

UCLA

UCLA Electronic Theses and Dissertations

Title

Investigation and modeling of impact ionization spatial-transient effects in silicon devices

Permalink

<https://escholarship.org/uc/item/01g1s341>

Author

Chau, Quan Nghia

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Investigation and Modeling of
Impact Ionization Spatial-Transient Effects in Silicon Devices

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Electrical Engineering

by

Quan Nghia Chau

2012

© Copyright by
Quan Nghia Chau
2012

ABSTRACT OF THE DISSERTATION

Investigation and Modeling of Impact Ionization Spatial-Transient Effects in Silicon Devices

by

Quan Nghia Chau

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2012

Professor Dee-Son Pan, Chair

Impact ionization (II) has played an important role in semiconductor devices; yet the understanding of II has not been mature. Abnormal behaviors related to II in deep sub-micrometer devices were observed and have not been fully explained. Existing models are not rigorously applicable to predicting II in different device structures and different operational regimes. Monte Carlo (MC) programs simulating transport of both electrons and holes are developed to investigate II in homogeneous electric field and in scaled devices. The programs' accuracy is verified by accurately producing many different transport parameters obtained from both experiments and previous MC simulations' results.

Impact ionization, for the first time, is modeled as a positive feedback loop in which electrons create holes, and the secondary holes feed back secondary electrons. This model is analytically proven to be valid for short devices due to the existence of the II dead-space. This

model is also numerically proven to be accurate by producing a good fit to the experimental data. It is easy to conclude from the positive feedback model that the breakdown voltage is the same for both the electron-initiating and hole-initiating II processes in a high field region. In addition, the positive feedback model also shows that the current gain from the electron-initiating II process is always higher than the current gain from the hole-initiating II process within the same high field region. More importantly, the positive feedback loop enables successful simulations of the II process in which both electrons and holes participate simultaneously. This is particularly important at high current gain. An efficient algorithm is also developed to speed up spatial transient simulations by implementing temporal meshes rather than the traditional spatial meshes.

The II current gain in short p-i-n diodes is studied. The calculated results fit well to the experimental data of diodes with different lengths. Various physical insights are learned from the simulations. The minimum breakdown voltage for highly doped junctions is extrapolated to be at least 4.41V. Franz-Keldysh effect plays a significant role at low bias, especially for short devices. For the first time, Franz-Keldysh effect is invoked to explain the experimental current gain. II threshold energy is not constant with respect to the electric field, which partly explains various values of the reported threshold energy. II threshold energy is higher for holes than for electrons. Both electron and hole II coefficients come to equilibrium with the electric field after a dead-space distance. This spatial transient effect is a major cause for the disagreements among the experimentally-extracted values of the II coefficients. The values extracted from the double drift p-n junction experiments are more reliable in terms of accounting for the II spatial-transient effect. The II spatial-transient effect is identified to be the main cause for the failures of different well-known II models for semiconductor devices. A pseudo-local electric field model and the

positive feedback model are proposed and proven to be sufficient in predicting the I_2 current gain in short devices.

MC simulations are conducted to study mixed tunneling and I_2 process in short p-n diodes, which are potential terahertz source devices. Tunneling current is treated as generation current, which is also subject to the tunneling dead-space distance. Another gain stage is added on top of the positive feedback model to account for the tunneling dead-space. I_2 is less important for more heavily doped p-n junctions. The contribution of the diffusion current and its I_2 is negligible compared to the tunneling counterparts.

Abnormal behaviors of I_2 in deep sub-micrometer MOSFETs are investigated and explained. Channel carrier distribution functions are generated by MC simulations employing the rare-state algorithm. The thermal tail of the distribution function is Maxwellian with the lattice temperature as the effective temperature. By formulating the thermal tails as functions of position and bias voltage, an analytical formula of the substrate current is successfully derived for the first time. The formula is then used to explain experimental results of the substrate current in a sub-micrometer pMOSFET. The newly-developed formula is able to explain different abnormal behaviors of the substrate current that cannot be explained by the conventional formulas.

The dissertation of Quan Nghia Chau is approved.

Oscar M. Stafsudd

Chi On Chui

Vidvuds Ozolins

Dee-Son Pan, Committee Chair

University of California, Los Angeles

2012

Contents

	Page
Abstract	ii
List of Figures	x
List of Tables	xiv
Acknowledgment	xv
Vita	xvi
 Chapters	
 1 Introduction	1
1.1 Motivations and objectives	1
1.2 Organization	5
Chapter bibliography	8
2 Monte Carlo models	11
2.1 Introduction	11
2.2 Monte Carlo model for electrons	13
2.2.1 Conduction band model	13
2.2.2 Electron-phonon scattering probability and impact ionization rate	16

2.2.3	Procedures of Monte Carlo simulations	22
2.2.4	Results of parameter calibrations for Monte Carlo simulations of electrons	27
2.3	Monte Carlo simulations for holes	34
2.4	Summary	39
	Chapter bibliography	40
3	A novel approach to Monte Carlo simulations of an impact ionization process in semiconductor devices	44
3.1	Introduction	44
3.2	Positive feedback model	46
3.3	Implementation of the positive feedback model in Monte Carlo simulations	54
3.4	An efficient Monte Carlo algorithm to simulate spatial transient transport in semiconductor devices	59
3.5	Summary	62
	Chapter bibliography	63
4	New perspectives on physics and modeling of impact ionization	64
4.1	Introduction	64
4.2	Analysis of models for impact ionization in semiconductor devices	67
4.2.1	Local electric field model	67
4.2.2	Delay model	68
4.2.3	Energy model	69
4.3	Monte Carlo investigation of spatial-transient impact ionization	73
4.3.1	Monte Carlo calculations of current gain in short p-i-n diodes	73
4.3.2	Franz-Keldysh effect on impact ionization	76

4.3.3	Spatial transient values of impact ionization coefficients	78
4.3.4	Evaluations of the experimentally extracted impact ionization coefficients	83
4.4	A proposal of simple and sufficient models for impact ionization in semiconductor devices	87
4.4.1	Pseudo-local electric field model	87
4.4.2	Positive feedback model	89
4.5	Summary	91
	Chapter bibliography	93
5	Study of mixed tunneling and impact ionization in short p-n diodes	95
5.1	Introduction	95
5.2	Modeling mixed tunneling and II current	97
5.2.1	Tunneling current	97
5.2.2	Modeling an II process initiated by the tunneling-generated electrons	101
5.3	Monte Carlo simulations of mixed tunneling and impact ionization current	104
5.3.1	Impact ionization due to tunneling-generated electrons	104
5.3.2	Impact ionization due to diffusing current	107
5.4	Summary	109
	Chapter bibliography	110
6	Investigating paradoxical behaviors of impact ionization and modeling substrate current in deep sub-micrometer MOSFETs	112
6.1	Introduction	112
6.2	Investigating and modeling low voltage impact ionization in deep sub-micrometer MOSFETs	114

6.3	Summary	122
	Chapter bibliography	123
7	Conclusion	125
7.1	Summary	125
7.2	Suggestions for future studies	129
	Appendix	131
A	Calculations of impact ionization rates in silicon	131
A.1	Non-local empirical pseudopotential method of calculating Si band structures	131
A.2	Integration over Brillouin zone	133
A.3	Calculations of impact ionization rates	140
	Chapter bibliography	145

