eq. 1.17 by accounting for the energy of each phonon emitted and absorbed during the simulation. Furthermore, since detailed spectral information about the type of the phonons involved in the scattering events is also available through the dispersion relationship, the heat generation can be classified by each phonon mode. Figure 3.7 illustrates the spectral composition of the volumetric heat generation in bulk silicon over a range of applied electric field strengths. Once again, comparisons are made with the work of Pop et al. [111] to validate the simulations performed, and an average error of less than 5% is observed for all four phonon modes.

It is observed that, for both the longitudinal and transverse modes, the optical modes have a higher contribution to the energy dissipated in the lattice than their acoustic counterparts. Consistent with previous observations [12], it is found that the optical phonon modes (LO and TO) contribute to roughly two-thirds of the energy dissipation, with the LO mode alone accounting for approximately 60% of the total energy dissipated at any given applied field strength. This is due to the lower group velocity of the optical phonon modes, which translates to a greater involvement in scattering events with the charge carriers. Scattering events result in the motion of the charge carriers becoming increasingly random instead of being aligned with the electric field, thereby resulting in increased internal energy and lattice temperature. By comparison, the faster acoustic modes present less of a hindrance to electron transport,

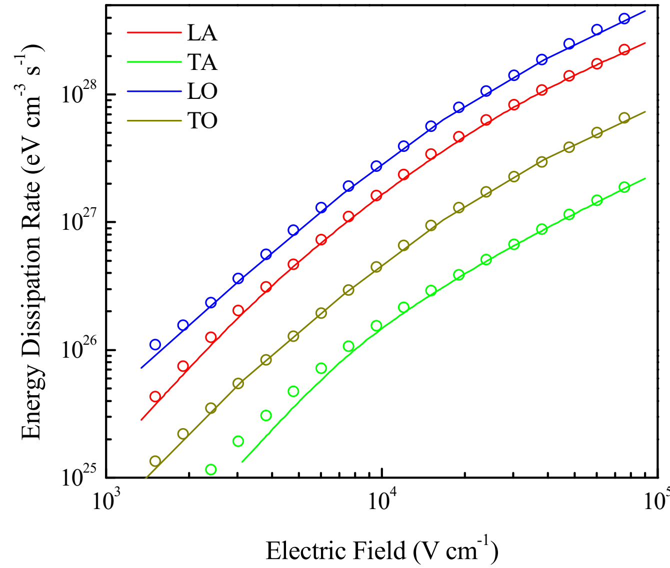

Figure 3.7: Validation of the variation with the applied electric field strength of the rate of energy dissipation through each phonon mode in 10^{17} cm^{-3} doped bulk silicon at 300 K. Solid lines are the results of the present work, while circles represent the results of Pop et al. [111].

allow for energy to be transported away from the point of phonon emission at a faster rate, and thereby result in reduced dissipation to the lattice. Additionally, it is also observed that the longitudinal modes (LA and LO) have a slightly less than 90% contribution to the total energy dissipated, explained by their larger population in the lattice, which is made possible by the much wider range of dispersion-permitted energies at which longitudinal phonons can be emitted compared to the range for transverse phonon emission, as indicated by Figure 3.6.

Further simulations were also performed to investigate the effects of the lattice temperature and doping concentration on the volumetric energy dissipation rate. Figure 3.8 indicates that the rate of energy dissipation increases significantly with the doping density at a constant temperature and field strength. This is explained by the increased ionized impurity scattering that occurs at high doping levels, and is physically consistent with the observations made from the results shown in Figure 3.3, where the carrier drift was impeded by increased dopant concentrations. Together, the two figures indicate that the increased scattering events result in a greater proportion of the supplied electric potential energy being dissipated as waste heat in the lattice, instead of being converted to useful kinetic energy in the form of electrical current. Furthermore, the minor improvement in the low-field drift mobility at low temperatures and near-degenerate doping levels, observed in Figure 3.5, is found to translate to a slight reduction in the rate of increase of the energy dissipation rate with the doping at near-degenerate doping levels in Figure 3.8, which is less pronounced at stronger electric fields.

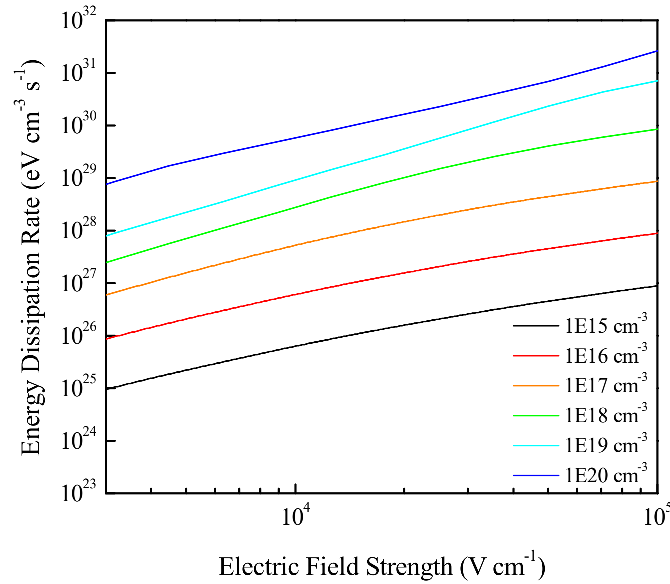

Figure 3.8: Effect of the doping density and applied electric field strength on the volumetric energy dissipation rate in bulk silicon at 300 K.

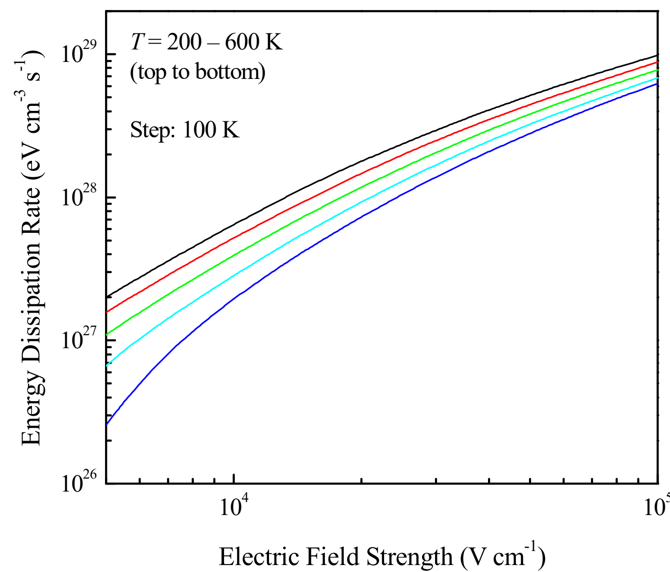

Figure 3.9: Effect of the lattice temperature and applied electric field strength on the volumetric energy dissipation rate in 10^{17} cm^{-3} doped bulk silicon.

Figure 3.9 illustrates the relationship between the energy dissipation rate in doped bulk silicon and the lattice temperature, and indicates that the heat generation rate is slightly reduced at higher temperatures. The hindrance of carrier motion due to the increased atomic vibrations in the lattice at higher temperatures, as observed in Figure 3.2 and Figure 3.5, means that electrons flow through the modeled resistor at a slower rate. Although, as discussed previously, an increase in the doping levels results in a similar reduction in drift velocity, the accompanying

increase in carrier concentration which occurs in that case is not achieved through an increase in the lattice temperature. For the same applied electric field, the current flowing through the resistor therefore decreases, whereas the electrical resistance increases, with increasing lattice temperature. This subsequently translates to the observed reduction in the rate of energy dissipation at a particular field strength.

3.6 Summary

The effects of several important parameters, including lattice temperature, doping density, and applied electric field strength, on the carrier transport properties of bulk silicon have been investigated in this chapter through Monte Carlo simulations. The computed values of the performance measures have been validated against results reported in published literature to verify the adopted methodology, and a high degree of agreement has been observed. The rate of energy dissipation to the silicon lattice through the four different phonon modes included in the numerical model has also been determined over a range of operating conditions, which is of particular importance to this study of Joule heating in electronic devices. The results presented in this chapter, important to the understanding of the electrical properties of bulk silicon, have additional significance in serving as a basis for the explanation of the phenomena observed in the study of more realistic electronic device models in the following chapter.

CHAPTER 4: JOULE HEATING IN BALLISTIC SILICON DEVICES

4.1 Introduction

Having studied the carrier transport and heat generation in bulk silicon under the influence of an externally-applied electric field, as discussed in the previous chapter, the same numerical methodology was extended to the investigation of the electrical and thermal properties of simple silicon devices. One-dimensional n^+nn^+ devices have been frequently used in literature as a simplified model of the cross section of a MOSFET channel [73], [80], [88], [115], [116], [127]. The energy band diagram for this simplified model is similar to that of a MOSFET channel, with an injection barrier and a highly peaked lateral electric field, but the multi-dimensional potential gradients and quantum confinement effects encountered in full device simulations are not present [72]. Furthermore, all scattering mechanisms and boundary conditions are incorporated as they would be realistically encountered in a characteristic cross section of the MOSFET channel. The one-dimensional simplification thus allows for the various methods used to simulate carrier transport (MC, drift-diffusion, hydrodynamic models etc.) to provide an understanding of the effects of operating conditions on transport phenomena in electronic devices in isolation from the complicating factors that arise in a full device simulation and which result in an increase in the computation time and cost.

Figure 4.1 shows a schematic representation of the 1D n^+nn^+ device investigated in this study. The n-type channel has a mild doping of 1×10^{16} cm^{-3} (almost intrinsic doping), while the initial device temperature, channel length, electrode lengths, electrode doping, and applied potential difference are varied for the purposes of this study. Through a review of the physical phenomena governing the operation of these electronic devices, as well as through the results presented in Chapter 3, these parameters have been identified to be potentially critical to the device performance, which is directly dependent upon the carrier drift velocity, carrier energy,

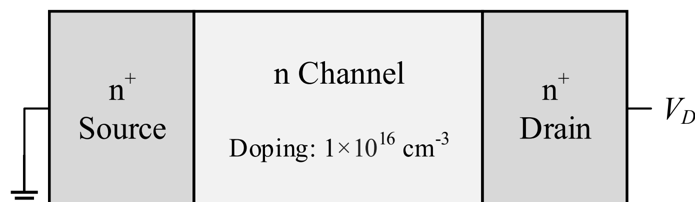

Figure 4.1: Schematic diagram of the n^+nn^+ device investigated in this study.

and rate of Joule heating. The applied voltage and device dimensions are both expected to have a significant role in the determination of carrier energies and velocities through their influence on the electric field distribution in the devices. The electrode doping, meanwhile, directly affects the number of available charge carriers in, and thus the electrical resistance of, the devices. The initial lattice temperature is anticipated to have a significant impact on carrier mobility and the rate of Joule heating in these devices through an extension of its impact on the electronic properties of bulk silicon.

The device geometries and corresponding meshes are created in the COMSOL Multiphysics software, through which preliminary simulations for carrier transport are performed until the achievement of steady state conditions. The information regarding the spatial distributions of the charge carrier concentration, dopant concentration, electric field, and electric potential—obtained from the output of each preliminary simulation—is then used to provide the initial conditions for the MC simulation performed at the same operating conditions in order to reduce the likelihood of encountering numerical convergence issues, as explained previously in Chapter 2. The potential difference across the device boundaries for the MC simulation is also fixed by the steady-state potential profile obtained from the preliminary simulation. Furthermore, a periodic boundary condition is applied to the device contacts, such that the particles escaping from one contact are reinjected with thermal energy and inwards momentum at the other contact to maintain current continuity through the device. A uniform grid, with grid independence checked beforehand, is used to simplify the charge assignment process and the formulation of other finite difference approximations discussed previously. Moreover, for the purposes of the following discussion, the start of the channel is taken as the reference point of the x-coordinate, except where stated otherwise.

4.2 Preliminary Setup

In order to create the device geometry, generate a suitable mesh, apply the desired initial conditions to the semiconductor lattice, and subsequently obtain the initial electric potential and charge carrier concentration profiles required for the numerical Monte Carlo method, the semiconductor module of the COMSOL Multiphysics software was used in the present work. This simulation package employs the drift-diffusion approach to model the transport of charge carriers, and its results are therefore only suitable to be directly interpreted for devices with characteristic feature sizes significantly larger than the phonon mean free path. However, while this solver does not incorporate hot carrier effects or velocity overshoot, and assumes the

electron energy bands to be parabolic [91], its steady-state results—even for the nanoscale devices investigated in this study—are sufficiently realistic to serve as initial conditions for the MC solver, which only uses this information to speed up the numerical convergence, as opposed to using randomly oriented momenta to initialize the super-particle distribution.

Important simulation setup information regarding the lattice temperature, applied electric field strength, total simulation time and the time step size is provided to the MC solver separately through an input text file to ensure that the primary simulation is not potentially affected by inaccurate data inputs in the preliminary COMSOL simulation. The geometry and mesh information, however, is entirely obtained through the grid coordinate outputs provided by COMSOL, so that the initial conditions obtained from the preliminary simulations can be directly mapped to the grid, although the same objective could be accomplished alternatively by specifying the minimum and maximum boundary extents together with the number of divisions or the spatial step size within the MC solver. To achieve an optimum convergence rate, the preliminary simulations are performed at the same operating conditions (lattice temperature, applied potential difference, dopant concentration, etc.) at which the actual MC simulation for the corresponding case is intended to be performed. The spatial distributions of the electric field strength, electric potential, charge carrier concentration and dopant concentration are then imported into the MC solver through another text file.

Within the COMSOL semiconductor module, the geometry is created with a 1D spatial dimension setting. This module solves the Poisson equation to update the electric potential distribution over the device geometry with the motion of the charge carriers, and both electron and hole transport is simulated through the drift-diffusion equations. The accompanying graphics window provides a visualization of the created geometry—a snapshot of a 10 nm channel $n^{+}nn^{+}$ device with 150 nm long source and drain regions, as created in this module, is shown in Figure 4.2 below. The source, channel and drain are created as separate intervals (left

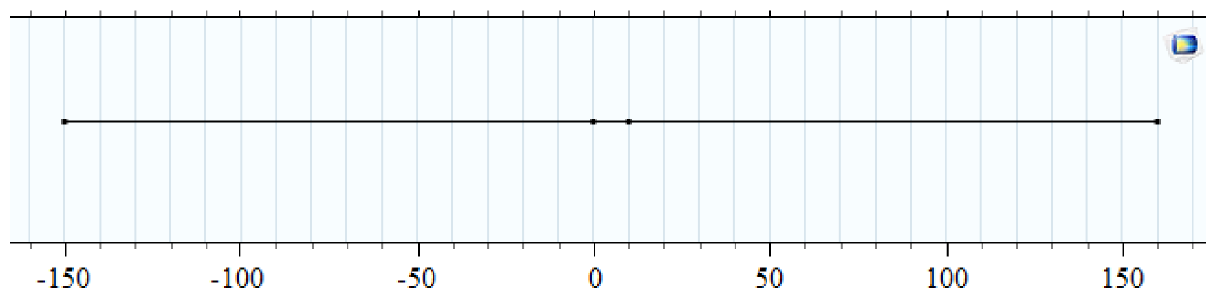

Figure 4.2: Snapshot of the geometry of a 10 nm channel 1D $n^{+}nn^{+}$ device created using the COMSOL Multiphysics software. The *x*-axis length scale is in nanometers.