List of figures

	Page
2.1 Schematic of a typical electron-phonon scattering event. An electron with wave vector $\mathbf{k}$ scatters with a phonon $\mathbf{q}$, resulting in an electron with a new wave vector $\mathbf{k}'$.	16
2.2 Flowchart of the MC programs for both electrons and holes.	22
2.3 Geometry for the determination of the new state after a scattering event. The direction of new state $\mathbf{k}'$ is stochastically generated relative to state $\mathbf{k}$ before the collision.	26
2.4 Experimental [6, 27] and simulated electron drift velocity as a function of the electric field along $\langle 100 \rangle$ and $\langle 111 \rangle$ crystallographic directions.	28
2.5 Experimental and simulated electron II coefficients as a function of electric field.	29
2.6 Total electron-phonon scattering rate and electron II rate as functions of energy. Rates from previous MC simulations are shown for comparison [9, 32-33]. Zero energy corresponds to the maximum point of the valance band edge.	30
2.7 Electron distribution function at electric field = 300kV/cm in $\langle 100 \rangle$ crystallographic direction. Dots are MC results from other groups (figure 17 by Higman and Hess, figure 18 by Fischetti and Laux, and figure 20 by Kunikiyo <i>et al.</i> in Ref. [35]). Zero energy corresponds to the minimum point of the conduction band edge.	31
2.8 Electron average energy as a function of electric field. For comparison, results from previous MC simulations are shown (Canali <i>et al.</i> [6]: analytical band MC, Fischetti <i>et al.</i> [9]: full band MC).	32
2.9 Experimental [36-38] and simulated quantum yield as a function of electron energy. Zero energy corresponds to the maximum point of the valance band edge.	33
2.10 Hole II rate and total phonon scattering rate weighted by DOS.	37
2.11 Hole drift velocity as a function of electric field along the $\langle 100 \rangle$ and $\langle 111 \rangle$ directions. Shown for comparison are the results of full-band MC simulations in Ref. [40] and the experimental results [43].	37

2.12	Hole average energy as a function of electric field. Results from previous MC simulations [39-41] are shown for comparison.	38
2.13	Experimental [28-31] and simulated hole II coefficients as a function of electric field.	38
3.1	Impact ionization process in a high field region. Electrons in the conduction band impact-ionize and generate holes in the valence band. Secondary holes then impact-ionize and generate electrons in the conduction band.	44
3.2	A simplified process of II in a high field region. M_n is the II gain due to electrons, and M_p is the II gain due to holes.	46
3.3	Block diagram of the positive feedback model for the II process in a p-n junction.	46
3.4	Band diagram of a p-n junction in reverse bias. Electrons cannot cause II from $[0, d_n]$, and holes cannot cause II from $[W_D - d_p, W_D]$; d_n and d_p are the “so-called” dead-space distances of electrons and holes, respectively.	49
3.5	Band diagram explaining how the MC program counts the number of II events by initiating electrons in a p-n diode.	55
3.6	Flowcharts of spatial mesh (a) and temporal mesh (b) methods in spatial transient MC simulations. For temporal mesh method, n is calculated before the program executes the internal loop and $n = \text{fFT} / t_{\text{step}}$.	61
4.1	A simplified p-i-n diode’s cross-section and its electric field profile. The electric field is constant in the i-region and varies according to Poisson’s equation in the p^+ and n^+ regions.	65
4.2	II current gain of p-i-n diodes versus bias. Lines are theoretical results using the local electric field model, and dots are experimental data [9].	68
4.3	II current gain of p-i-n diodes versus reverse bias. Lines are the measured data [9]. Filled and empty dots are the simulated results with and without hole feed back, respectively.	74
4.4	II current gain of electron and hole versus the intrinsic length of p-i-n diodes at avalanche breakdown condition.	74
4.5	Bias voltages and the i-region’s electric field at avalanche breakdown conditions for diodes of different lengths.	75

4.6	II current gain in p-i-n diodes versus reverse bias. Lines are the measured data [9]. Filled and empty dots are the results of the MC simulations with and without Franz-Keldysh effect on the electron II rate, respectively.	77
4.7	(a) Electron average energy and II coefficient versus energy at $\mathcal{E}=10^6\text{V/cm}$. (b) Distribution functions at different positions versus energy in a homogeneous field $\mathcal{E}=10^6\text{V/cm}$. 20nm: before the first peak in average energy is reached. 37nm and 55nm: the first peak and trough of the average energy, respectively. 200nm and 250nm: electron energy is in equilibrium with the electric field.	80
4.8	Electron (a) and hole (b) spatial-transient II coefficients in different homogeneous electric field.	81
4.9	Electron's and hole's dead-space versus electric field.	82
4.10	Experimental II coefficients extracted in Ref. [10, 21, 23-24].	83
4.11	Typical electric field profiles of diodes in the photo-multiplication experiments in Ref. [10, 21, 24]. Electrons move from $x=0$ to $x=W_D$, and electrons' II can only occur from $x=d_n$ to $x=W_D$.	86
4.12	II current gain in p-i-n diodes versus bias. Lines are calculated results using the pseudo-local electric field model. Dots are the experimental data [9].	89
4.13	II current gain in p-i-n diodes versus bias. Lines and dots are the theoretical calculations using the pseudo-local electric field model and positive feedback model, respectively.	90
5.1	Band diagram of a p-n diode showing a typical tunneling process. A valance band electron tunnels from $x=0$ to the conduction band at position $x=X_{\text{start}}$.	97
5.2	Experimental [11] (dots) and simulated current (lines) of p-n diodes versus reverse bias. The theoretical current is calculated as the tunneling generation rate with the effective electric field and including the dead-space effect.	100
5.3	(a) Energy band diagram of a p-n junction describing a mixed tunneling and II process. (b) Block diagram modeling the mixed tunneling and II process.	101
5.4	Mixed tunneling and II current versus reverse bias in different p-n diodes. Dots are the measured values [11]. Broken lines are the simulated tunneling current. Solid lines are the results of MC calculations of mixed tunneling and II.	106

6.1	Typical p-MOSFET channel's lateral electric field in the velocity saturation mode. Channel length is $0.1\mu\text{m}$, and the source-drain bias voltage is 0.8 V . The electric field increases linearly near the source ($x=0$) and increases exponentially near the drain ($x=0.1$).	116
6.2	Hole distribution as a function of energy at different channel positions in a $0.1\mu\text{m}$ p-MOSFET with 0.8V source-drain bias voltage at 300K lattice temperature. The high-energy tails of distribution curves have the effective temperature equal to the lattice temperature. Filled dots are the values of holes at position x whose energy is qV/x .	117
6.3	Experimental (dots) and theoretical (lines) values of the multiplication factor versus the lattice temperature at different source-drain biases with the gate bias at $V_G=0.4\text{V}$ (a) and $V_G=0.8\text{V}$ (b). The solid lines and dash lines show the results of the theoretical calculations with and without taking into account the temperature-dependent bandgap, respectively.	121
A.1	Silicon band structures along different high-symmetry directions.	132
A.2	The first BZ of silicon (a) and its cuboid model (b). The cuboid is divided into smaller cubes (shaded region).	134
A.3	(a) The magnified version of the shaded cube in figure A.2b. (b) Six equal-volume tetrahedral composition of a cube.	135
A.4	Three cases of intersections between an equal-energy plane and a tetrahedron.	136
A.5	Electron DOS of silicon contributed by the four lowest conduction bands. Energy equal to 0 corresponds to the lowest energy of the first conduction band. Contributions from each of the four lowest conduction bands are also plotted.	139
A.6	Hole DOS of silicon from the four highest valance bands versus energy. Contributions from each of the four highest valance bands are also plotted.	139
A.7	A typical process of an II event. An electron in state 1 in the conduction band collides with an electron in state 4 in the valance band, resulting in two conduction band electrons in states 2, 3 and one valance band hole in state 4	140
A.8	(a): Electron II rate versus the initial electron energy. Energy is set to 0 at the lowest point of the first CB. The solid line is the best fit curve. (b): The best-fit electron II rate versus energy. Previous reports of II rates (Kamakura et. al.[7] and Dinh and Jungemann [8]) are shown for comparison.	143
A.9	Hole II rate as a function of the primary hole energy.	144

List of tables

	Page
2.1	Parameters used in calculating the electron-phonon intravalley scattering rate. 18
2.2	X-X intervalley phonon energy and coupling constants ($\times 10^8 \text{eV/cm}$). The values from other groups' MC simulations are shown for comparison purpose. 20
2.3	X-L intervalley phonon energy and coupling constants ($\times 10^8 \text{eV/cm}$). The values from other groups' MC simulations are shown for comparison purpose. 20
2.4	Parameters used in the Monte Carlo program for holes. 36
4.1	Breakdown voltage of p-i-n diodes obtained from the theoretical calculation (equation 4.17) and from the experimental data [15]. 76
5.1	Summary of fitting parameters in calculating the tunneling current of diodes IN746, IN749, and IN751 in Ref. [11]. 100
5.2	Comparison of current gain between II due to the tunneling-generated electrons and II due to of the secondary electrons in sample IN751. 105
5.3	Comparisons of current gain between II due to primary diffusion current and II due to primary tunneling current in sample IN751. 108

Acknowledgements

I would like to express my profound gratitude towards my parents for their infinite love, support and encouragements. Their sacrifice and life principles have deeply shaped my moral compass. I wish to thank my brother, sister, and uncles, who never hesitate to extend the helping hands when I am in need. I am especially grateful for the love and faith of my wife, Sunwind Le, who has shared all the ups and downs with me ever since our first encounter.

I would like to address my sincere appreciation to my advisor, Professor Dee-Son Pan, for his invaluable guidance and continuous support throughout the course of this work. His endless encouragement and patience have helped me overcome insurmountably difficult times to complete this course of higher education. His insightful academic expertise and life philosophy are always informative and inspiring.

I am indebted to Professors Oscar M. Stafstudd, Chi On Chui, and Vidvuds Ozolins for kindly spending time to serve on my committee.

My appreciation goes to Deona Columbia for her administrative help and to my colleague, Dr. Edward Chen, for his support and friendship.

VITA

2003	B.S., Electrical Engineering University of California, Berkeley Berkeley, California, USA
2007	M.S., Electrical Engineering University of California, Los Angeles Los Angeles, California, USA

Chapter 1

Introduction

1.1 Motivations and objectives

Impact ionization (II) has played a very important role in many semiconductor devices. All levels of II current gain have corresponding applications in scaled devices. II at breakdown conditions has been implemented in designing avalanche photodiodes (APDs). APDs are superior to the conventional p-i-n diodes in detecting low-level light. However, noise in APDs is higher due to the statistical fluctuations of the multiplication process of individual carriers. The excess noise factor was derived to be inversely proportional to the ratio of II of holes and electrons [1]. This is a drawback for III-V semiconductors since both types of carriers' II coefficients are close in magnitude [2]. APDs' noise has been observed to decrease as the multiplication width decreases [3-4]. Therefore, short Si-APDs can potentially be engineered for low-noise applications.

Mixed tunneling and impact ionization in p-n junctions can be used as high-Q, low-noise inductances [5-6], potentially replacing the passive inductance components in high-performance RF-CMOS circuits [6]. In addition, mixed tunneling and avalanche transit time (MITATT) diodes are capable of generating useful power with low noise in the terahertz frequency regime

[7-8], which has attracted increasing interests for its promising applications in communications, imaging, sensing, and spectroscopy [9].

As DRAM scaling continues, it has become more challenging to incorporate and miniaturize the charge-storage capacitors to the process. One-transistor random access memory (1T-RAM) is a very potential candidate to overcome the challenge of the capacitor integration of conventional DRAM [10-12]. The floating body of a FINFET or a SOI transistor is used as the charge-storing capacitor. The parasitic BJT current is utilized for writing and reading data. The whole operation relies on the latch condition $\beta(M-1) \sim 1$, where β is the BJT current gain and M is the impact-ionization current gain near the drain. The operational drain bias in the intrinsic BJT is smaller than the avalanche breakdown condition for the isolated drain-body junction. For example, with the BJT gain of 100, the drain needs to be biased high enough just to yield $M-1 \sim 0.01$.

In general, II in MOSFETs is of great concern for the reliability of the CMOS technology. The increased electric field in the channel increases the injection rate of hot carriers into the gate's dielectric. Consequently, the interface states and the oxide charges are accumulated near the drain junction, degrading devices' performance and stability [13-14]. This is usually referred to as the hot-carrier effect. Substrate current, caused by low II current gain of the channel current, is commonly used to monitor the hot-carrier effect in MOSFET's [13-14]. Hot-carrier effect has been a major challenge for device reliability for three decades. As MOSFET's downscaling is carried on beyond the 45 nm technologies, hot-carrier effect is still an important issue for multiple-gate MOSFETs, such as FINFETs. Usually, it is more difficult to measure the substrate current due to II in FINFETs than in the conventional MOSFETs. The hot-

carrier effect in conventional MOSFETs shows interesting anomalous behaviors in the deep sub-micrometer regime and is still not well understood [15-18].

Despite the important role of II, basic II parameter values such as threshold energy, II rates, and II coefficients have not been fully agreed among researchers. For example, different reported values of electron threshold energy are 1.1eV [19], 1.65eV [20], 1.8eV [21-22], 3.2eV [23]. Different values of II coefficients were experimentally extracted among different groups [20-22, 24], and there is still not enough comprehensive analysis on the causes of the differences and the reliability of the reported values.

More complications involving II arise in scaled devices where the spatial-transient transport becomes more important. The local electric field model describing the II process in devices is no longer valid [25-26]. A few models were attempted to modify the local electric field model [27-30]. The delay model [27-29] cannot be generally solved analytically but requires computer simulations to follow each charge carrier's trajectory. Moreover, this approach is not self-consistent, and its validity is not physically justified. The energy model [30] is not rigorously applicable to predicting the high II current gain where both electron and hole cause significant II [30]. Furthermore, this model is based on carriers' average energy. Therefore, the energy model is not applicable in the low-bias regime, especially at bias lower than the sub-bandgap energy.

Overall, all of the models above fail to predict the avalanche-breakdown conditions in short devices. They also fail to predict the abnormalities of II current at low bias in sub-micrometer MOSFETs [15-18]. In addition, all of the models above do not include the Franz-Keldysh effect, which will be shown later to be very important in scaled devices.

The objectives of this work are investigating and modeling the spatial-transient effects of an II process. This work involves studying II at a very microscopic level using Monte Carlo (MC) simulations. Fundamental II parameters including threshold energy, II rates, and II coefficients will be investigated and clarified. The spatial-transient values of the II coefficients will be investigated and calculated by MC simulations. The mixed tunneling and II process and the II process in different device structures will also be studied by means of MC simulations. Causes of the abnormal II behaviors in short devices will be identified and analyzed. From the basis of understanding the II process as a result of MC simulations, several widely-used models of II for devices will be evaluated with respect to their applicability limits. Modifications to the II fundamental parameters will be proposed to correctly predict current gain in short devices at low-supply bias. Different models will also be proposed to calculate the II current at different gain levels. Finally, an analytical formula will be derived for the substrate current in deep sub-micrometer MOSFETs operating at sub-bandgap voltage.

1.2 Organization

In chapter 2, MC simulations are set up for both electrons and holes using Matlab. Analytical band-structure models are used. Six X valleys and eight L valleys are included to model the conduction band edge, and they are modeled as an elliptical nonparabolic analytical band. Both acoustic and optical intervalley and acoustic intravalley phonon scattering types are taken into account with six X-X intervalley scattering phonons and four X-L intervalley scattering phonons. For holes, a single warped parabolic analytical band is used to model the valance band edge. Both acoustic and optical intravalley phonon scattering types are included. All phonon scattering types are considered isotropic, and their scattering probability as a function of energy is weighted by the density of states. II is treated as an additional scattering mechanism. A fictitious self-scattering scheme is employed to generate random free-flight time. Parameters for MC programs are calibrated by exhaustively fitting not only both analytical band and full band MC simulations' results but also experimental values of charge transport parameters.

In chapter 3, an II process is modeled as a positive feedback loop. Inside a depletion region, electrons impact-ionize and create secondary holes. Those holes, in turn, impact-ionize and create secondary electrons. The coupling of electrons and holes in the self-sustaining manner can eventually lead to avalanche breakdown if the supply bias is high enough. An analytical formula of II current gain is derived using the positive feedback model. The validity of this analytical formula is confirmed by comparing with a well-known formula derived directly from the current continuity equation. Besides, a new and more efficient algorithm for spatial-transient MC simulations is also developed by dividing the charge carriers' free-flight time into smaller

time steps. In addition, a low-cost method is also introduced to account for the β of the secondary carriers.

Chapter 4 analyzes different well-known models of β in devices. This chapter identifies the main factors responsible for those models' limited applicability in predicting the current gain for different devices' structures and different operating regimes. Microscopic physical insights are investigated by means of MC simulations. Short p-i-n diodes are used for the study in this chapter due to their important roles in avalanche photodiodes and BJTs. Besides, short p-i-n diodes are ideal candidates because of their simple device structures and their trivial electric field profile, which reduce the confusion due to parasitic effects. Franz-Keldysh effect is shown to have a strong influence on β at low bias, especially for short devices. The spatial-transient β coefficients are simulated by MC programs for electrons and holes. Causes of the difference in the experimentally-extracted β coefficients are shown, and the reliability of these β coefficient values is discussed. Based on the physical findings from the MC simulations, more efficient models of β in semiconductor devices are proposed.