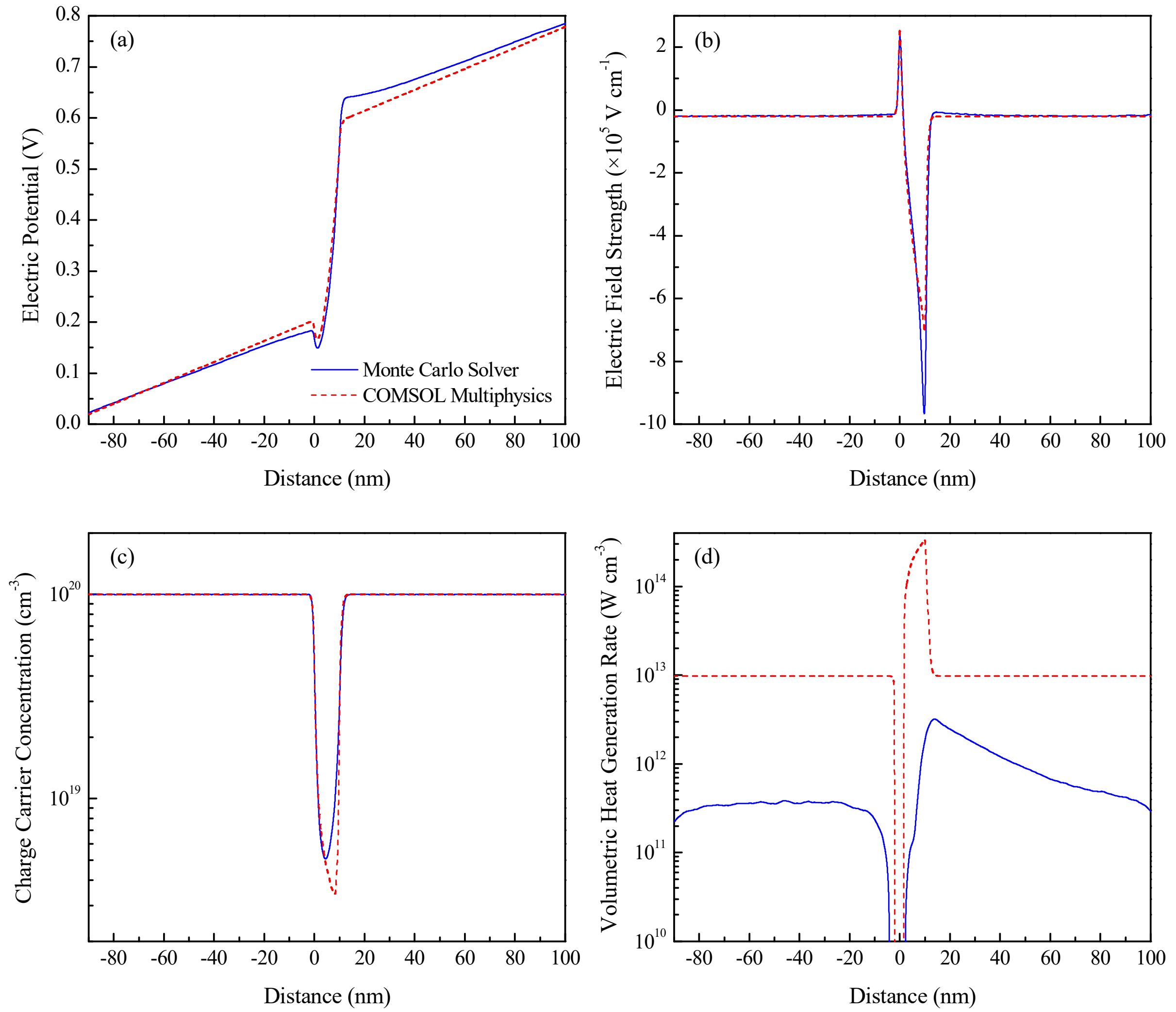

Figure 4.3: Comparisons of the distributions of characteristic parameters over a section of the length of a 10 nm channel n^+nn^+ device, as determined through a MC simulation, with the preliminary drift-diffusion simulation (COMSOL Multiphysics) results for the corresponding case. The device, with 100 nm long electrodes doped to 10^{20} cm^{-3}, is subjected to an applied potential difference of 0.8 V and an initial uniform temperature of 300 K.

to right) by providing the appropriate start and end coordinate values, and creating a union at their mutual interfaces. Separate doping types and concentration levels can then be specified for each of these intervals. Furthermore, the ends of the device are treated as ideal ohmic metal contacts, across which the desired constant potential difference V_D is applied.

Figure 4.3 presents the results obtained for a representative case through a Monte Carlo simulation performed using the initial conditions generated from the steady-state results of a preliminary drift-diffusion simulation. These steady-state results are also shown to facilitate comparison between the MC and drift-diffusion methods. Figure 4.3 (a) indicates that the

electric potential profiles determined through both methods are mostly similar, although there are slight differences in the magnitudes of localized peak values. The electric field strength distributions, shown in Figure 4.3 (b), which are the negative gradient of their corresponding potential trends, as expressed in Eq. 2.14, are thus also mostly similar, with the localized peak magnitude difference becoming pronounced particularly at the drain end of the channel.

In Figure 4.3 (c), a difference in the spatial location of the minimum charge carrier concentration in the device, as determined by the two differing simulation techniques, can be observed. With the current density constrained to be constant throughout the device to maintain current continuity, Eq. 1.12 indicates that a higher carrier velocity implies a lower carrier concentration. In the MC simulation, the maximum carrier velocity is skewed towards the source side of the channel due to the injection of cooler electrons from the drain [72], as shown in Figure 4.7 (c), which results in a lower carrier concentration at the source side of the channel than the drain side—opposite to the predictions made by the drift-diffusion approach. Because the COMSOL solver does not consider velocity overshoot, the maximum carrier velocity in its calculations is limited to the saturation velocity of the material, which is attained near the drain end of the channel. Moreover, note that the MC method allows for velocity overshoot, which is a significant phenomenon in mesoscopic devices, and thus computes a higher peak carrier velocity—the value of the current density computed by the two methods thus also differs.

The most significant indication of the inability of the drift-diffusion approach to accurately predict carrier transport characteristics in the vicinity of highly peaked electric fields is found in the results of Figure 4.3 (d). With the $Q''' = \boldsymbol{J} \cdot \boldsymbol{E}$ approach employed by the drift-diffusion method, the volumetric heat generation rate is found to be almost an order of magnitude higher than the MC results throughout the length of the device. The drift-diffusion calculations for heat generation follow the trend of the electric field distribution, which incurs errors in ultra-short device geometries. By contrast, the MC method, which computes the total heat generation rate through an aggregation of the energy exchanged via individual collision/scattering events, is much more accurate because it accounts for the distance that the energized charge carriers must travel before they are able to collide with the lattice atoms and release a discrete amount of energy per collision (constrained by the energy of the emitted phonon). Thus, while the drift-diffusion approach is sufficiently accurate for larger electronic devices with characteristic length scales in the continuum limit, this comparison indicates that it cannot be used to accurately predict the behavior of mesoscopic devices. Note that a semi-log plot is used in Figure 4.3 (d) to allow for a visual comparison between the trends that differ

by over an order of magnitude, so regions of negative heat generation, i.e. localized lattice cooling, cannot be observed, but a detailed discussion of the MC results containing the complete trends is presented in the following subsections.

4.3 Time Step and Grid Spacing Sensitivity

To avoid encountering divergence errors during algorithm execution and to ensure the reliability of results obtained from numerical simulations, it is essential that insensitivity of the results to the temporal and spatial discretization resolutions is established. The incorporation of the Poisson equation in the model imposes a maximum time step size limit, given by [93]:

$$\Delta t < \frac{1}{2}\sqrt{\frac{\epsilon_s m_t}{e^2 n}} \tag{4.1}$$

where m_t is the transverse mass of electrons in silicon and n is the mobile electron density. The satisfaction of this condition is necessary to avoid charge imbalance due to plasma oscillations. Since ϵ_s, m_t and e are constants, a higher electron density implies a stricter time step restriction. For $n = 10^{20}$ cm^{-3}, which is the maximum value of semiconductor doping density considered in most of this study, it is found that $\Delta t < 1$ fs is necessary.

Moreover, to accurately capture the motion of charge carriers within the discretized domain, it is also preferable that the selected time step size is not so large as to allow the charge carriers to move by a length of more than one grid cell before the parameter distributions are reevaluated. The time step Δt is thus also limited by the grid spacing Δx as:

$$\Delta t < \frac{\Delta x}{v_{max}} \tag{4.2}$$

where v_{max} is the maximum velocity of the electrons in the device. The magnitude of v_{max} to be used in the calculation of the time step limit can either be based off values reported in literature, or selected following test runs of the MC solver—in the present case, it is taken to be 6×10^5 m s^{-1}.

Figure 4.4 presents the sensitivity of the heat generation rate to the time step size for a representative 1D device geometry and a uniform horizontal grid spacing of $\Delta x = 0.25$ nm. The device electrodes are doped to 10^{20} cm^{-3}, and a potential difference of 0.8 V is applied across the device contacts. It is evident from the figure that a significant amount of error is incurred when $\Delta t > 1$ fs, as predicted by Eq. 4.1. A measurable reduction in the value of the

maximum heat generation rate is observed even for $\Delta t < 1$ fs, since the condition of Eq. 4.2 requires $\Delta t < 0.42$ fs. For this reason, several time steps that are significantly shorter than the maximum limit are also considered. After considering the tradeoff between the improvement in accuracy and the accompanying increase in computation time, a time step size of $\Delta t = 0.2$ fs is selected for the present study, since a choice of $\Delta t = 0.1$ fs improves accuracy by less than 2.5% but roughly doubles the simulation time.

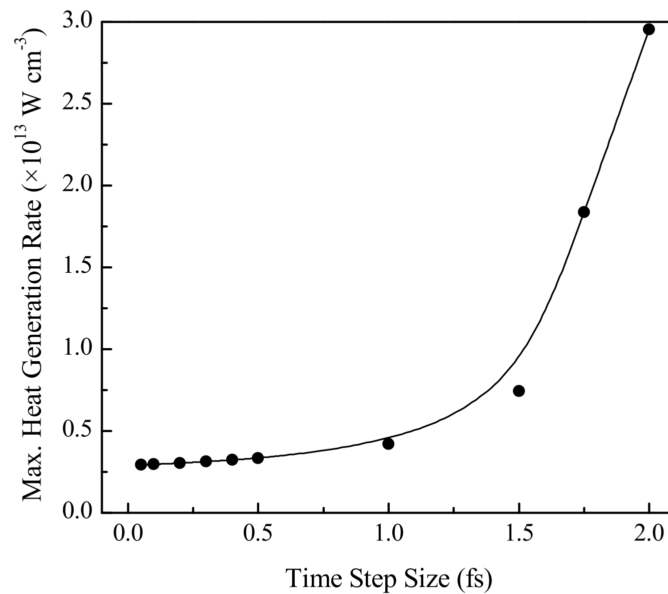

Figure 4.4: Variation of the maximum volumetric heat generation rate in a 10 nm channel n^+nn^+ device at 300 K with the simulation time step size.

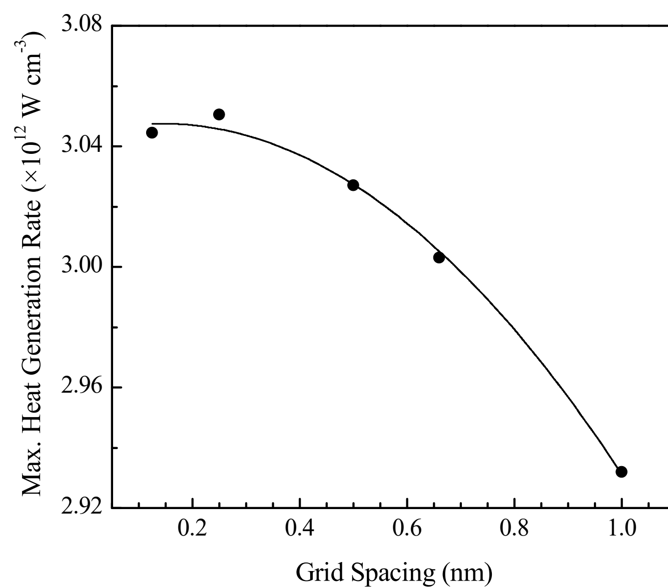

Figure 4.5: Variation of the maximum volumetric heat generation rate in a 10 nm channel n^+nn^+ device at 300 K with the grid spacing.

Figure 4.5 presents the sensitivity of the heat generation rate to the grid spacing for the same representative case and the selected time step size of 0.2 fs. A uniform grid spacing of 0.25 nm is selected for the present study, since the improvement in accuracy that results from selecting the 0.125 nm spacing is only 0.20%, compared to the improvement of 0.98% offered by choosing $\Delta x = 0.25$ nm over $\Delta x = 0.5$ nm. Furthermore, a smaller grid spacing requires a larger storage space to handle the information at the increased grid nodes, and imposes a more stringent requirement on the time step size in accordance with Eq. 4.2, which, in turn, increases the computation time. A larger grid spacing is not selected since the minimum characteristic feature length of the devices to be investigated in this study was anticipated to be as small as 2 nm, and an adequate number of grid nodes must be present to accurately capture the spatial distributions of the parameters of interest.

4.4 Validation

Prior to solving the problem of interest, simulations were performed to validate the performance characteristics of 1D quasi-ballistic n^+nn^+ devices. The results of these simulations, together with the validation performed in Chapter 3 for the carrier transport characteristics of bulk silicon at a range of operating conditions, help to establish a level of confidence in the adopted methodology by confirming that the geometry and mesh information and the carrier transport mechanisms are being handled correctly by the MC solver. The validation exercise also helps to tune simulation setup parameters such as the sub-history size, total simulation time and the number of simulated super-particles—all of which need to be carefully considered to ensure that the results of the stochastic MC simulation process are adequately converged.

Figure 4.6 (a) compares the results of the present work with the published dataset of the heat generation profiles at three different applied potentials for a 20 nm long channel n^+nn^+ device with 10^{20} cm^{-3} electrode doping density. Moreover, as mentioned earlier, the employed methodology accounts for the heat generation through the optical and acoustic phonons separately to accommodate the influence of their varying group velocities on the localized heat generation distribution. Therefore, validation for the contribution of both types of phonons has also been performed. Figure 4.6 (b) presents the volumetric heat generation by phonon type along the entire length of an n^+nn^+ device with a 500 nm long channel, wherein the device is subjected to a 2.5 V drain voltage (V_D) with an electrode doping density of 10^{18} cm^{-3}. The obtained trends in Figure 4.6 are in good agreement with the benchmark published values, with

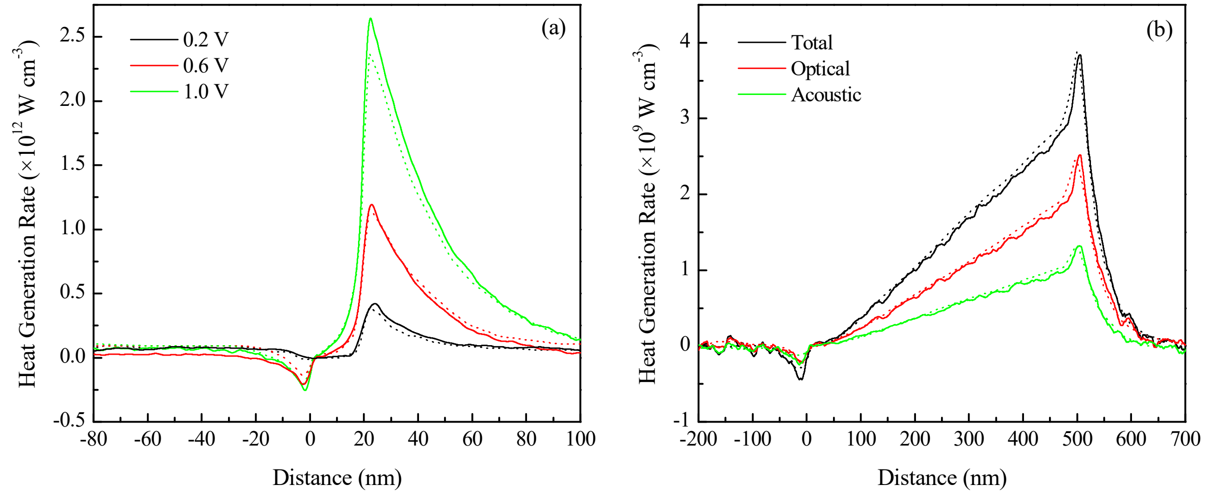

Figure 4.6: Validation of the volumetric heat generation rates at 300 K over the lengths of (a) 20 nm and (b) 500 nm channel n^+nn^+ silicon devices. Solid lines are the results of the present work, while dotted lines are the MC results obtained from Pop et al. [73].

only minor differences observed in the magnitudes of the heat generation. Errors of 13.2%, 5.1% and 11.8% are observed in Figure 4.6 (a) for the maximum heat generation rates at 0.2 V, 0.6 V and 1.0 V (close to the non-parabolic band approximation limit) respectively, while a maximum error of less than 6% is observed in the comparisons shown in Figure 4.6 (b). These deviations are within reasonable expectations due to the stochastic nature of the MC method.