Chapter 5 investigates mixed tunneling and β in short p-n diodes using MC simulations. Tunneling current is treated as the generation current inside the depletion region. The spatial-transient effect is taken into account when calculating the tunneling current. Extra gain stage is added to the positive feedback model to account for the tunneling dead-space. Theoretical calculations of the mixed tunneling and β current in short p-n diodes are compared against the experimental values. Even though the diffusion current is much smaller than the tunneling current, diffusion carriers have longer available space for β due to the tunneling dead-space effect. The contribution of diffusion current and its β current gain is discussed in this chapter.

Chapter 6 studies II in deep sub-micrometer MOSFETs. The derivation of an analytical formula for II current in deep sub-micrometer MOSFETs at low-bias voltage is performed. This formula is used to fit the experimental data, taking into consideration of the temperature-dependent bandgap and the source-channel barrier height modulation. This formula is able to explain the abnormal behaviors of the substrate current such as increasing substrate current at higher lattice temperature and II occurrences at sub-bandgap voltage.

Finally, chapter 7 summarizes important results of this work. Suggestions on studying related subjects and related fields are given.

Appendix A shows detailed calculations of electrons' and holes' density of states and II rates.

Bibliography

- [1] R.J. McIntyre, "Multiplication noise in uniform avalanche diodes," *Transactions on Electron Devices*, Vol. 13, pp. 164, 1966.
- [2] S.P. Wilson, S. Brand, and R.A. Abram, "Avalanche multiplication properties of GaAs calculated from spatially transient ionization coefficients," *Solid-State Electronics*, Vol. 38, pp. 2095, 1995.
- [3] M.M. Hayat, B.E.A. Saleh, and M.C. Teich, "Effect of dead space on gain and noise of double-carrier-multiplication avalanche photodiodes," *Transactions on Electron Devices*, Vol. 38, pp. 546, 1992.
- [4] C. Hu, K.A. Anselm, B.G. Streetman, and J.C. Campbell, "noise characteristics of thin multiplication region GaAs avalanche photodiodes," *Applied Physics Letters*, Vol. 69, pp. 3734, 1996.
- [5] G.A.M. Hurkx, P.G.M Baltus, J.A.M. Boet, J.A.M Geelen, and J.J.E.M. Hageraats, "The pn-junction as an inductive design component in silicon IC process," *Bipolar/BiCMOS Circuit and Technology Meeting*, pp. 125, 1995.
- [6] Chie-In Lee, "Investigation of p-n junction inductors for RF CMOS circuits," Ph.D. Dissertation, University of California at Los Angeles, 2005.
- [7] X. Bi, J.R. East, U. Ravaioli, G.I. Haddad, "Analysis and design of Si terahertz transit-time diodes," *Solid-State Electronics*, Vol. 50, pp. 889, 2006.
- [8] X. Bi, J.R. East, U. Ravaioli, G.I. Haddad, "A Monte Carlo study of Si/SiGe MITATT diodes for terahertz power generation" *Solid-State Electronics*, Vol. 52, pp. 689, 2008.
- [9] P.H. Siegel, "THz technology: an overview," in "Terahertz sensing technology, Vol. 1: Electronic devices and advanced systems technology," edited by D.L. Woolard, W.R. Loerop, M.S. Shur, World Scientific, 2003
- [10] S. Okhonin, M. Nagoga, E. Carman, R. Beffa, E. Faraoni, "New Generation of Z-RAM," *International Electron Devices Meeting*, pp. 925, 2007.
- [11] D. Moon, S. Choi, J. Han, S. Kim, and Y. Choi, "Fin-width dependence of BJT-based 1T-DRAM implemented on FinFET," *Electron Device Letters*, Vol. 31, pp. 909, 2010.
- [12] A. Hubert, M. Bawedin, G. Guegan, T. Ernst, O. Faynot, S. Cristoloveanu, "SOI 1T-DRAM cells with variable channel length and thickness: Experimental comparison of programming mechanisms," *Solid-State Electronics*, Vol. 65, pp. 256, 2011.

- [13] C. Hu, S.C. Tam, F. Hsu, P. Ko, T. Chan, and K.W. Terrill, "Hot-electron-induced MOSFET degradation—model, monitor, and improvement," *Transactions on Electron Devices*, Vol. 32, pp. 375, 1985.
- [14] E. Takeda, C.Y. Yang, and A.M. Hamed, "Hot-Carrier effects in MOS devices," New York: Academic, 1995.
- [15] M. Koyanagi, T. Matsumoto, M. Tsuno, T. Shimatani, Y. Yoshida, and H. Watanabe, "Impact ionization phenomenon in 0.1 μ m MOSFET at low temperature and low voltage," *International Electron Devices Meeting*, pp. 341-344, 1993.
- [16] M. Mastrapasqua, J. Bude, M. Pinto, L. Manchanda, and K. F. Lee, "Temperature and channel-length dependence of impact ionization in p-channel MOSFETs," in *1994 SympD. VLSI Tech. Dig. Tech. Papers*, p. 125-126.
- [17] K. G. Anil, S. Mahapatra, and I. Eisele, "Experimental verification of the nature of the high energy tail in the electron energy distribution in n-channel MOSFETs," *Electron Device Letters*, Vol. 22, pp. 478, 2001.
- [18] Pin Su, Ken_Ichi Goto, Toshihiro Sugii, and Chenming Hu, "A Thermal Activation View of Low Voltage Impact Ionization in MOSFETs," *Electron Device Letter*, Vol. 23, pp. 550, 2002.
- [19] Y. Okuto and C.R. Crowell, "Threshold energy effect on avalanche breakdown voltage in semiconductor junctions," *Solid-State Electronics*, Vol. 18, pp. 161, 1975.
- [20] M. H. Woods, W. C. Johnson and M. A. Lampert, "Use of a Schottky Barrier to Measure Impact Ionization Coefficients in Semiconductors," *Solid-State Electronics*, Vol. 16, pp. 81, 1973.
- [21] R. Van Overstraeten and H. De Man, "Measurement of the ionization rates in diffused silicon p-n junctions," *Solid-State Electronics*, Vol. 13, pp. 583-608, 1970.
- [22] W. N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields," *Solid-State Electronics*, Vol. 16, pp. 1189, 1973.
- [23] A. Spinelli, A. Pacelli, and A.L. Lacaita, "Dead space approximation for impact ionization in silicon," *Applied Physics Letters*, Vol. 69, pp. 3707, 1996.
- [24] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Holes and Electrons in Silicon," *Physical Review*, Vol. 134, No. 3A, 1964.
- [25] E.F. Crabbe, J.M.C Stork, G. Baccarani, M.V. Fischetti, and S.E. Laux, "The impact of non-equilibrium transport on breakdown and transit time in bipolar transistors," *International Electron Devices Meeting*, pp. 463-469, 1990.

- [26] M Ershov and V Ryzhii, "Temperature dependence of the electron impact ionization coefficient in silicon," *Semiconductor Science and Technology*, Vol. 10, pp. 138, 1995.
- [27] Wieslaw A. Lukaszek, Aldert Van Der Ziel, and Eugene R. Chenette, "Investigation of the transition from tunneling to impact ionization multiplication in Silicon p-n junctions," *Solid-State Electronics*, Vol. 19, pp. 57-71, 1976.
- [28] A. Di Carlo and P. Lugli, "Dead-space effects under near-breakdown conditions in AlGaAs/GaAs HBT's", *Electron Device Letters*, Vol. 14, pp. 103, 1993.
- [29] C. Canali, P. Pavan, A. Di Carlo, P. Lugli, R. Malik, M. Manfredi, A. Neviani, L. Vendrame, E. Zanoni, and G. Zandler, "Experimental and Monte Carlo analysis of impact-ionization in AlGaAs/GaAs HBT's," *Transactions on Electron Devices*, Vol. 43, pp. 1769, 1996.
- [30] P. Agarwal, M.J. Goosens, V. Zieren, E. Aksen, and J.W. Slotboom, "Impact Ionization in Thin Silicon Diodes," *Electron Device Letters*, Vol. 25, pp. 807, 2004.

Chapter 2

Monte Carlo models

2.1 Introduction

Charge transport in semiconductors can be described by the integro-differential Boltzmann transport equation (BTE), whose solutions provide insights into the microscopic properties of transport phenomena. Unfortunately, solving BTE analytically is almost impossible without approximations. In the early stage of modeling semiconductor devices, BTE was approximated and simplified to the drift-diffusion model. This model can only be used to describe charge transport in long semiconductor devices where charge carriers are in equilibrium with the electric field. However, as devices are scaled down to the sub-micrometer regime, non-stationary or non-local transport becomes more important. The hydrodynamic model has often been considered an alternative approach to the drift-diffusion model in scaled devices. This model, employing higher moments of BTE, captures information of charge transport based on average parameters such as energy, velocity, and relaxation time. However, only high-energy carriers (energy higher than the threshold energy) can cause impact ionization, and full knowledge of the energy distribution function is often needed, especially at low biases.

Therefore, hydrodynamic model is not a rigorously suitable model for the purpose of impact ionization investigations.

Even though BTE is extremely difficult to solve analytically and often involves broad approximations, exact numerical solutions to BTE can be provided by MC simulations. MC method is a statistical numerical technique that directly simulates charge dynamics. Kurosawa [1] was the first to employ the MC method to study charge transport in semiconductors. MC soon became popular and adopted to study complicated charge dynamics in devices [2-3]. MC has been regarded as one of the few most comprehensive approaches to device modeling. This technique follows the trajectories of charge carriers in semiconductors in response to external forces such as electric field or magnetic field. Charge carriers' motion is influenced by different scattering mechanisms including phonon scatterings, carrier-carrier scatterings, impurity scatterings, etc. In MC simulations, the free-flight time (fFT) between scattering events is randomly generated based on the total scattering probability. Similarly, a scattering type at the end of fFT is stochastically selected according to the distribution function of the scattering probability.

One major obstacle discouraging the use of MC simulations is its long computation time. However, with faster computing capacity and larger memory in computers nowadays, MC has become more and more of an excellent tool for charge-transport study.

2.2 Monte Carlo model for electrons

MC simulations were explained in great details by Jacobonie *et al.* [4] and Price [5]. This work closely follows the MC algorithm introduced in Ref. [4].

2.2.1 Conduction band model

For low to intermediate electric field (up to about $5 \times 10^4 \text{V/cm}$), electrons mostly stay in the six equivalent X valleys. The minima points of the X valleys, which are also the conduction band minima, locate at about 85% of the distance from the middle to the edge of the first Brillouin zone (BZ) along $\langle 100 \rangle$ crystallographic directions. In their MC program, Canali *et al.* [6] used an ellipsoidal, parabolic analytical band to model the X valleys:

$$E^i = \frac{\hbar^2}{2} \left(\frac{(\mathbf{k} - \mathbf{k}_0^i)_l^2}{m_l} + \frac{(\mathbf{k} - \mathbf{k}_0^i)_t^2}{m_t} \right) \quad (2.1)$$

where $\hbar$ is the reduced Planck constant; $\mathbf{k}_0^i$ corresponds to the minimum energy of valley i ; l and t subscripts represent the longitudinal and transverse components with respect to symmetry axis of valley i .

Canali *et al.* was among the first to include intervalley scattering among the six X valleys in MC simulations. They successfully fit the measured electron velocity in silicon at $T=300\text{K}$ for electric field up to about 10^4V/cm . At higher electric field, their calculated drift velocity is higher than the measured result [6]. This is due to the fact that E_c - k relationship is not simply parabolic at higher electron energy.

In their MC simulations, Jacobonie *et al.* [7] modified the X valleys' band structures as an elliptical, non-parabolic analytical band, which was first introduced by Conwell *et al.* [8]:

$$E^i(1 + \alpha E^i) = \frac{\hbar^2}{2} \left(\frac{(\mathbf{k} - \mathbf{k}_0^i)_l^2}{m_l} + \frac{(\mathbf{k} - \mathbf{k}_0^i)_t^2}{m_t} \right) \quad (2.2)$$

where α (eV⁻¹) is the non-parabolicity constant term.

With $\alpha = 0.5$ eV⁻¹, Jacobonie *et al.* [7] was able to fit the entire experimental range of drift velocity as a function of the electric field (the maximum electric field was about 5x10⁴V/cm).

However, this result from Jacobonie *et al.* [7] does not guarantee that the analytical band in Ref. [7] can be used to simulate II accurately. The average electron energy at 5x10⁴V/cm and T=300K is about 0.25eV [6, 9], which is much lower than the II threshold energy ($E_T \approx$ bandgap energy ≈ 1.1 eV). The eight secondary minima (L valleys), located along the <111> directions, are in 1eV separation with the minima of X valleys. Therefore, at energy high enough for II to occur, a significant number of electrons also reside in the L valleys. Consequently, MC programs need to include the L valleys for reliable investigations of II.

In their MC simulations of II, Sano *et al.* [11] used the elliptical, nonparabolic analytical band model with both X and L valleys. In addition, at high energy, the electron density of states (DOS) does not increase monotonically with energy as suggested by the analytical band model. From the non-local pseudopotential calculations of Si electronic structures [10], the electron DOS starts decreasing at about 1.8eV and increasing at about 2.5eV. Sano *et al.* [11] also weighted the total electron-phonon scattering rate according to the realistic DOS for electron energy larger than 1.8eV. They were able to fit both the experimental drift velocity and the experimentally-extracted II coefficient for the entire experimental range of the electric field (maximum electric field is about 7x10⁵V/cm).

In light of Sano *et al.*'s [11] success, the MC program in this work employs the elliptical, non-parabolic analytical band for six X valleys and eight L valleys. The phonon scattering rates are also weighted by the realistic DOS.

To simplify the calculations dealing with the elliptical band structures, it is useful to introduce Herring-Vogt transformation method [12]. The transformation matrix is described in the frame centered at the minimum point of the valley with z-axis along its symmetry axis:

$$T = \begin{bmatrix} (\frac{m_0}{m_t})^{1/2} & 0 & 0 \\ 0 & (\frac{m_0}{m_t})^{1/2} & 0 \\ 0 & 0 & (\frac{m_0}{m_l})^{1/2} \end{bmatrix} \quad (2.3)$$

where m_0 is the electron rest mass.

Each component of the transformed wave vector $\mathbf{k}^*$ is expressed as $k_i^* = T_{ij}k_j$.

With the transformation matrix, the elliptical band structure can be formulated as

$$E(\mathbf{k}) = \frac{\hbar^2}{2m_0} T_{ij}T_{il}k_jk_l = \frac{\hbar^2 k^{*2}}{2m_0} \quad (2.4)$$

Including the nonparabolic factor, the elliptical, nonparabolic band structure has the form

$$E(1 + \alpha E) = \frac{\hbar^2 k^{*2}}{2m_0} \quad (2.5)$$

To summarize this section, the MC program for electrons in this work uses the elliptical, non-parabolic analytical band. All of the six X valleys and the eight L valleys are included. For the convenience of calculations, Herring-Vogt transformation is employed. The phonon scattering rates are also weighted by the realistic DOS.

2.2.2 Electron-phonon scattering probability and impact ionization rate

In silicon, electrons interact with phonons via the deformation potential. The deformation potential describes the shift of the electronic band structure per unit crystal deformation due to phonons.

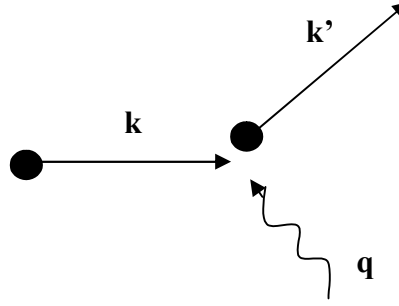


Figure 2.1: Schematic of a typical electron-phonon scattering event. An electron with wave vector $\mathbf{k}$ scatters with a phonon $\mathbf{q}$, resulting in an electron with a new wave vector $\mathbf{k}'$.