4.5 Effect of Voltage

From the results reported in literature [73], [88], as well as from the results of the validation exercise presented in Figure 4.6 (a), it is evident that a variation in the potential difference applied across the device contacts has a pronounced effect on the volumetric heat generation rate in quasi-ballistic silicon devices. Using the validated methodology, additional simulations were thus performed on a shorter channel device to investigate whether the effects of the applied potential difference on the device performance change as a result of miniaturization into the ballistic transport regime. For this purpose, a 10 nm channel silicon device with 100 nm long electrodes doped to 10^{20} cm^{-3} is used. Because the implemented MC model uses the non-parabolic band approximation and neglects the role of impact ionization, with the source terminal kept grounded, the possible range of applied drain voltages to investigate is limited to $V_D \leq 1.1$ V (below the band gap energy of Si).

Figure 4.7 (a), 4.7 (b) and 4.7 (c) present the variations of the volumetric heat generation profile, the electron energy distribution, and the electron drift velocity distribution,

respectively, with the applied potential difference. The insets of these figures present the variations of the corresponding parameter maxima with the applied voltage; the maximum heat generation and the maximum electron energy are both found to increase linearly with the applied voltage, as evident from the relevant insets. Interestingly, this linear correlation is consistent with the results reported by Pop et al. [73] for a longer channel silicon device, which indicates that the effect of the applied voltage on device performance does not break trend upon device miniaturization into the ballistic transport regime, unlike for the other control parameters investigated in this study.

The distributions of the electric potential, electric field strength, and charge carrier concentration over the device geometry for the $V_D = 0.8$ V case were presented in Figure 4.3. The shape of the electric field strength distribution, which is observed to peak at the source-channel and the channel-drain contacts, remains similar for different V_D, but the magnitudes involved change. For a particular applied voltage, the electric field strength is found to have a maximum absolute value at the channel-drain interface ($x = 10$ nm), as in Figure 4.3 (b), due to the creation of a potential barrier as a result of contact between the mildly-doped channel region and the heavily-doped drain region. Despite the source having the same dopant concentration as the drain region, the magnitude of the electric field is smaller, and its direction opposite, at the source-channel interface ($x = 0$ nm), since the junction is biased in the opposite direction to the channel-drain junction for a particular applied potential difference.

Under the influence of the electric field, some of the charge carriers present in the source (which are originally in thermal equilibrium with the lattice) gain sufficient energy to surmount the interfacial potential barrier formed between the highly doped source and the mildly doped channel region. Due to the potential barrier, only the high energy electrons are able to pass through the interface, leaving behind lower energy electrons on the source side, which now have an average energy below the thermal equilibrium level. As a result, the electrons remaining on the source side in the vicinity of the potential barrier extract energy from the lattice through net phonon absorption to restore thermal equilibrium [80], effectively cooling the lattice. This phenomenon explains the region of negative heat generation observed near the end of the source ($x \approx 0$ nm) in Figure 4.7 (a), which is accompanied by a notable reduction in the electron energy, shown in Figure 4.7 (b), at the same location.

The electrons injected from the source into the channel are accelerated under the influence of the electric field in the mildly doped channel region. A higher potential difference applied across the device terminals causes the charge carriers to gain more energy as they

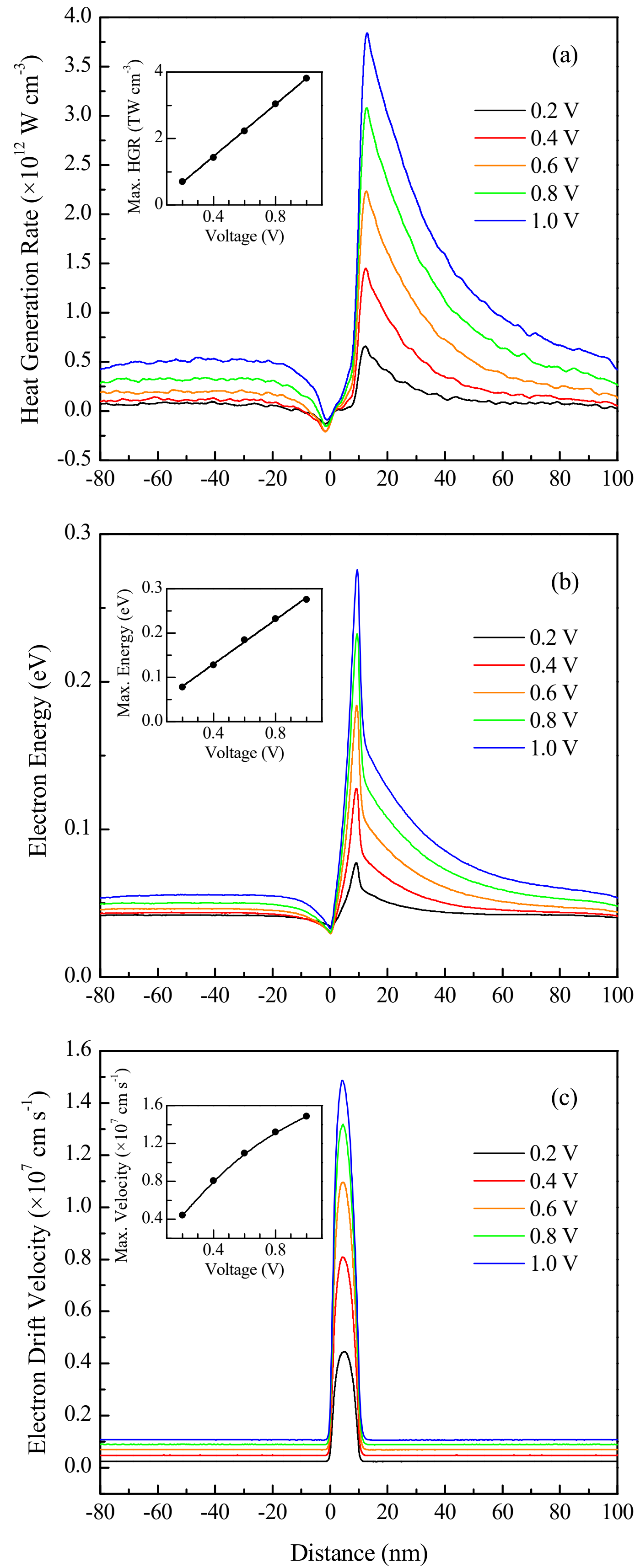

Figure 4.7: Effect of voltage on the distributions of (a) heat generation, (b) electron energy and (c) electron velocity over the length of a 10 nm channel n^+nn^+ silicon device at 300 K.

traverse the channel, which explains the increase in the average and maximum electron energies with increasing V_D, as observed in Figure 4.7 (b). Higher energy electrons undergo a greater number of inelastic scattering events with the lattice atoms, and therefore dissipate a larger amount of waste energy, which translates to the observed increase in the average and maximum heat generation rates with increasing V_D in Figure 4.7 (a).

A lower doping in the channel region, together with the existence of potential barriers at the source-channel and channel-drain interfaces, means that the mobile charge carrier density in the channel is limited in comparison to the source and drain, as seen in Figure 4.3 (c). In a geometrically continuous device, current conservation mandates that the electron velocity is higher in a region with low electron density, which explains why the drift velocity is maximum in the channel, as shown in Figure 4.7 (c). As aforementioned, the injection of lower-energy electrons from the drain into the channel causes the maximum electron drift velocity to be skewed towards the source end of the channel. Furthermore, in Figure 4.7 (c), the average and maximum electron drift velocities are found to increase at a decreasing rate as the applied voltage is increased. From Figure 4.3 (b), it can be noted that the magnitude of the maximum electric field strength at $V_D = 0.8$ V is approximately 9.77×10^5 V cm^{-1}. Upon reconsideration of the electrical properties of bulk silicon that were investigated in Chapter 3, Figure 3.3 indicates that for 10^{16} cm^{-3} doped silicon, this peak field strength value is well into the velocity saturation regime. For the considered range of applied voltages and shortened device dimensions, velocity saturation is therefore a significant concern, and is why an increase in the applied voltage is unable to cause a proportional increase in the maximum drift velocity.

The drain, which initially contains electrons that are in thermal equilibrium with the lattice, gets heated through scattering of the higher-energy electrons that are injected across the channel-drain interface due to the applied electric field. As discussed previously, the electric field peaks at this interface ($x = 10$ nm), causing electrons to gain a significant amount of energy in the vicinity of this location. These electrons subsequently lose their excess energy and return to thermal equilibrium via net phonon emission during collisions with the lattice atoms; the energy released in a single collision is constrained to a discrete amount dictated by the energy of the emitted phonon. Since the energized electrons must undergo multiple scattering events to release the amount of energy requisite for the restoration of thermal equilibrium, they must traverse a multiple of the inelastic electron-phonon mean free path—which has been shown by previous studies to be between 5–10 nm for the ranges of scattering times and average electron velocities involved in this case [73]—for this purpose. For ultra-

short devices like the one studied here, energy is therefore dissipated by the injected electrons over a significant portion of the drain length. Moreover, as the average electron temperature gets closer to the lattice temperature further along the length of the drain through scattering events, which is reflected in the trends of electron energy shown in Figure 4.7 (b), lower-energy phonons are released as a result of subsequent collisions, resulting in the decreasing rate of heat generation observed in Figure 4.7 (a). Because all other types of scattering events mainly influence only the electron momenta, and not their energies [2], only electron-phonon scattering contributes to the Joule heating of electronic devices.

4.6 Effect of Channel Length

A comparison of Figure 4.6 (a) and Figure 4.7 (a) indicates that a reduction in the channel length from 20 nm to 10 nm causes a significant increase in the heat generation rate for the same applied voltages. The effects of changing the channel length on the device performance were thus studied in greater detail by simulating devices with channel lengths varying from 2 nm to 120 nm. The minimum investigated channel length is limited by the scale constraints on the Boltzmann transport equation [60], which serves as the basis for the MC method. The devices are subjected to an applied V_D of 0.8 V and an initial temperature of 300 K, and have 100 nm long electrodes that are doped to 10^{20} cm^{-3}.

Figure 4.8 illustrates the effects of changing the channel length on the distributions of volumetric heat generation rate, electron energy and electron velocity. The reference point of the x-coordinate is taken at the start of the device in this figure. Note that a normalized distance $\left({}^{x}/_{x_{total}}\right)$ of 0.5 corresponds to the center of the devices, which coincides with the center of their channel, since the source and drain have equal lengths of 100 nm. From Figure 4.8 (a) and 4.8 (b), it is observed that the maximum heat generation and maximum electron energy for any given channel length occur along the channel-drain contact, consistent with the results for the 10 nm channel device shown in Figure 4.7. This emphasizes the need for careful drain design to ensure that thermal equilibrium between the charge carriers and the lattice atoms is restored before the electrons exit the device through the drain contact. Moreover, these results also support the design of a drain-centric heat sink to improve the durability of scaled devices. Meanwhile, the electron velocity, shown in Figure 4.8 (c), is found to be maximum near the source end of the channel region, which is also consistent with previous observations.

Figure 4.9 presents the variations of the parameter maxima with the device channel length. In Figure 4.9 (a), the contribution of the optical phonon mode to the heat generation

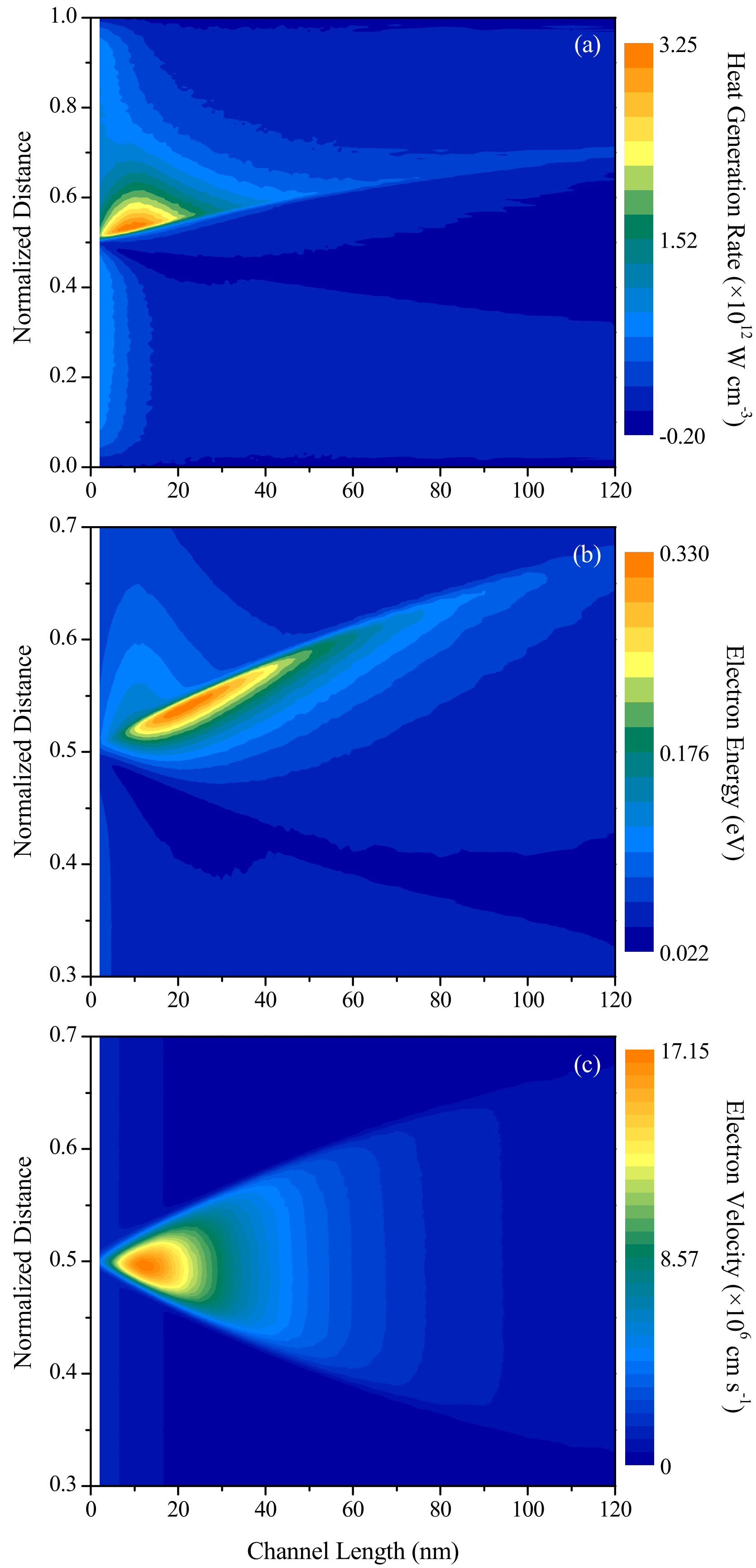

Figure 4.8: Variation of the spatial distributions of the (a) heat generation rate, (b) electron energy, and (c) electron velocity with the channel length of the n^+nn^+ silicon devices.