Figure 2.1 shows a typical phonon scattering event: an initial electron in state $\mathbf{k}$ interacts with a phonon $\mathbf{q}$ and results in an electron in a new state $\mathbf{k}'$. In an electron-phonon scattering event, both energy and momentum are conserved. The general formula for electron-phonon scattering probability can be expressed as [4]:

$$P(\mathbf{k}, \mathbf{k}') = \frac{\pi}{\rho V \omega_q} \left[\frac{N_q}{N_q + 1} \right] \zeta \left| \zeta_{ij} q_j \xi_i \right|^2 \delta[\varepsilon(\mathbf{k}') - \varepsilon(\mathbf{k}) \mp \hbar \omega_q] \quad (2.6)$$

where the upper and lower symbols correspond to the phonon absorptions and emissions, respectively.

ρ : density of crystal.

V : crystal volume.

ω_q : phonon angular frequency.

$$N_q = \frac{1}{\exp(\frac{\hbar\omega_q}{k_B T_0}) - 1} \text{ number of phonons.}$$

ζ : overlap of cell periodic part of the initial and final electron states.

ς : tensor describing the shift of electron band structure per unit deformation.

ξ : phonon polarization.

Intravalley elastic acoustic phonon scattering In silicon, intravalley phonon scattering usually involves only acoustic phonons due to the restrictive conservations of energy and momentum [13]. The acoustic phonon energy involved in this type of scattering is much smaller than the average electron energy in high electric field, so it is often treated as an elastic scattering type [4]. Herring and Vogt [12] pointed out that, for valleys along $\langle 100 \rangle$ and $\langle 111 \rangle$ directions, the deformation potential tensor has the shear component Ξ_u and the dilation component Ξ_d . For electron-acoustic phonon interactions, the deformation potential has both a transverse mode Ξ_{TA} and a longitudinal mode Ξ_{LA} , which can be expressed as [12]:

$$\Xi_{TA}(\theta) = \Xi_u \sin \theta \cos \theta \quad (2.7a)$$

$$\Xi_{LA}(\theta) = \Xi_d + \Xi_u \cos^2 \theta \quad (2.7b)$$

where θ is the angle between the phonon wave vector and the longitudinal direction of the valley. The effect of anisotropy is very small [14], so the deformation potential can be taken as the average over angle θ . The overlap factor is exactly equal to unity due to the small change in electron wave vectors $\mathbf{k}' - \mathbf{k}$ compared to the size of BZ [4]. The probability of intravalley elastic acoustic phonon scattering is [4]:

$$P(E) = \frac{\sqrt{2}k_B T_0 \zeta_1^2}{\pi \hbar^4 \rho u^2} m_d^{3/2} \sqrt{E(1 + \alpha E)} (1 + 2\alpha E) \quad (2.8)$$

where

$$m_d = (m_t m_l^2)^{1/3}$$

where m_t and m_l are electron transverse and longitudinal effective masses in Si.

k_B : Boltzmann constant.

T_0 : lattice temperature.

ζ : deformation potential.

$$u = \frac{2u_t + u_l}{3},$$

where u_t and u_l are transverse and longitudinal sound velocity in Si.

ρ	2.33 gcm ⁻³	[15]
u_t	5.3x10 ⁵ V/cm	[15]
u_l	9x10 ⁵ V/cm	[15]
m_l	0.916m ₀ (X valley) 1.59m ₀ (L valley)	[11] [11]
m_t	0.19m ₀ (X valley) 0.12m ₀ (L valley)	[11] [11]
α	0.5 eV ⁻¹	[4]
ζ	9eV	[4, 6, 11, 16, 17]

Table 2.1: Parameters used in calculating the electron-phonon intravalley scattering rate.

Intervalley phonon scattering

- X-X intervalley scattering can occur between valleys on the same symmetry axis (g-type) and between valleys on perpendicular axis (f-type). By the crystal symmetry, Streitwolf [18] and Lax and Birman [19] argued that only high-energy phonons (around 650 K) could participate in X-X intervalley scattering in silicon. However, low-energy phonons (around 200 K) were detected by magnitophonon resonance measurements [20]. Ferry [21] attributed the low-energy phonons to the first-order interaction in the phonon wave vector. Therefore, both low-energy and high-energy phonons have been included in MC simulations [4, 6, 11, and 17].

Based on the geometrical calculations, phonon wave vectors involved in transferring electrons at the bottoms of X valleys can be determined [22]. There are three g-type phonons and three f-type phonons. Since phonon wave vectors, involved in transferring electrons between the X valleys, are almost the same as the distance between the minima of the valleys [4], X-X intervalley phonon energies are treated as constants. Those phonons' energy can be retrieved from the phonon dispersion curve.

The overlap factor depends mostly on the valleys involved in the transferring process [23]. It is often treated as a constant for each scattering type and included in the coupling constants $(D_i K)^2$. The coupling constants for X-X intervalley transitions were empirically found by fitting drift velocity and the diffusion coefficients as functions of electric field strength and directions [6, 16]. Table 2.2 summarizes the X-X intervalley phonon parameters used in this work in comparison with other MC groups.

Phonon Type and Energy	Ref. [6, 11]	Ref. [4,16,17]	Ref. [24]	This work and Ref. [25]
f1 (19meV)	0.15	0.3	2.5	0.5
f2 (51meV)	3.4	2	—	3.5
f3 (57meV)	4	2	8	1.5
g1(10meV)	0.5	0.5	—	0.3
g2 (19meV)	0.8	0.8	4	1.5
g3 (62meV)	3	11	8	6

Table 2.2: X-X intervalley phonon energy and coupling constants ($\times 10^8 \text{eV/cm}$). The values from other groups' MC simulations are shown for comparison purpose.

- X-L intervalley phonons' energy can be calculated based on geometrical calculations, similar to X-X intervalley scattering. Since there is little knowledge of X-L coupling constants, they are treated as empirical fitting parameters to reproduce the experimental Π coefficients [11, 17]. There are four possible phonons in X-L intervalley scatterings [11, 17]. Table 2.3 summarizes the X-L intervalley phonon parameters used by other MC groups and used in this work.

Phonon Energy (meV)	Ref. [11]	Ref. [17]	This work
17	4	2.5	0.65
41	4	2.5	0.65
55	4	2.5	0.65
60	4	2.5	0.65

Table 2.3: X-L intervalley phonon energy and coupling constants ($\times 10^8 \text{eV/cm}$). The values from other groups' MC simulations are shown for comparison purpose.

The probability of intervalley phonon scattering derived in Ref. [4] is given by

$$P(E) = \frac{Z_f (D_t K)_i^2 m_d^{3/2}}{\sqrt{2\pi\hbar^3 \rho \omega_i}} \left[\frac{N_i}{N_i + 1} \right] \sqrt{(E \pm \hbar\omega_i - \Delta E_{fi}) [1 + \alpha(E \pm \hbar\omega_i - \Delta E_{fi})] [1 + 2\alpha(E \pm \hbar\omega_i - \Delta E_{fi})]} \quad (2.9)$$

where

Z_f : the number of equivalent final valleys.

ΔE_{fi} : the energy difference between the bottom of final valley and initial valley.

$D_t K$: coupling constant.

$N_i = \frac{1}{\exp(\frac{\hbar\omega_i}{k_B T_0}) - 1}$: the number of phonons.

ω_i : phonon frequency.

Impact ionization rate II is treated as an additional electron scattering. II usually occurs at very high electric field ($>10^5 \text{V/cm}$), and only carriers whose energy is higher than the threshold energy ($\sim E_G = 1.1 \text{eV}$) can cause II. For carriers whose energy is much higher than optical phonon energy, the momentum relaxation time is much smaller than energy relaxation time. Therefore, it is appropriate to average II rates over wave vector directions. The detailed mathematical calculations in the appendix A yield the electrons' II rate:

$$P_{II}(s^{-1}) = 9.5 \times 10^{10} (E - 1.1 \text{eV})^{4.79} \quad (2.10)$$

2.2.3 Procedures of Monte Carlo simulations

This work closely follows the algorithm developed in Ref. [4]. Figure 2.2 shows the flow chart of the MC programs for both electrons and holes.