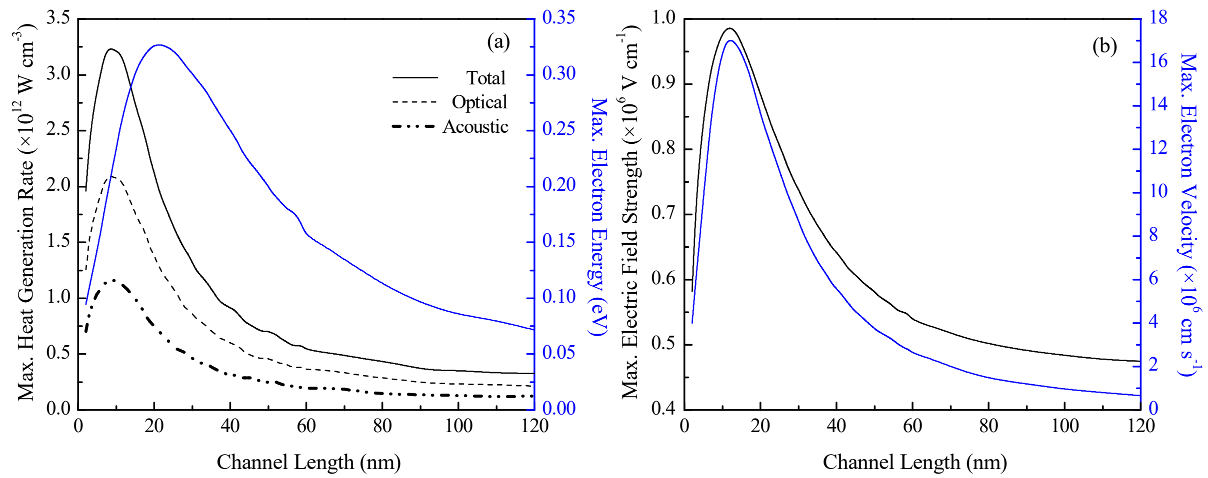

Figure 4.9: Variation of the device performance parameter maxima in an n^+nn^+ device with the channel length at 300 K. In (a), the separate contributions of the optical and acoustic phonon modes to the total heat generation rate are shown.

rate is observed to be almost twice as much as that of the acoustic mode at any given channel length. This behavior, also observed in the results for bulk silicon presented in Figure 3.7, is explained by the difference in the phonon group velocities [12], with the slower optical phonons contributing little to thermal conductivity and dissipating a greater portion of their energy to the lattice locally compared to their acoustic counterparts.

The maximum volumetric heat generation rate, maximum electron velocity and maximum electron energy are found to peak at channel lengths of 8 nm, 12 nm, and 20 nm respectively. As discussed previously, the inelastic electron-phonon scattering length is found to lie between 5 nm and 10 nm for the simulated conditions, considering the encountered range of maximum electron velocities together with the average scattering time reported in literature [73]. For channel lengths shorter than this scattering length, which is the length scale of primary interest in this study, the number of scattering events greatly decreases as ballistic transport starts to dominate. This causes the electrons to lose less energy to the lattice, and in turn, results in reduced heat generation. These findings also verify the observations of Rowlette et al. [128], who indicated the prevalence of near-ballistic conditions to be the likely reason behind the trends of their results for their 20 nm channel silicon n^+nn^+ device.

Figure 4.10 confirms that the reduction in the heat generation rate for channel lengths smaller than 8 nm is not due to a reduction in current, since the current density is observed to increase continuously as the channel length is reduced (the cross-sectional areas of the devices are constant). On the other hand, the reduction in both the dynamic and leakage power with

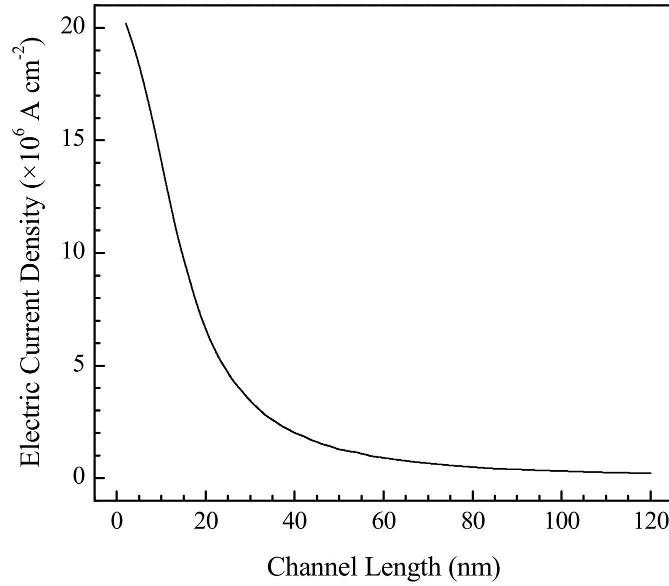

Figure 4.10: Effect of channel length on the electric current flowing through the n^+nn^+ device.

increasing channel length above 10 nm is a well-documented phenomenon [12], [27]. For the same applied V_D, a longer channel implies a weakened driving electric field, which subsequently results in a decrease in the current flowing through the device. The reduced current corresponds to a reduced rate of power dissipation, which also explains the observed peak in the maximum heat generation rate.

The maximum electric field in devices subjected to the same applied potential difference is found to decrease with a shortening channel below 12 nm, as shown in Figure 4.9 (b). This reduction in the maximum electric field (which occurs at the channel-drain interface, as aforementioned) is accompanied by an increase in the electric field strength across the electrodes, which directly affects the maximum electron velocity, and in turn, the electric current. Since the maximum electric field strength is reduced, electrons undergo a smaller amount of acceleration in the channel, thereby resulting in a coinciding peak in electron velocity, also shown in Figure 4.9 (b). The increase in the electric field strength across the electrodes is attributed to the direct source-drain tunneling that results from an increasingly ballistic nature of carrier transport as device dimensions are downscaled, and can potentially result in thermal runaway and breakdown. Confinement effects serve to increase the threshold voltage and contribute to a reduction in thermal conductivity, while the tunneling phenomenon gives rise to a large leakage current, which can ultimately cause an increase in power dissipation [55]. The reduction in the heat generation rate for channel lengths below 8 nm does not, therefore, guarantee an improvement in device performance.

4.7 Effect of Electrode Lengths

As aforementioned, it is observed that the region of heat generation in the 1D devices investigated thus far extends well into the drain. Furthermore, a change in the channel length is found to have a pronounced influence on the device performance parameters through a variation of the electric field distribution in the device for the same applied potential difference. Therefore, the effects of changing the electrode lengths on the device performance were also studied. Simulations were performed on 20 nm channel devices subjected to an applied V_D of 0.8 V, with equally long source and drain regions doped to 10^{20} cm^{-3}. The chosen channel length (20 nm) corresponds to the length at which the electron energy was observed to be maximal in Figure 4.9 (a), since it was anticipated that the primary influence of a reduction in the electrode lengths would be to hamper the ability of the energized charge carriers to completely dissipate their excess energy and return to thermal equilibrium. The heat generation rate and electron energy profiles over the lengths of the devices are presented in Figure 4.11.

As shown in Figure 4.11 (a) and 4.11 (b) respectively, the heat generation rate and the electron energy in the electrodes are both higher for the devices with longer source and drain regions. Elongating the electrodes results in a slightly larger potential drop across them for the same applied potential—causing a small increase in the electron energy—while also increasing the number of scattering events, therefore resulting in increased heat dissipation. For the shorter devices (20 to 40 nm electrode length), the heat generation trends indicate that the charge carriers injected across the channel-drain interface are unable to completely dissipate their

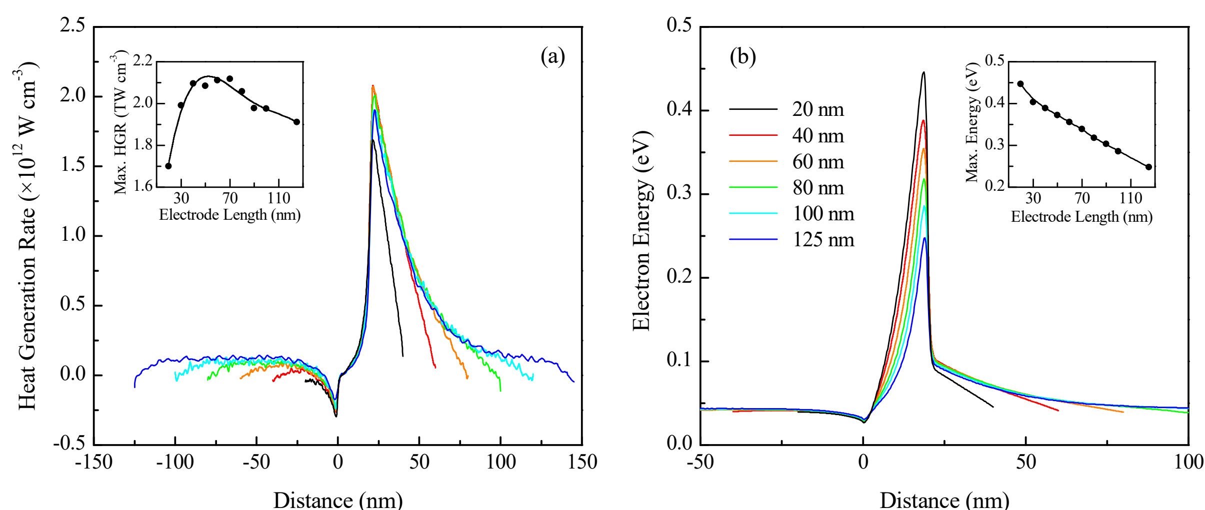

Figure 4.11: Effect of electrode length on the (a) heat generation rate and (b) electron energy over the length of a 20 nm channel device at 300 K (legends identical).

excess energy to the lattice in the drain region. Since the electrons can only lose energy in discrete quantities via phonon emission during scattering events, it is important that the drain length and device contacts be adequately designed to ensure safe and efficient operation. These findings are consistent with those of previously reported studies [54], [73], in which it has been noted that the heat generation in mesoscopic devices is expected to be significant along the device contacts instead of in the active device region.

With increasing electrode lengths, the maximum heat generation rate, which occurs in the channel and is shown in the inset of Figure 4.11 (a), is found to increase significantly for the shorter electrode devices, followed by a relatively gradual decrease. The inset of Figure 4.11 (b), meanwhile, indicates a decreasing maximum electron energy for increasing electrode lengths, with the trend being linear for electrodes longer than 30 nm. Longer electrodes result in a weaker electric field across the channel region for the same applied V_D, causing the charge carriers to undergo less acceleration across the intrinsically doped channel. This subsequently causes a reduction in the maximum electron energy and maximum heat generation rate, both of which are observed at the drain end of the channel. Despite the reduced maximum heat generation at smaller electrode lengths, the charge carriers do not come into thermal equilibrium with the semiconductor lattice prior to leaving the device, and the excess energy is dissipated at the critical device junctions, which can result in inefficient operation and thermal breakdown. These findings, therefore, emphasize the need for adequately long electrodes to ensure that thermal equilibrium is achieved within the drain region.

4.8 Effect of Temperature

In Chapter 3, the results of simulations performed on bulk silicon at various operating conditions indicated that the initial lattice temperature has a significant influence on the electron drift velocity, electron drift mobility and the volumetric energy dissipation rate in the semiconductor, as shown in Figure 3.2, Figure 3.4, Figure 3.5, and Figure 3.9. The consequences of a variation in the initial lattice temperature on the performance parameters of n^+nn^+ devices were therefore investigated by performing simulations on a 10 nm channel device subjected to $V_D = 0.8$ V with 150 nm long source and drain regions doped to 10^{20} cm^{-3}.

Figure 4.12 illustrates the effects of temperature on the distributions of electron energy and drift velocity over a section of the device length, while Figure 4.13 illustrates the variation with temperature of the parameter maxima evaluated over the entire device length. From Figure 4.12 (a), an increase in the average electron energy over the length of the device with the

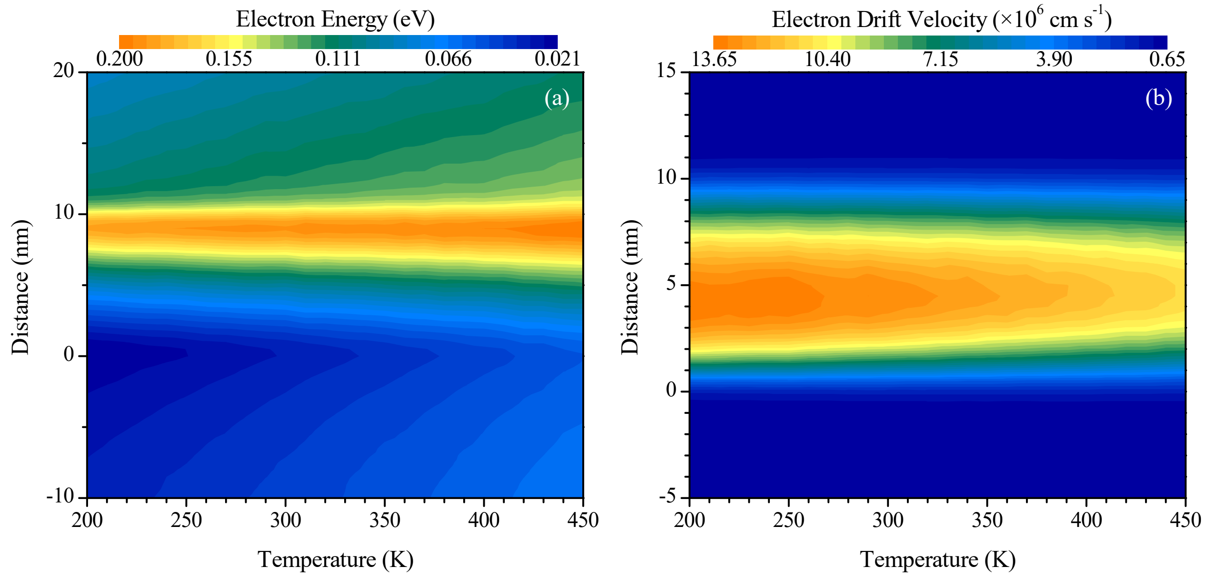

Figure 4.12: Variation of the device performance parameter distributions over a section of the total device length with the initial temperature.

temperature is evident, while Figure 4.12 (b) indicates a decrease in the average electron drift velocity as the initial lattice temperature is raised. Please note that Figure 4.12 shows the parameter distributions only in the vicinity of the channel, which is the region of significant changes in the distributions due to the presence of the peaked electric fields at either end of the channel (channel ranges from $0 \leq x \leq 10$ nm). It is also observed that, consistent with the results previously shown in Figure 4.8, the maximum electron energy occurs near the drain end of the channel, while the maximum drift velocity occurs closer to the source end of the channel for the range of temperatures investigated in the present study.