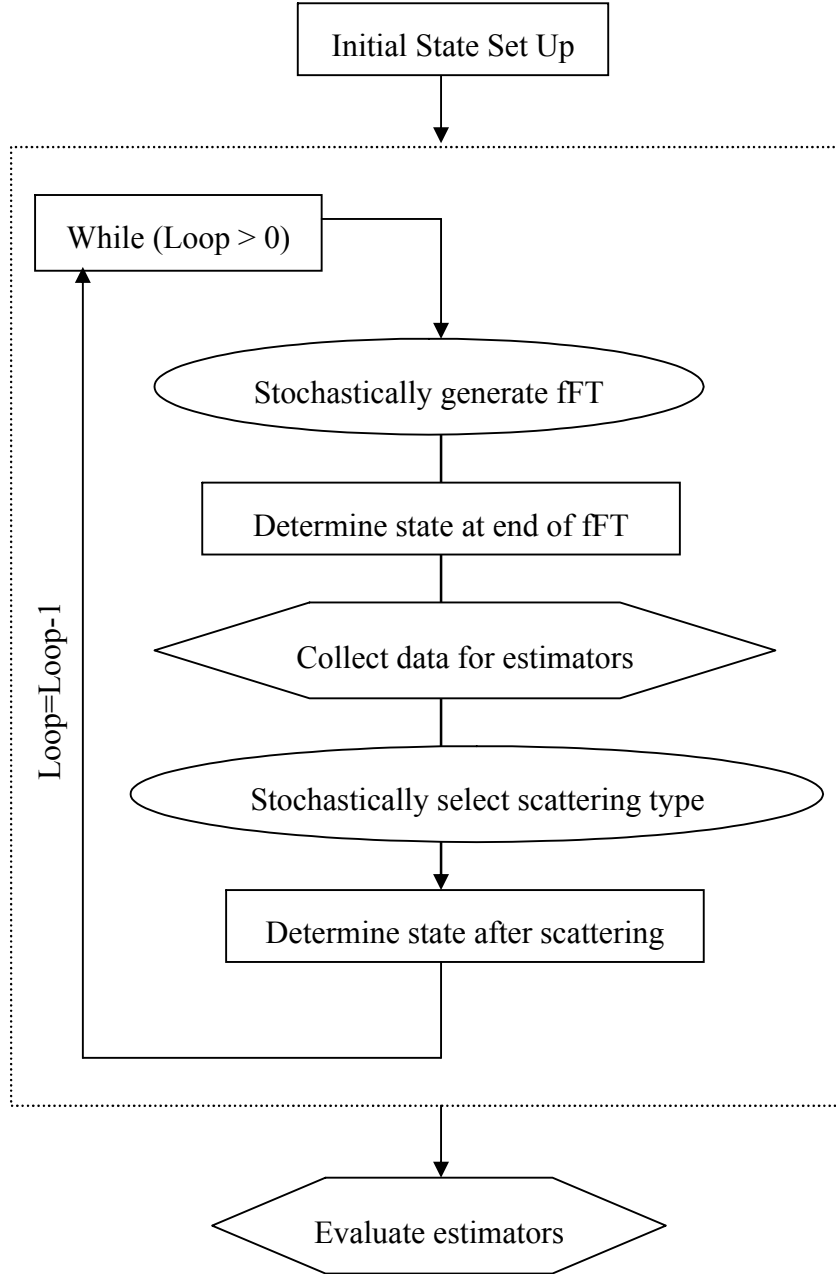


Figure 2.2: Flowchart of the MC programs for both electrons and holes.

Initial setup Appropriate energy, momentum, and positions of the initial electrons and holes are assigned, depending upon the interested transport parameters and device structures.

Stochastic generation of free-flight time The probability of an electron that starts out a free flight at time $t=0$ and experiences the next scattering event during dt at time t has the form [4]:

$$P(t)dt = P[E(t)]\exp\left(-\int_0^t P[E(t')]dt'\right)dt \quad (2.11)$$

where $P[E(t)]$ is the total electron scattering rate.

The total electron scattering rate is a function of energy. Electron energy depends on the duration of free-flight time, during which the electrons gain energy from the electric field. It requires a large number of calculations to generate stochastic free-flight time (fFT) according to the distribution function in equation (2.11). This difficulty can be easily avoided by using the self-scattering technique introduced in Ref. [26]. By introducing the self-scattering rate ($P_{self-scatter}$), the total electron scattering rate (P_{total}) is treated as a constant, independent of energy:

$$P_{total} = \sum P_{e-phonon}(E) + P_{II}(E) + P_{self-scatter}(E) \quad (2.12)$$

So the fFT can easily be generated stochastically by

$$fFT = -\frac{1}{P_{total}} \log(rand) \quad (2.13)$$

At the end of a fFT, if a self-scattering event is chosen among all of the scattering types, then the electron's state before and after the scattering remains unchanged. P_{total} needs to be larger than the maximum electron scattering probability during the entire simulation. If P_{total} is chosen to be

smaller than total electron scattering, the generated fFT may be artificially made longer than it should be. On the other hand, if P_{total} is too large, the MC program wastes time handling self-scattering cases.

Determination of state before scattering (at the end of a free-flight duration)

For

most of this work's MC simulations, only electrons' wave vector and energy are needed to prepare for the next step. An electron's wave vector is composed of the parallel and perpendicular components with respect to the electric field. During free-flight time, only the parallel component of the wave vector changes:

$$k_{\parallel}(t = fFT) = k_{\parallel}(t = 0) + HV \frac{q}{\hbar} \int_0^{fFT} \varepsilon dt \quad (2.14a)$$

$$k_{\perp}(t = fFT) = k_{\perp}(t = 0) \quad (2.14b)$$

where HV : Herring-Vogt transformation constant.

Electron energy is calculated as a function of wave vector:

$$E = \frac{\sqrt{1 + 2\alpha \hbar^2 |\mathbf{k}|^2 / m_0} - 1}{2\alpha} \quad (2.15)$$

Data collection for estimators

Different transport parameters can be extracted at this step, depending on the goals of the MC simulations. For example, average drift velocity and distance traveled can be calculated. Derivation of average drift velocity is shown as follows:

$$v[\mathbf{k}(t)] = \frac{1}{\hbar} \frac{\partial E}{\partial \mathbf{k}} \quad (2.16)$$

$$\partial \mathbf{k} = \frac{1}{\hbar} e \mathcal{E} dt \quad (2.17)$$

$$\bar{v} = \frac{1}{fFT} \frac{E(t = fFT) - E(t = 0)}{q\mathcal{E}} \quad (2.18)$$

The distance traveled within the duration of fFT can be calculated as follows.

$$dx = v(t) dt \quad (2.19)$$

Substituting equation (2.18) into equation (2.19) gives

$$dx = \frac{1}{\hbar} \frac{\partial E(t)}{\partial \mathbf{k}} dt \quad (2.20)$$

Substituting equation (2.17) into equation (2.20) gives

$$dx = \frac{\frac{1}{\hbar} \partial E(t) dt}{\frac{1}{\hbar} q \mathcal{E} dt} \quad (2.21)$$

or

$$dx = \frac{\partial E(t)}{q\mathcal{E}} \quad (2.21a)$$

Taking integration of equation (2.21a) yields

$$x = \frac{E(t = fFT) - E(t = 0)}{q\mathcal{E}} \quad (2.22)$$

Stochastic selection of scattering type After calculating the energy at the end of fFT, the value of each scattering rate can be determined. The discrete case technique [4] is then used to select a specific scattering type. If all realistic scattering types are not chosen, a self-scattering event occurs. In a self-scattering event, the electron's state before and after the collision remains unchanged.

Determination of state after scattering (at the beginning of the next free flight)

Determination of electrons' energy after collisions is trivial due to the energy conservation:

$$E' = E \pm \hbar\omega_q - \Delta E_{fi} \quad (2.23)$$

where E' and E are the energy values after and before a scattering event; $\hbar\omega_q$ is the phonon energy; ΔE_{fi} is the energy difference between the bottom of final valley and initial valley.

The determination of an electron's wave vector after a scattering event is a little more involving. Because of the isotropic characteristic of all phonon scattering mechanisms discussed above, any new state $\mathbf{k}'$ that belongs to the energy-conserving sphere has equal probability. The probability of $\mathbf{k}'$ forming an angle θ with $\mathbf{k}$ is proportional to the available states on the circumference of the circle. The circle's radius is $k' \sin \theta$ (figure 2.3).