Figure 4.13 (a) indicates that no significant correlation exists between the temperature and the maximum heat generation rate in the device. Considering the range of the maximum field strengths that occur in the devices for the applied V_D and investigated range of temperatures, as shown in Figure 4.13 (d), it is found that the weakest field strength distribution (at $T = 200$ K) has a maximum field strength value of ~ 8.45×10^5 V cm^{-1}. From the results presented in Figure 3.9, it is clear that the effect of temperature on the heat generation rate in bulk silicon decreases in significance at higher applied electric field strengths, and that the maximum field strengths encountered in the investigated devices are much higher than those discussed previously for bulk silicon. The relative insensitivity of the rate of heat generation in the devices to the initial lattice temperature is thus anticipated for the range of conditions presently investigated, and can be attributed to the effects of increasing temperature on the

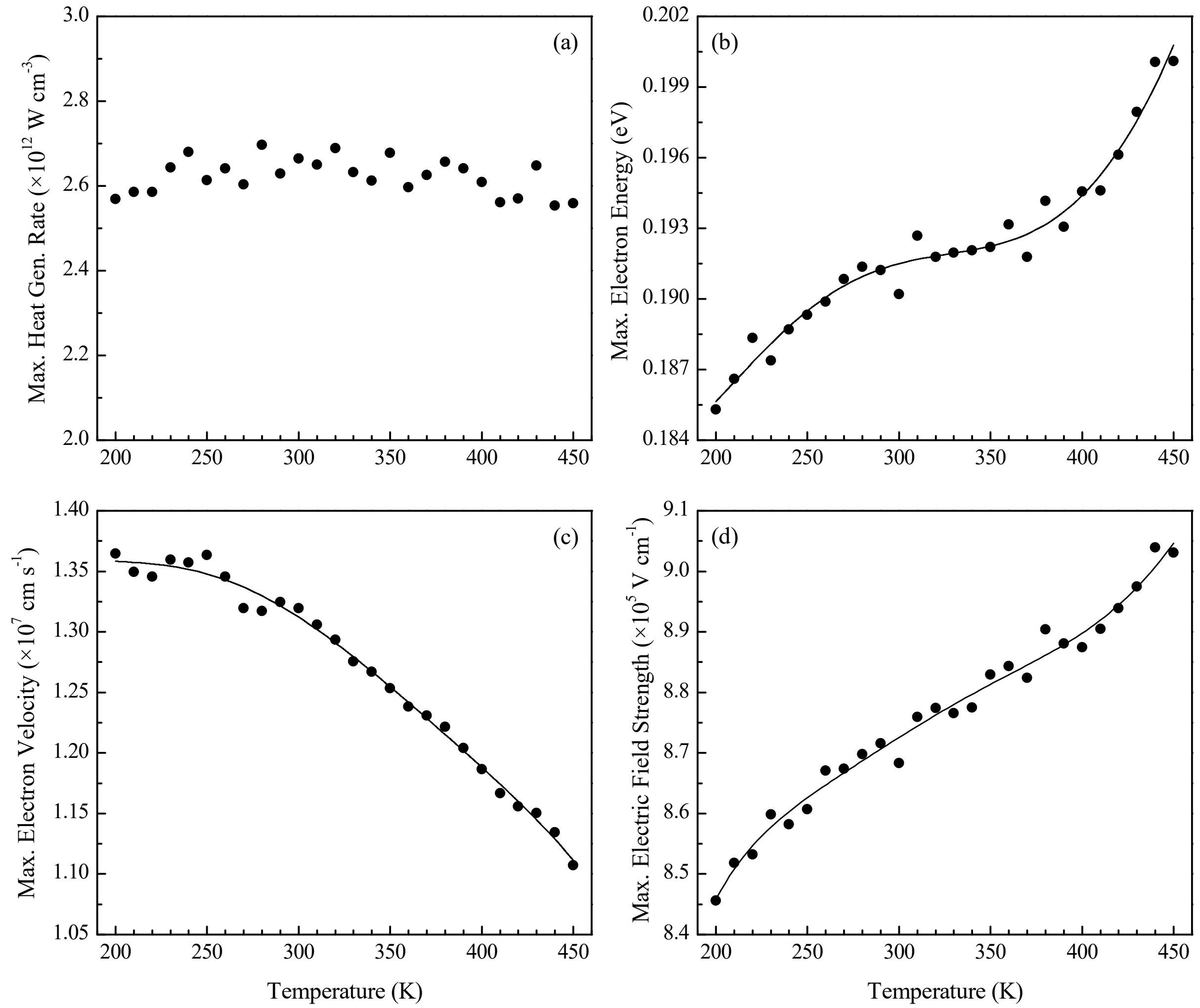

Figure 4.13: Effect of initial temperature on the maximum (a) volumetric heat generation rate, (b) electron energy, (c) electron drift velocity and (d) electric field strength in a 10 nm channel n^+nn^+ device.

maximum electron energy (found to increase, as indicated by Figure 4.13 (b)), maximum electron drift velocity (found to decrease, as shown in Figure 4.13 (c)), and electric current.

The maximum (and average) electron energy increases with increasing temperature since temperature is directly linked to the internal energy of the lattice, with a higher temperature involving promotion of electrons to higher energy bands [17]. A higher average electron energy subsequently leads to an increase in heat generation, since a larger amount of energy can be dissipated to the lattice through inelastic electron-phonon scattering events. The maximum (and average) drift velocity, on the other hand, decreases due to the increased vibrational motion of the lattice atoms, which impedes the motion of the charge carriers and results in earlier velocity saturation [42], [110]. This behavior (also observed for bulk silicon in Figure 3.2), in turn, tends to reduce the current flowing through the device for the same applied potential difference.

The minimum electron density in the device, which occurs in the channel region due to the doping profile, is found to increase continuously with an increase in temperature, as shown in Figure 4.14 (a). This is because the increased average electron energy enables a greater number of charge carriers to overcome the potential barrier formed at the source-channel interface, thereby increasing the number of mobile charge carriers in the channel region. The current density, calculated as in Eq. 1.12, depends on both the electron density and drift velocity; the trend observed in Figure 4.14 (b) can thus be explained in light of Figure 4.13 (c) and Figure 4.14 (a) for the investigated range of temperatures. Since, with an increasing temperature, the electron velocity decreases at an increasing rate, whereas the electron density increases almost linearly throughout, the electric current density initially increases at a decreasing rate. For $T > 290$ K, however, the rate of decrease of the electron velocity outweighs the rate of increase of the electron density, and, as a consequence, the electric current density starts to decrease at an increasing rate. The overall effect of a higher initial temperature is thus a decreased rate of flow of higher energy carriers, which culminates in a negligible net effect on the heat generation rate for the range of temperatures considered.

It is interesting to note that the variation of the maximum drift velocity with temperature, as observed through the present simulations on short channel silicon devices, is phenomenologically consistent with the effect of temperature on ZnO diodes with a 250 nm long channel, as reported by Adeleh and Reza [88]. However, in their investigations, the maximum electron energy was found to decrease slightly with an increase in temperature, unlike the present findings. This was despite the average electron energy being higher at higher

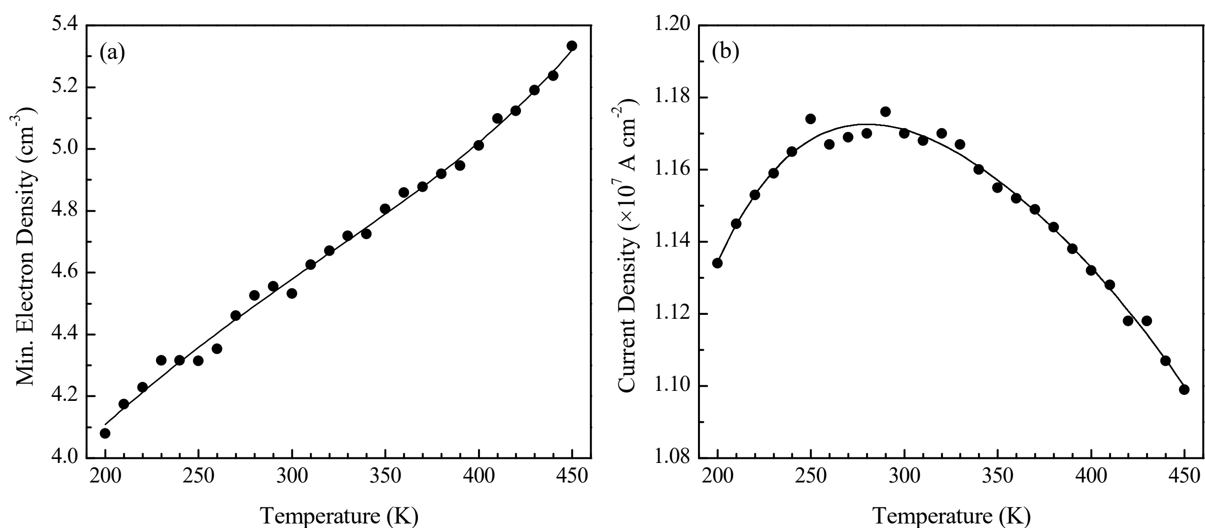

Figure 4.14: Effect of temperature on the (a) minimum electron density in a 10 nm channel n^+nn^+ device and (b) density of the electric current flowing through the device.

temperatures, similar to the present results shown in Figure 4.12 (a). This difference is best explained by the significantly different band structures and band gap energies (3.4 eV for ZnO [88], compared to 1.1 eV for silicon [110]) of the two materials.

4.9 Effect of Electrode Doping Density

The doping density directly influences both the number of charge carriers and the number of impurity atoms present in a semiconductor. The effects of the resulting changes in the mobile charge carrier density and ionized impurity scattering probability on the electron drift velocity, electron drift mobility and volumetric heat generation in bulk silicon have been presented previously in Figure 3.3, Figure 3.5 and Figure 3.8. Furthermore, it is evident from the previous discussion of the characteristics of n^+nn^+ devices that the device performance is determined primarily by the properties of the electrons that are injected from the highly doped source into the mildly doped channel when an electric field is applied across the device contacts. The impact of changing the electrode doping density on the device performance is thus investigated by performing simulations on 10 nm channel silicon devices with 150 nm long source and drain regions doped equally to various levels. The devices, which are subjected to $V_D = 0.8$ V, have an initial temperature of 300 K. The electrode doping levels for the simulations range from 10^{16} cm^{-3} (equal to the channel doping) to 5×10^{21} cm^{-3}.

Figure 4.15 shows the effects of the electrode doping density on the different device parameter maxima. Please note that these plots are shown on log-log and semi-log scales. Figure 4.15 (a) indicates that the maximum rate of volumetric heat generation in the device increases continuously with increasing electrode doping, which is consistent with the trend for bulk silicon previously observed in Figure 3.8, and occurs due to an increased scattering of electrons with ionized impurities, since an increased dopant concentration means an increase in the number of irregularly-sized impurity atoms in the silicon lattice.

An initial gradual decrease in electron energy is observed with increasing electrode doping at low electrode doping levels, followed by a gradual increase that becomes sharper as the doping level becomes more degenerate, as shown in Figure 4.15 (b). Figure 4.15 (c), meanwhile, indicates that the magnitude of the maximum electric field strength encountered over the device length increases as the electrode doping density is increased. This is because at an electrode doping level equal to the channel doping (10^{16} cm^{-3}), the potential barrier between the channel and the electrode region is nonexistent, with the device essentially being one continuous semiconducting element. As the electrodes become more heavily doped, however,

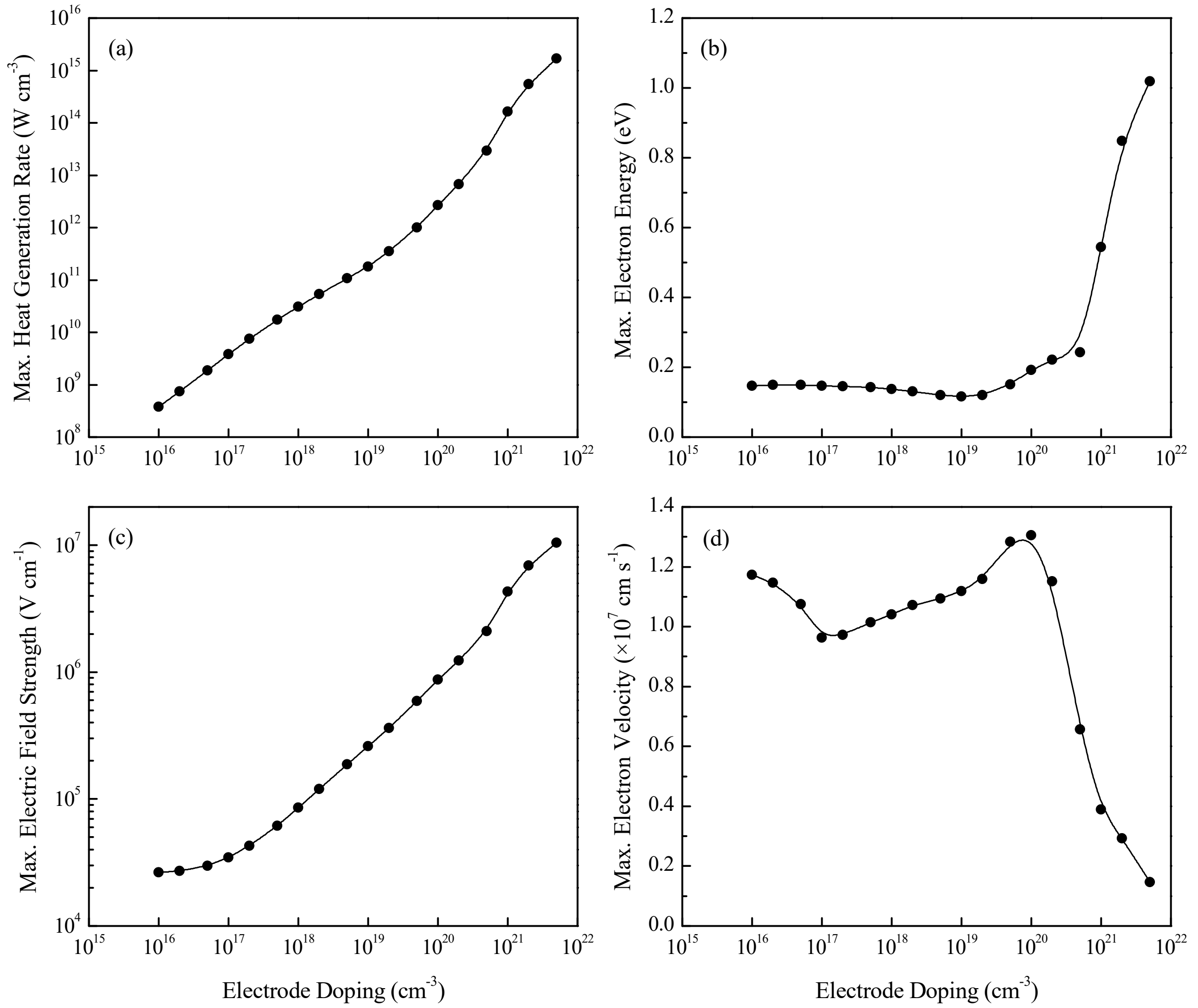

Figure 4.15: Effect of electrode doping on the maximum (a) volumetric heat generation rate, (b) electron energy, (c) electric field strength and (d) electron drift velocity in a 10 nm channel n^+nn^+ device.

the potential barrier increases in strength as the depletion region develops and widens, resulting in a highly peaked electric field at the channel-drain interface.

The carrier drift velocity and mobility in bulk silicon are known to decrease with increasing doping levels due to enhanced carrier scattering [129], as also indicated by the results of Figure 3.3 and Figure 3.5. This explains the reduction in the maximum drift velocity at high electrode doping levels ($> 2\times10^{20}$ cm^{-3}) observed in Figure 4.15 (d); at such near-degenerate electrode doping levels, the number of mobile charge carriers injected into the intrinsically doped channel is also very high (which can be noted from Figure 4.16 (a)), which, in turn, results in acceleration of the increased number of charge carriers to a slower average velocity through the channel. At low doping levels, meanwhile, the drift velocity is high as

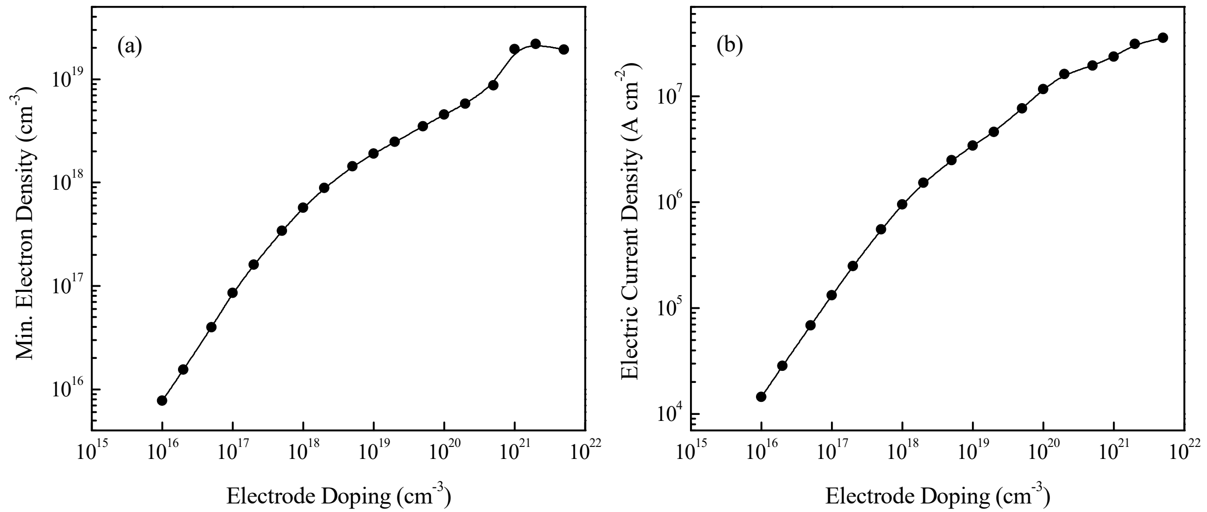

Figure 4.16: Effect of electrode doping on the (a) minimum electron density in a 10 nm channel n^+nn^+ device and (b) density of the electric current flowing through the device.

there is no potential barrier to overcome, and the velocity decreases as the barrier develops. The drift velocity then increases again at moderate doping levels, since the carriers that overcome the barrier at the source-channel interface undergo a greater acceleration due to a sharper reduction in the conduction band energy level in the channel.

The relevance of the electrode doping density to the concentration of available mobile charge carriers in the device is made evident by Figure 4.16 (a), in which a clear increase in the minimum electron density (which occurs in the channel due to the constant 10^{16} cm^{-3} doping applied in that region, as opposed to the equal or higher doping in the electrodes) with increasing electrode doping is observed. Even at very high doping levels, the increase in the minimum mobile charge carrier concentration that results from an increase in the electrode doping is found to outweigh the reduction in the carrier drift velocity that occurs due to increased ionized impurity scattering (which is also somewhat mitigated by the strengthening of the electric field, as discussed previously). Together, this culminates in a continuous increase in the electric current density with increasing electrode doping, as shown in Figure 4.16 (b).

4.10 Two-Dimensional Simulations

Having investigated the influence of the various control parameters on the performance of 1D nanoscale n^+nn^+ silicon devices—which serve as a realistic representation of the cross section of a MOSFET channel and have been frequently used by previous researchers to study the effects of changing various parameters on device performance while minimizing computational costs—and thereby having fulfilled the research objectives of this study, as set

out in Section 1.7, efforts were made to extend the present investigations to a two-dimensional MOSFET geometry. The MC solver developed by Pop et al. [111], which was used to obtain the results presented previously, does not incorporate a Poisson equation solver for the 2D mode, unlike for 1D simulations. This means that the electric field distribution over the 2D device geometry is not updated with the movement of the charge carriers, so that the entire simulation is performed using the initial field distribution information imported from the drift-diffusion solver, which significantly reduces the accuracy advantage offered by the Monte Carlo method over the drift-diffusion approach [98].

A simplistic two-dimensional rectangular transistor geometry was created in the semiconductor module of the COMSOL Multiphysics software for the purpose of importing the subsequent steady-state simulation results as initial conditions into the MC solver. This transistor, with an initial temperature of 300 K, had a channel length of 20 nm, electrode lengths of 50 nm, electrode thicknesses of 5 nm, total height of 30 nm and total width of 120 nm. The channel had an n-type doping of 10^{16} cm^{-3}, while the source and drain were subjected to n-type doping with a donor concentration of 10^{20} cm^{-3}. A drain voltage of 0.8 V and a gate voltage of 1.5 V were applied through the appropriate ideal ohmic metal contacts, while the source and the substrate were grounded, and a silicon dioxide gate was used.

Figure 4.17 presents the steady state distribution of the electrons in the transistor, as determined through the COMSOL drift-diffusion simulation. It can be noted that the channel, which extends from $50 \leq x \leq 70$ nm (the start of the source is taken as the reference point of the x-coordinate here), is properly developed for the applied terminal voltages, since a

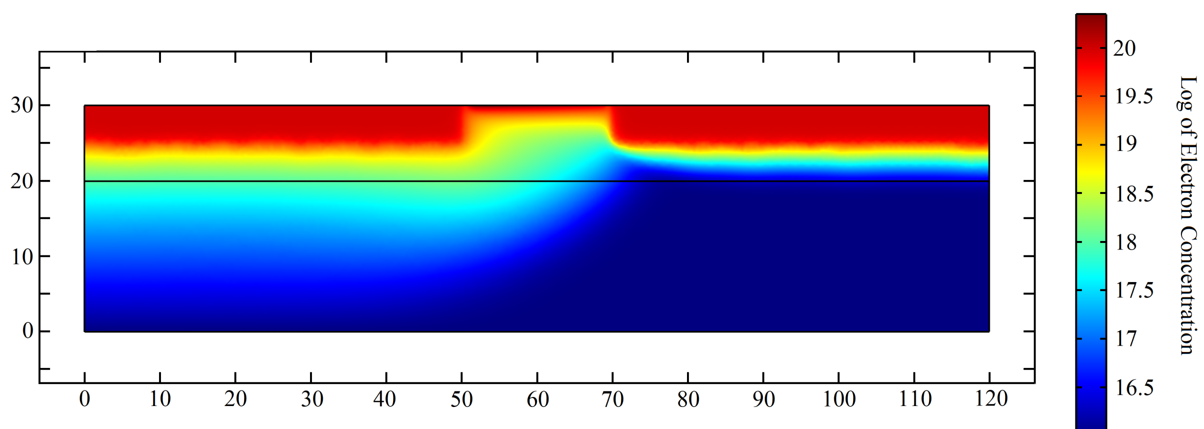

Figure 4.17: Steady-state distribution of electrons in the two-dimensional transistor geometry. The x- and y-axes represent the horizontal and vertical distance, respectively, from the bottom left corner of the transistor in nanometer units.

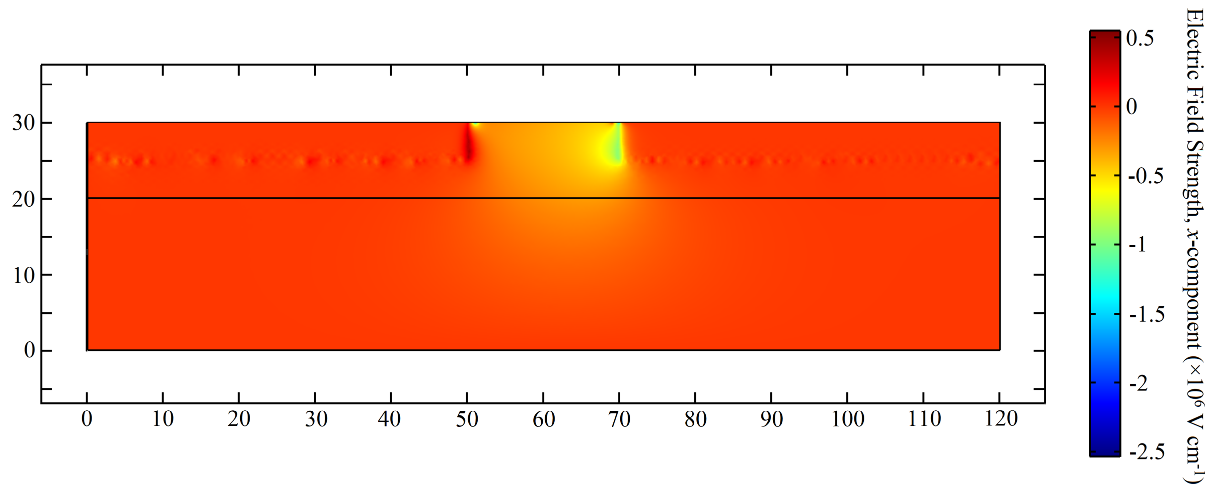

Figure 4.18: Steady-state distribution of the x-component of the electric field strength over the two-dimensional transistor geometry. The x- and y-axes represent the horizontal and vertical distance, respectively, from the bottom left corner of the transistor in nanometer units.

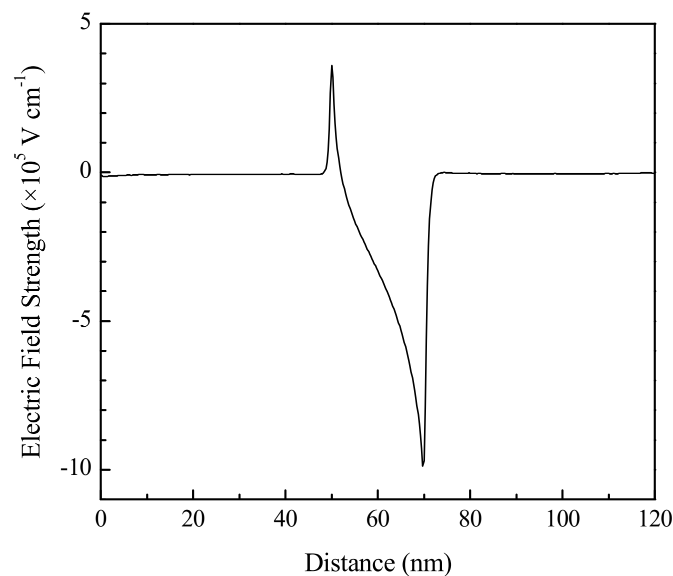

Figure 4.19: Steady-state distribution of the electric field strength over the length of a 1D n^+nn^+ device subjected to an applied potential difference of 0.8 V with a 20 nm long channel and 50 nm long electrodes doped to 10^{20} cm^{-3}, as determined through a Monte Carlo simulation.

significant charge carrier population is present in this lightly-doped region and channel pinching is not observable. A measurable forward current thus flows through the channel. Figure 4.18 presents the distribution of the electric field strength in the x-direction over the device geometry, and peaks in the field strength magnitude can be observed at the interfaces between the channel and the electrodes. Significantly, this distribution in the 2D device geometry is consistent with the electric field distribution determined through a MC simulation

for the equivalent one-dimensional representation of this geometry, which is presented in Figure 4.19. A positive peak in the field strength is observed at the source-channel interface while a stronger negative peak is observed at the channel-drain interface for the 2D geometry, and the magnitudes of both peaks are very similar to those determined for the n^+nn^+ device.

In a manner similar to the procedure adopted earlier for the 1D simulations, the steady-state distributions of the electric field strength, dopant concentration, electric potential and charge carrier concentration were imported into the MC solver through a text file. The computational cost associated with performing a single simulation through the MC method was, however, found to increase significantly for the two-dimensional geometry in comparison to its one-dimensional counterpart, even though the Poisson equation was not solved and the electric field distribution was therefore not updated self-consistently during the 2D simulation. Having taken into consideration the increased computational time associated with the 2D simulations, that the non-self-consistent two-dimensional MC solver does not offer a significant advantage in accuracy over the much faster drift-diffusion method, and that the 1D simulation results presented earlier provide sufficient and accurate information about the variation of the device performance with control parameters of interest in the critical device regions, it was decided that conducting an investigation into the properties of semiconductor devices operating in the ballistic transport regime using 2D MC simulations with the employed solver would not be justified by the addition of meaningful information to the results already obtained for the 1D n^+nn^+ devices, at least unless a self-consistent Poisson equation solver was incorporated.

4.11 Summary

In this chapter, the effects of multiple device parameters—applied potential difference, channel length, electrode lengths, electrode doping and initial lattice temperature—that were identified to be potentially critical to the performance of nanoscale devices, particularly as device dimensions are downscaled into the ballistic transport regime, were investigated using an established self-consistent ensemble Monte Carlo simulation model. The simulations were performed on one-dimensional n^+nn^+ silicon devices that served as a simplified representation of the cross-section of a MOSFET channel. An attempt to extend the MC simulations to two-dimensional geometries was also made, but limitations in the available computational resources, a very minor improvement in accuracy over the significantly faster drift-diffusion

approach, and insignificant additional information in comparison to the 1D simulations meant that 2D MC simulations were not included in the scope of the present work.

The maximum volumetric heat generation rate and electron energy in the devices were both found to vary linearly with the applied voltage, whereas mutually offset maxima were observed due to a variation of the channel length because of a transition of the characteristic feature size into the ballistic carrier transport regime. These maxima indicate a significant departure from the performance trends known for longer devices operating in the quasi-ballistic transport regime. Furthermore, results for the impact of the electrode lengths were found to be primarily dependent upon the variation of the electric field profile that develops across the three device regions with the electrode length, which emphasizes the necessity of careful drain and contact design in ultra-short devices since most of the heat dissipation occurs in the drain region. An increase in the initial lattice temperature did not have a significant effect on the heat generation rate due to the counteracting contributions of its effects on carrier energy and velocity. The heat generation rate was found, however, to increase markedly with an increase in the electrode doping.

CHAPTER 5: CONCLUSION

5.1 Summary

In this thesis, the ensemble Monte Carlo method has been used to perform numerical simulations of the transport of charge carriers in nanoscale electronic devices. The importance of minimizing Joule heating in order to improve the performance of existing devices, and particularly to allow for the continued miniaturization of modern electronic components without sacrificing device reliability and efficiency, was highlighted in Chapter 1. An explanation of the physical phenomena at the atomic scale that result in, and regulate, heat dissipation during carrier transport under the influence of a driving electric field was also presented. Using this knowledge, several parameters were identified that could potentially have a significant influence on electrical and thermal conduction in the device structure. The different well-established techniques that are used to study Joule heating and its influence on device performance at various length scales and with varying detail were also compared.

Of the reviewed methods, the numerical Monte Carlo technique was selected for the present study after considering the availability (or lack thereof) of the experimental and computational resources necessary to achieve the planned research objectives which would suitably address gaps in the existing knowledge of the properties of ultra-short electronic devices. Since the development of a new MC solver did not lie in the scope of the present study, a thorough review of literature was performed to compare the various MC models developed by past researchers—which incorporated the electron band structure, phonon dispersion relationship, and scattering mechanisms with varying levels of detail—on the basis of accuracy and computational costs. As described in Chapter 2, the employed methodology tracked a large number of super-particles, each representing billions of real electrons, as they moved through the geometry of the simulated device over time under the influence of an externally applied electric field. To quantify the rate of heat generation, the number of phonons emitted and absorbed as a result of each collision of each super-particle with the defects, impurities, and atoms in the semiconductor lattice, weighted according to the energies and classified by the types of the phonons involved, was aggregated. Analytical descriptions were used for both the phonon dispersion relationship and the electron band structure. Moreover, the Poisson equation was solved at regular intervals during the simulation to self-consistently update the electric field distribution with the motion of the charge carriers.

The selected Monte Carlo model was first used to investigate the properties of bulk silicon, which was modelled as an infinite resistor and subjected to varying initial temperatures, applied electric field strengths, and doping densities, as discussed in Chapter 3. This exercise successfully validated the employed methodology against the widely-reported characteristics for the bulk semiconductor, and provided additional data regarding the effects of these parameters on the electron drift velocity, electron drift mobility, energy of emitted phonons, and heat generation rate in bulk silicon, which served as a basis for understanding the results of the subsequent investigation of realistic silicon device geometries. The rate of Joule heating in bulk silicon was found to increase with an increasing applied electric field strength and doping density, but was found to decrease with an increasing initial lattice temperature over the range of conditions investigated. Furthermore, it was found that the optical phonon modes, which lie at the higher energy end of the phonon generation spectrum for silicon and have lower group velocities than acoustic phonons, contribute primarily to heat generation in the semiconductor. Saturation of the carrier drift velocity was also observed at sufficiently strong electric fields, and the unsaturated drift velocity was found to be adversely impacted by increasing temperatures and doping densities.

The same methodology was then extended to the simulation of nanoscale one-dimensional n^+nn^+ devices, which have been used as a simplified model of the cross section of a MOSFET channel, thereby allowing for the reasonably accurate investigation of the device performance parameters in the critical device region at a significantly lower computational cost in comparison to full three-dimensional simulations. Given the trend of aggressive miniaturization of electronic devices, the focus of this work was on investigating the impact of a potential crossover of the carrier transport mechanism into the ballistic regime, and to compare the behavior of the ultra-short devices with that of larger devices operating in the quasi-ballistic transport regime. The maximum volumetric heat generation rate in the 1D devices was found to increase linearly with the applied voltage, and to increase significantly with an increase in the electrode doping density, but was found to be relatively insensitive to a change in the initial temperature of the device, as discussed in Chapter 4. Moreover, in a notable change of behavior from longer devices that results from a marked reduction in the number of scattering events due to an increasingly ballistic nature of carrier transport, the maximum rate of heat generation was found to decrease for shortening channel lengths below 8 nm. However, from a perspective of device design, it is essential to consider the trade-off between a reduction in the maximum heat generation rate and a simultaneous reduction in the

thermal conductivity, together with an increase in the leakage current that occurs as device dimensions are scaled down into the ballistic transport regime, which may lead to inefficient operation and thermal runaway.

5.2 Future Work

From the spatial distributions of the heat generation over the lengths of the nanoscale 1D devices studied in this work, it is evident that most of the energy dissipation occurs in the drain region of these devices, and may also occur in the device contacts if the electrodes are not sufficiently long to allow for thermal equilibrium to be restored between the energy carriers and the device lattice. The results therefore emphasize the necessity of careful drain electrode and contact design for ultra-short devices. Furthermore, given the spatial constraints that are compounded by miniaturization and a subsequent increase in the on-chip transistor density, the incorporation of a heat sink in the vicinity of the drain should be of preferential importance in future device designs. The understanding developed by this work of the changes in the intensities and spatial locations of thermal hotspots that are caused by modifications to the device geometry thus has significance for the future design of chip cooling systems and technologies.

The Monte Carlo solver employed in the present work used an analytical, non-parabolic band approximation for the electron band structure and an analytical description for the phonon dispersion relationship. While these approximations help to reduce the computation time, they also limit the maximum range of applied voltages for which simulations can be performed, particularly since impact ionization—which is especially significant at high electric field strengths—was also neglected. The silicon band gap energy represents the maximum limit of the applicable range of voltages for this solver, which should nevertheless be sufficient for the investigation of modern nanoscale electronic devices, since these require low voltages to operate. If permitted by the availability of computational resources, the use of full band models and the incorporation of additional scattering mechanisms (such as electron-electron scattering) would improve the accuracy of the results, and allow for the investigation to be extended to high-field applications. The methodology used in this study could also be used to investigate the properties of electronic devices made of different semiconducting materials of interest, such as ZnO and GaAs, by appropriately modifying the mathematical relationships used to model the electron band structure and phonon dispersion of the material. Such a study could allow for comparisons to be drawn with the present results for silicon devices at varying length scales

and operating conditions and thereby potentially culminate in the identification of a range of conditions were one semiconductor has better properties than the other, and vice versa.

An attempt to extend the present work to simple two-dimensional MOSFET geometries was also made, but the investigation was confined to 1D devices due to the lack of a 2D Poisson equation module to self-consistently update the electric field distribution in the employed methodology, which meant that accurate quantitative results could not be obtained. Moreover, the limited availability of the computational resources needed to perform the significantly more demanding 2D numerical simulations meant that such a solver could not be incorporated in the MC solver during the course of this work. Although the 1D simplification is reasonably accurate and has been frequently used previously in literature for the numerical investigation of the properties of electronic devices, such a simplification does not account for degenerate carrier statistics, for the vertical quantum confinement effects that can be encountered in three-dimensional nanoscale MOSFETs, or for carrier transport through more complicated geometric configurations like multi-gate and raised-electrode transistors. An extension of the present work to more complex device geometries could therefore provide additional information of greater significance to the future miniaturization and development of electronic devices.

REFERENCES

[1] Z.-Y. Ong and M.-H. Bae, “Energy dissipation in van der Waals 2D devices,” *2D Mater.*, vol. 6, no. 3, p. 032005, Jun. 2019.

[2] E. Pop, S. Sinha, and K. E. Goodson, “Heat generation and transport in nanometer-scale transistors,” *Proc. IEEE*, vol. 94, no. 8, pp. 1587–1601, Aug. 2006.

[3] J. Choi and M. Jeong, “Compact, lightweight, and highly efficient circular heat sink design for high-end PCs,” *Appl. Therm. Eng.*, vol. 92, pp. 162–171, Jan. 2016.

[4] S. Sinha and K. E. Goodson, “Review: Multiscale thermal modeling in nanoelectronics,” *Int. J. Multiscale Comput. Eng.*, vol. 3, no. 1, pp. 107–133, 2005.

[5] G. E. Moore, “Cramming more components onto integrated circuits,” *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jan. 1998.

[6] G. E. Moore, “Progress in digital integrated electronics,” *IEEE Solid-State Circuits Soc. Newsl.*, vol. 11, no. 3, pp. 36–37, Sep. 2006.

[7] “Intel® product specifications.” [Online]. Available: https://ark.intel.com. [Accessed: 14-Feb-2020].

[8] K. Rupp, “42 Years of microprocessor trend data,” 2018. [Online]. Available: https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/. [Accessed: 28-Mar-2020].

[9] J. G. Koomey, S. Berard, M. Sanchez, and H. Wong, “Implications of historical trends in the electrical efficiency of computing,” *IEEE Ann. Hist. Comput.*, vol. 33, no. 3, pp. 46–54, Jul. 2011.

[10] R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. R. Leblanc, “Design of ion-implanted MOSFET’s with very small physical dimensions,” *IEEE J. Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, Oct. 1974.

[11] J. M. P. Cardoso, J. G. F. Coutinho, and P. C. Diniz, “High-performance embedded computing,” in *Embedded Computing for High Performance*, Elsevier, 2017, pp. 17–56.

[12] E. Pop, “Energy dissipation and transport in nanoscale devices,” *Nano Res.*, vol. 3, no. 3, pp. 147–169, 2010.

[13] J. Koomey and S. Naffziger, “Efficiency’s brief reprieve,” *IEEE Spectrum*, Apr-2015.

[14] R. R. Schaller, “Moore’s law: Past, present, and future,” *IEEE Spectr.*, vol. 34, no. 6, pp. 52–59, Jun. 1997.

[15] D. Rotman, “We’re not prepared for the end of Moore’s law,” *MIT Technology Review*, Mar-2020.

[16] K. F. Brennan, *The Physics of Semiconductors: With Applications to Optoelectronic Devices*. Cambridge University Press, 1999.

[17] K. F. Brennan, *Introduction to Semiconductor Devices: For Computing and Telecommunications Applications*. Cambridge: Cambridge University Press, 2005.

[18] S. S. Li, Ed., *Semiconductor Physical Electronics*, 2nd ed. New York: Springer New York, 2006.

[19] P. Siffert and E. Krimmel, Eds., *Silicon: Evolution and Future of a Technology*, 1st ed. Springer-Verlag Berlin Heidelberg, 2004.

[20] T. L. Bergman, A. S. Lavine, F. P. Incropera, and D. P. Dewitt, *Fundamentals of Heat and Mass Transfer*, 7th ed. Wiley, 2011.

[21] M. I. Flik, B. I. Choi, and K. E. Goodson, “Heat transfer regimes in microstructures,” *J. Heat Transfer*, vol. 114, no. 3, pp. 666–674, Aug. 1992.

[22] W. M. Rohsenow and H. Y. Choi, *Heat, Mass, and Momentum Transfer*, vol. 108. Prentice-Hall Englewood Cliffs, NJ, 1961.

[23] L. Brillouin, *Wave Propagation in Periodic Structures: Electric Filters and Crystal Lattices*, 2nd ed. New York: Dover Publications, 1953.

[24] S. Wei and M. Y. Chou, “Phonon dispersions of silicon and germanium from first-principles calculations,” *Phys. Rev. B*, vol. 50, no. 4, pp. 2221–2226, Jul. 1994.

[25] D. R. Lide, Ed., *CRC Handbook of Chemistry and Physics*, 85th ed. Boca Raton: CRC Press, 2005.

[26] A. A. Balandin, “Thermal properties of graphene and nanostructured carbon materials,” *Nat. Mater.*, vol. 10, no. 8, pp. 569–581, Aug. 2011.

[27] W. Haensch *et al.*, “Silicon CMOS devices beyond scaling,” *IBM J. Res. Dev.*, vol. 50, no. 4–5, pp. 339–361, Jul. 2006.

[28] S. Hanson *et al.*, "Ultralow-voltage minimum-energy CMOS," *IBM J. Res. Dev.*, vol. 50, no. 4–5, pp. 469–490, Jul. 2006.

[29] F. Voelklein and T. Starz, "Thermal conductivity of thin films - Experimental methods and theoretical interpretation," in *International Conference on Thermoelectrics, ICT, Proceedings*, 1997, pp. 711–718.

[30] D. G. Cahill, H. E. Fischer, T. Klitsner, E. T. Swartz, and R. O. Pohl, "Thermal conductivity of thin films: Measurements and understanding," *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.*, vol. 7, no. 3, pp. 1259–1266, May 1989.

[31] P. Schelling, "Thermal properties of metallic films," in *Metallic Films for Electronic, Optical and Magnetic Applications: Structure, Processing and Properties*, K. Barmak and K. Coffey, Eds. Cambridge: Woodhead Publishing, 2014.

[32] G. Keresztury, "Raman spectroscopy: Theory," in *Handbook of Vibrational Spectroscopy*, vol. 1, J. M. Chalmers and P. R. Griffiths, Eds. Chichester: Wiley, 2002.

[33] Y. Itoh and T. Hasegawa, "Polarization dependence of raman scattering from a thin film involving optical anisotropy theorized for molecular orientation analysis," *J. Phys. Chem. A*, vol. 116, no. 23, pp. 5560–5570, Jun. 2012.

[34] M. Gu, Y. Zhou, L. Pan, Z. Sun, S. Wang, and C. Q. Sun, "Temperature dependence of the elastic and vibronic behavior of Si, Ge, and diamond crystals," *J. Appl. Phys.*, vol. 102, no. 8, p. 083524, Oct. 2007.

[35] B. C. Johnson, B. Haberl, J. E. Bradby, J. C. McCallum, and J. S. Williams, "Temperature dependence of Raman scattering from the high-pressure phases of Si induced by indentation," *Phys. Rev. B - Condens. Matter Mater. Phys.*, vol. 83, no. 23, p. 235205, Jun. 2011.

[36] Z.-Y. Cao and X.-J. Chen, "Phonon scattering processes in molybdenum disulfide," *Appl. Phys. Lett.*, vol. 114, no. 5, p. 52102, Feb. 2019.

[37] A. Lugstein *et al.*, "In situ monitoring of Joule heating effects in germanium nanowires by µ-Raman spectroscopy," *Nanotechnology*, vol. 24, no. 6, p. 65701, Jan. 2013.

[38] F. Ahmed *et al.*, "High electric field carrier transport and power dissipation in multilayer black phosphorus field effect transistor with dielectric engineering," *Adv. Funct. Mater.*, vol. 27, no. 4, Jan. 2017.

[39] N. J. Everall, "Raman spectroscopy of the condensed phase," in *Handbook of Vibrational Spectroscopy*, J. M. Chalmers and P. R. Griffiths, Eds. Chichester: Wiley, 2002.

[40] L. M. Malard, M. A. Pimenta, G. Dresselhaus, and M. S. Dresselhaus, "Raman spectroscopy in graphene," *Phys. Rep.*, vol. 473, no. 5–6, pp. 51–87, Apr. 2009.

[41] C. Canali, G. Ottaviani, and A. Alberigi Quaranta, "Drift velocity of electrons and holes and associated anisotropic effects in silicon," *J. Phys. Chem. Solids*, vol. 32, no. 8, pp. 1707–1720, Jan. 1971.

[42] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta, "Electron drift velocity in silicon," *Phys. Rev. B*, vol. 12, no. 6, pp. 2265–2284, 1975.

[43] A. A. Quaranta, C. Jacoboni, and G. Ottaviani, "Negative differential mobility in III-V and II-VI semiconducting compounds," *La Riv. Del Nuovo Cim.*, vol. 1, no. 4, pp. 445–495, Oct. 1971.

[44] J. L. Su, Y. Nishi, J. L. Moll, and A. Neukermans, "High-field drift velocity of electrons in p-type silicon," *Solid State Electron.*, vol. 13, no. 7, pp. 1115–1117, Jul. 1970.

[45] R. Cecchi, A. Loria, M. Martini, and G. Ottaviani, "Electrons and holes drift velocity in silicon at very low temperature," *Solid State Commun.*, vol. 6, no. 10, pp. 727–728, Oct. 1968.

[46] D. G. Cahill, K. Goodson, and A. Majumdar, "Thermometry and thermal transport in micro/nanoscale solid-state devices and structures," *J. Heat Transfer*, vol. 124, no. 2, pp. 223–241, Apr. 2002.

[47] A. Majumdar, J. Lai, M. Chandrachood, O. Nakabeppu, Y. Wu, and Z. Shi, "Thermal imaging by atomic force microscopy using thermocouple cantilever probes," *Rev. Sci. Instrum.*, vol. 66, no. 6, pp. 3584–3592, Jun. 1995.

[48] J. Lai, A. Majumdar, M. Chandrachood, and J. P. Carrejo, "Thermal Detection of Device Failure by Atomic Force Microscopy," *IEEE Electron Device Lett.*, vol. 16, no. 7, pp. 312–315, Jul. 1995.

[49] Y. Zhang, P. S. Dobson, and J. M. R. Weaver, "High temperature imaging using a thermally compensated cantilever resistive probe for scanning thermal microscopy," *J. Vac. Sci. Technol. B, Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.*, vol. 30, no. 1, p. 010601, Jan. 2012.

[50] Y. S. Ju and K. E. Goodson, "Short-time-scale thermal mapping of microdevices using a scanning thermoreflectance technique," *J. Heat Transfer*, vol. 120, no. 2, pp. 306–313, May 1998.

[51] J. Bodzenta, "Nanoscale heat transport," *Materials Science-Poland*, vol. 26, no. 1, pp. 95–103, 2008.

[52] R. C. Mclaren, "Thermal conductivity anisotropy in molybdenum disulfide thin films," University of Illinois at Urbana-Champaign, Urbana, 2009.

[53] C. A. Paddock and G. L. Eesley, "Transient thermoreflectance from thin metal films," *J. Appl. Phys.*, vol. 60, no. 1, pp. 285–290, Jul. 1986.

[54] R. Lake and S. Datta, "Energy balance and heat exchange in mesoscopic systems," *Phys. Rev. B*, vol. 46, no. 8, pp. 4757–4763, Aug. 1992.

[55] S. Oda and D. Ferry, Eds., *Silicon Nanoelectronics*. Boca Raton: CRC Press Taylor & Francis Group, 2006.

[56] R. W. Keyes, "Physical limits in semiconductor electronics," *Science*, vol. 195, no. 4283, pp. 1230–1235, Mar. 1977.

[57] X. Luo and B. Wang, "Structural and elastic properties of LaAlO3 from first-principles calculations," *J. Appl. Phys.*, vol. 104, no. 7, p. 073518, 2008.

[58] "Silicon (Si), Debye temperature, heat capacity, density, hardness, melting point," in *Group IV Elements, IV-IV and III-V Compounds. Part b - Electronic, Transport, Optical and Other Properties*, vol. 41A1β, Springer-Verlag, 2005, pp. 1–16.

[59] Y. S. Ju and K. E. Goodson, "Phonon scattering in silicon films with thickness of order 100 nm," *Appl. Phys. Lett.*, vol. 74, no. 20, pp. 3005–3007, May 1999.

[60] D. Vasileska, S. M. Goodnick, and K. Raleva, "Self-consistent simulation of heating effects in nanoscale devices," in *Proceedings - 2009 13th International Workshop on Computational Electronics, IWCE 2009*, 2009, pp. 1–4.

[61] Z. Aksamija, Ed., *Nanophononics: Thermal Generation, Transport, and Conversion at the Nanoscale*, First edit. New York: Pan Stanford Publishing, 2018.

[62] G. Chen, "Particularities of heat conduction in nanostructures," *J. Nanoparticle Res.*, vol. 2, no. 2, pp. 199–204, 2000.

[63] A. C. Sparavigna, "The Boltzmann equation of phonon thermal transport solved in the

relaxation time approximation - I - Theory," *Mater. Sci. Eng. J.*, pp. 34–45, Mar. 2016.

[64] F. Seitz and D. Turnbull, Eds., *Solid State Physics: Advances in Research and Applications*, 1st ed., vol. 7. New York: Academic Press, 1958.

[65] A. A. Joshi and A. Majumdar, "Transient ballistic and diffusive phonon heat transport in thin films," *J. Appl. Phys.*, vol. 74, no. 1, pp. 31–39, Jul. 1993.

[66] R. E. Peierls, *Quantum Theory of Solids*. Oxford: Clarendon Press, 1955.

[67] J. R. Howell, M. P. Mengüç, and R. Siegel, *Thermal Radiation Heat Transfer*, 6th ed. Boca Raton: CRC Press, 2015.

[68] G. Chen, D. Borca-Tasciuc, and R. G. Yang, "Nanoscale Heat Transfer," in *Encyclopedia of Nanoscience and Nanotechnology*, H. S. Nalwa, Ed. American Scientific Publishers, 2004, pp. 429–460.

[69] S. A. Ali and S. Mazumder, "Phonon Boltzmann transport equation based modeling of time domain thermo-reflectance experiments," *Int. J. Heat Mass Transf.*, vol. 107, pp. 607–621, Apr. 2017.

[70] P. G. Sverdrup, Y. S. Ju, and K. E. Goodson, "Sub-continuum simulations of heat conduction in silicon-on-insulator transistors," *J. Heat Transfer*, vol. 123, no. 1, pp. 130–137, Feb. 2001.

[71] G. K. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 9, no. 11, pp. 1141–1149, 1990.

[72] E. Pop, "Self-heating and scaling of thin body transistors," Stanford University, 2004.

[73] E. Pop, J. A. Rowlette, R. W. Dutton, and K. E. Goodson, "Joule heating under quasi-ballistic transport conditions in bulk and strained silicon devices," in *International Conference on Simulation of Semiconductor Processes and Devices, SISPAD*, 2005, vol. 2005, pp. 307–310.

[74] R. Stratton, "Diffusion of hot and cold electrons in semiconductor barriers," *Phys. Rev.*, vol. 126, no. 6, pp. 2002–2014, Jun. 1962.

[75] K. Blotekjaer, "Transport Equations for Electrons in Two-Valley Semiconductors," *IEEE Trans. Electron Devices*, vol. 17, no. 1, pp. 38–47, Jan. 1970.

[76] J. Lai and A. Majumdar, "Concurrent thermal and electrical modeling of sub-micrometer

silicon devices," *J. Appl. Phys.*, vol. 79, no. 9, pp. 7353–7361, May 1996.

[77] M. Rudan, M. Lorenzini, and R. Brunetti, "Hydrodynamic simulation of semiconductor devices," in *Theory of Transport Properties of Semiconductor Nanostructures*, vol. 4, E. Schöll, Ed. Boston, MA: Springer US, 1998, pp. 27–57.

[78] C. Moglestue, F. A. Buot, and W. T. Anderson, "Ensemble Monte Carlo particle investigation of hot electron induced source-drain burnout characteristics of GaAs field-effect transistors," *J. Appl. Phys.*, vol. 78, no. 4, pp. 2343–2348, Aug. 1995.

[79] T. Grasser and S. Selberherr, "Limitations of hydrodynamic and energy-transport models," in *Proceedings of the International Society for Optical Engineering*, 2002, vol. 1, pp. 584–591.

[80] O. Muscato and V. Di Stefano, "Heat generation and transport in nanoscale semiconductor devices via Monte Carlo and hydrodynamic simulations," *COMPEL - Int. J. Comput. Math. Electr. Electron. Eng.*, vol. 30, no. 2, pp. 519–537, Mar. 2011.

[81] A. M. Anile, M. Junk, V. Romano, and G. Russo, "Cross-validation of numerical schemes for extended hydrodynamical models of semiconductors," *Math. Model. Methods Appl. Sci.*, vol. 10, no. 6, pp. 833–861, Aug. 2000.

[82] N. Metropolis and S. Ulam, "The Monte Carlo Method," *J. Am. Stat. Assoc.*, vol. 44, no. 247, pp. 335–341, Sep. 1949.

[83] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*, 1st ed. Springer Vienna, 1989.

[84] H. Kosina, M. Nedjalkov, and S. Selberherr, "Theory of the Monte Carlo method for semiconductor device simulation," *IEEE Trans. Electron Devices*, vol. 47, no. 10, pp. 1898–1908, Oct. 2000.

[85] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Rev. Mod. Phys.*, vol. 55, no. 3, pp. 645–705, Jul. 1983.

[86] 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.

[87] F. Nofeli and H. Arabshahi, "Electronic transport properties in bulk ZnO and Zn1-

xMgxO using Monte Carlo simulation," *Glob. J. Sci. Front. Res.*, vol. 15, no. 3, pp. 118–125, 2015.

[88] M. G. Adeleh and K. M. Reza, "Assessment of Monte Carlo simulation of electron transport in ZnO diode in intelligent information systems," *Int. J. Intell. Inf. Syst.*, vol. 8, no. 2, p. 43, 2019.

[89] P. J. Price, "Monte Carlo calculation of electron transport in solids," in *Semiconductors and Semimetals*, vol. 14, no. C, R. K. Willardson and A. C. Beer, Eds. 1979, pp. 249–308.

[90] W. Fawcett, A. D. Boardman, and S. Swain, "Monte Carlo determination of electron transport properties in gallium arsenide," *J. Phys. Chem. Solids*, vol. 31, no. 9, pp. 1963–1990, Sep. 1970.

[91] "Semiconductor module user's guide," COMSOL, 5.4, 2018.

[92] F. Venturi, R. K. Smith, E. C. Sangiorgi, M. R. Pinto, and B. Ricco, "A General Purpose Device Simulator Coupling Poisson and Monte Carlo Transport with Applications to Deep Submicron MOSFET's," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 8, no. 4, pp. 360–369, Apr. 1989.

[93] M. Lundstrom, *Fundamentals of Carrier Transport*. Cambridge University Press, 2000.

[94] H. D. Rees, "Calculation of steady state distribution functions by exploiting stability," *Phys. Lett. A*, vol. 26, no. 9, pp. 416–417, Mar. 1968.

[95] H. D. Rees, "Calculation of distribution functions by exploiting the stability of the steady state," *J. Phys. Chem. Solids*, vol. 30, no. 3, pp. 643–655, Mar. 1969.

[96] M. A. Littlejohn, R. J. Trew, J. R. Hauser, and J. M. Golio, "Electron transport in planar-doped barrier structures using an ensemble Monte Carlo method," *J. Vac. Sci. Technol. B Microelectron. Process. Phenom.*, vol. 1, no. 2, pp. 449–454, 1983.

[97] J. M. Higman, K. Hess, C. G. Hwang, and R. W. Dutton, "Coupled Monte Carlo-Drift Diffusion Analysis of Hot-Electron Effects in MOSFET's," *IEEE Trans. Electron Devices*, vol. 36, no. 5, pp. 930–937, May 1989.

[98] C. Jungemann and B. Meinerzhagen, "On the applicability of nonself-consistent Monte Carlo device simulations," *IEEE Trans. Electron Devices*, vol. 49, no. 6, pp. 1072–1074, Jun. 2002.

[99] M. Shur, "Semiconductors," in *The Electrical Engineering Handbook*, W.-K. Chen, Ed. Burlington: Academic Press, 2005, pp. 153–162.

[100] M. C. Cheng and E. E. Kunhardt, "Monte Carlo investigation of acoustic, intervalley and intravalley deformation potentials for GaAs," *Solid State Commun.*, vol. 79, no. 8, pp. 651–655, Aug. 1991.

[101] L. Tirino, M. Weber, K. F. Brennan, and E. Bellotti, "A general Monte Carlo model including the effect of the acoustic deformation potential on the transport properties," *J. Comput. Electron.*, vol. 3, no. 2, pp. 81–93, Apr. 2004.

[102] R. Brunetti, C. Jacoboni, F. Nava, L. Reggiani, G. Bosman, and R. J. J. Zijlstra, "Diffusion coefficient of electrons in silicon," *J. Appl. Phys.*, vol. 52, no. 11, pp. 6713–6722, 1981.

[103] 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.

[104] E. Sangiorgi, B. Ricco, and F. Venturi, "MOS/sup 2/: An efficient Monte Carlo simulator for MOS devices," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 7, no. 2, pp. 259–271, 1988.

[105] N. Sano, T. Aoki, M. Tomizawa, and A. Yoshii, "Electron transport and impact ionization in Si," *Phys. Rev. B*, vol. 41, no. 17, pp. 12122–12128, Jun. 1990.

[106] P. D. Yoder and K. Hess, "First-principles Monte Carlo simulation of transport in Si," *Semicond. Sci. Technol.*, vol. 9, no. 5S, p. 854, 1994.

[107] T. Kunikiyo *et al.*, "A Monte Carlo simulation of anisotropic electron transport in silicon including full band structure and anisotropic impact-ionization model," *J. Appl. Phys.*, vol. 75, no. 1, pp. 297–312, Jan. 1994.

[108] A. Duncan, U. Ravaioli, and J. Jakumeit, "Full-band Monte Carlo investigation of hot carrier trends in the scaling of metal-oxide-semiconductor field-effect transistors," *IEEE Trans. Electron Devices*, vol. 45, no. 4, pp. 867–876, Apr. 1998.

[109] H. Kosina, "Method to reduce small-angle scattering in Monte Carlo device analysis," *IEEE Trans. Electron Devices*, vol. 46, no. 6, pp. 1196–1200, 1999.

[110] E. Pop, R. W. Dutton, and K. E. Goodson, "Analytic band Monte Carlo model for electron transport in Si including acoustic and optical phonon dispersion," *J. Appl. Phys.*,

vol. 96, no. 9, pp. 4998–5005, Nov. 2004.

[111] E. Pop, R. W. Dutton, and K. E. Goodson, “Monte Carlo simulation of Joule heating in bulk and strained silicon,” *Appl. Phys. Lett.*, vol. 86, no. 8, pp. 1–3, Feb. 2005.

[112] Z. Aksamija and U. Ravaioli, “Joule heating and phonon transport in silicon MOSFETs,” *J. Comput. Electron.*, vol. 5, no. 4, pp. 431–434, Dec. 2006.

[113] Z. Aksamija and U. Ravaioli, “Joule heating and phonon transport in nanoscale silicon MOSFETs,” in *2007 IEEE International Conference on Electro/Information Technology, EIT 2007*, 2007, pp. 70–72.

[114] K. Raleva, D. Vasileska, S. M. Goodnick, and M. Nedjalkov, “Modeling thermal effects in nanodevices,” *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1306–1316, Jun. 2008.

[115] A. Ashok, D. Vasileska, O. Hartin, and S. M. Goodnick, “Monte Carlo simulation of GaN n+nn+ diode including intercarrier interactions,” in *2007 7th IEEE Conference on Nanotechnology (IEEE NANO)*, 2007, pp. 338–341.

[116] N. Harada, Y. Awano, S. Sato, and N. Yokoyama, “Monte Carlo simulation of electron transport in a graphene diode with a linear energy band dispersion,” *J. Appl. Phys.*, vol. 109, no. 10, p. 104509, May 2011.

[117] Z. Shomali, B. Pedar, J. Ghazanfarian, and A. Abbassi, “Monte-Carlo parallel simulation of phonon transport for 3D silicon nano-devices,” *Int. J. Therm. Sci.*, vol. 114, pp. 139–154, Apr. 2017.

[118] J. Fang *et al.*, “Understanding the average electron-hole pair-creation energy in silicon and germanium based on full-band Monte Carlo simulations,” *IEEE Trans. Nucl. Sci.*, vol. 66, no. 1, pp. 444–451, Jan. 2019.

[119] T. T. Nghiem, N. Trannoy, and J. Randrianalisoa, “Monte Carlo prediction of ballistic effect on phonon transport in silicon in the presence of small localized heat source,” *Nanotechnology*, vol. 30, no. 41, p. 415403, Jul. 2019.

[120] M. A. Green, “Intrinsic concentration, effective densities of states, and effective mass in silicon,” *J. Appl. Phys.*, vol. 67, no. 6, pp. 2944–2954, Mar. 1990.

[121] B. Fischer and K. R. Hofmann, “A full-band Monte Carlo model for the temperature dependence of electron and hole transport in silicon,” *Appl. Phys. Lett.*, vol. 76, no. 5,

pp. 583–585, Jan. 2000.

[122] C. Hamaguchi, *Basic Semiconductor Physics*, 3rd ed. Cham: Springer International Publishing, 2017.

[123] C. Herring and E. Vogt, "Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering," *Phys. Rev.*, vol. 101, no. 3, pp. 944–961, Feb. 1956.

[124] V. Misra and M. C. Öztürk, "Field Effect Transistors," in *The Electrical Engineering Handbook*, W.-K. Chen, Ed. Burlington: Academic Press, 2005, pp. 109–126.

[125] S. Reggiani, M. Valdinoci, L. Colalongo, M. Rudan, and G. Baccarani, "An analytical, temperature-dependent model for majority- and minority-carrier mobility in silicon devices," *VLSI Des.*, vol. 10, p. 52147, 2000.

[126] J. Pernot, "Carrier mobility in diamond: From material to devices," in *Power Electronics Device Applications of Diamond Semiconductors*, S. Koizumi and J. Pernot, Eds. Woodhead Publishing, 2018, pp. 174–189.

[127] R. Granzner *et al.*, "Simulation of nanoscale MOSFETs using modified drift-diffusion and hydrodynamic models and comparison with Monte Carlo results," *Microelectron. Eng.*, vol. 83, no. 2, pp. 241–246, Feb. 2006.

[128] J. Rowlette, E. Pop, S. Sinha, M. Panzer, and K. Goodson, "Thermal simulation techniques for nanoscale transistors," in *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, 2005, vol. 2005, pp. 225–228.

[129] H. K. Sy and C. K. Ong, "Electron mobility in heavily doped silicon," *Solid State Commun.*, vol. 52, no. 10, pp. 881–883, Dec. 1984.

CERTIFICATE OF COMPLETENESS

It is hereby certified that the dissertation submitted by NS Faraz Kaiser Malik, Reg. No. **00000276894**, titled: ***Numerical Analysis of Self-Heating in Silicon Devices*** has been checked/reviewed, and its contents are complete in all respects.

Supervisor's Name: **Dr. Tariq Talha**

Signature: ________________

Date: ________________