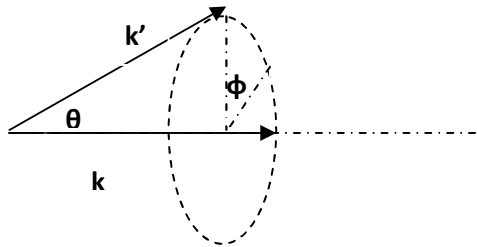


Figure 2.3: Geometry for the determination of the new state after a scattering event. The direction of new state $\mathbf{k}'$ is stochastically generated relative to state $\mathbf{k}$ before the collision.

Applying the direct technique [4], θ can be chosen by the relationship:

$$r_1 = \frac{\int_0^\theta k' \sin \theta d\theta}{\int_0^\pi k' \sin \theta d\theta} \quad (2.24)$$

where r_1 is a random number between [0,1].

Solving equation (2.24) gives

$$\cos \theta = 1 - 2r_1 \quad (2.24a)$$

$$\phi = 2\pi r_2 \quad (2.24b)$$

where r_2 is a random number between [0,1].

2.2.4 Results of parameter calibrations for MC simulations of electrons

MC parameter calibration is usually done through fitting bulk transport parameters. Figure 2.4 shows this work's calculations of drift velocity as a function of electric field in comparison with the experimental data. Drift velocity was experimentally measured at electric field up to 10^5 V/cm [5, 27]. The calculated results match the measured data well in homogeneous electric field oriented in different crystallographic directions. The theoretical results correctly show that drift velocity in $\langle 111 \rangle$ direction is higher than in $\langle 100 \rangle$. This helps confirm the accurate anisotropy of the CB model.

For electric field $>10^5$ V/cm, MC calibration is usually done by fitting experimental II coefficients. Figure 2.5 shows that this work's calculated electron II coefficient agrees well with the experimental data [28-31]. There is a certain disagreement of the II coefficients reported by different groups, especially at low electric field. One of the main factors for the disagreement

among the experimental data is the dead-space effect. The electron and hole dead-space and its effect on the techniques of experimentally extracting II coefficients will be discussed in details in chapter 4.

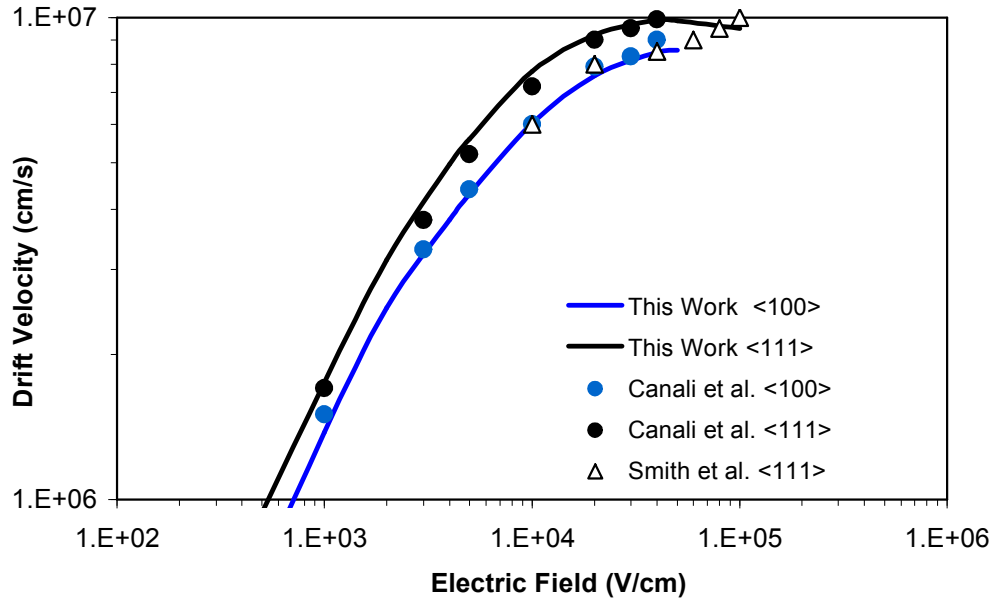


Figure 2.4: Experimental [6, 27] and simulated electron drift velocity as a function of the electric field along <100> and <111> crystallographic directions.

Being able to fit both experimental drift velocity and II coefficient still does not ensure the accuracy of either phonon scattering rate or II rate. Tang and Hess [17] demonstrated that both experimental drift velocity and II coefficient can be fit with different sets of phonon scattering rates and II rates. The workers in Ref. [17, 32] argued that the uniqueness of phonon scattering rate and II rate can only be justified if a MC program can also predict experimental data that reflects the electron distribution function beside drift velocity and II coefficient. Cartier *et al.* [32] claimed to have determined the unique electron-phonon scattering rate and the unique

II rate by fitting all possible high field experimental findings including XPS 2p-core level lines, II coefficient, and quantum yield.

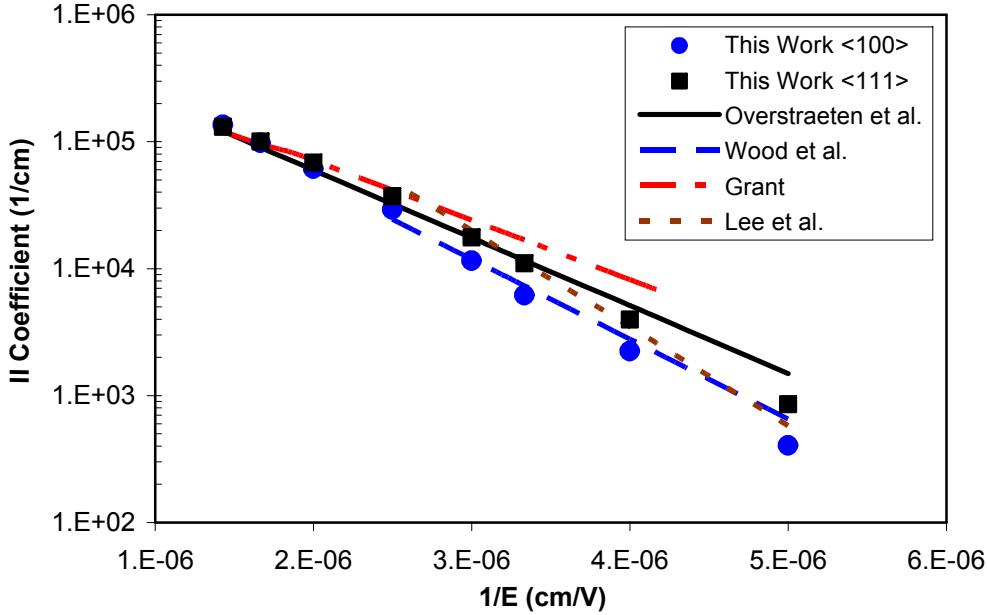


Figure 2.5: Experimental and simulated electron II coefficients as a function of electric field.

For electron-phonon scattering rate, Cartier *et al.* used the rate reported in Ref. [9]. Figure 2.6 shows that the total phonon scattering rate in this work agrees well with that of full-band MC simulations [9, 33]. The positions of peaks and valleys correspond to the realistic DOS. [10].

For II rate, many different MC groups [9, 11, 17] used the simple, well-known Keldysh II formula [34]:

$$P_{II} = P_0(E - E_T)^2 \quad (2.25)$$

Cartier *et al.* [32] extracted the II rate from their experiment using soft x-ray photoemission spectroscopy. They experimentally showed that II rate derived by Keldysh [34] is not accurately applicable to Si. This is because Keldysh derived the II rate based on the simple isotropic

parabolic CB model. Cartier *et al.* [32] used multiple terms of Keldysh's II formula to fit their experimental data:

$$P_{II} = \sum_{i=1}^3 P_i (E - E_{Ti})^2 \quad (2.26)$$

where $E_{Ti}=1.2, 1.8, 3.45$ eV for $i=1,2,3$, respectively.

Figure 2.6 shows that the II rate, derived in this work by full-band calculations, agrees well with Cartier's experimental result [32]. This work's II rate also agrees with the II rate from Ref. [33], which was the result of full-band calculations as well.

In short, this work's phonon scattering rate and II rate agree well with the uniqueness of the phonon scattering rate and II rate extracted in Ref. [32].

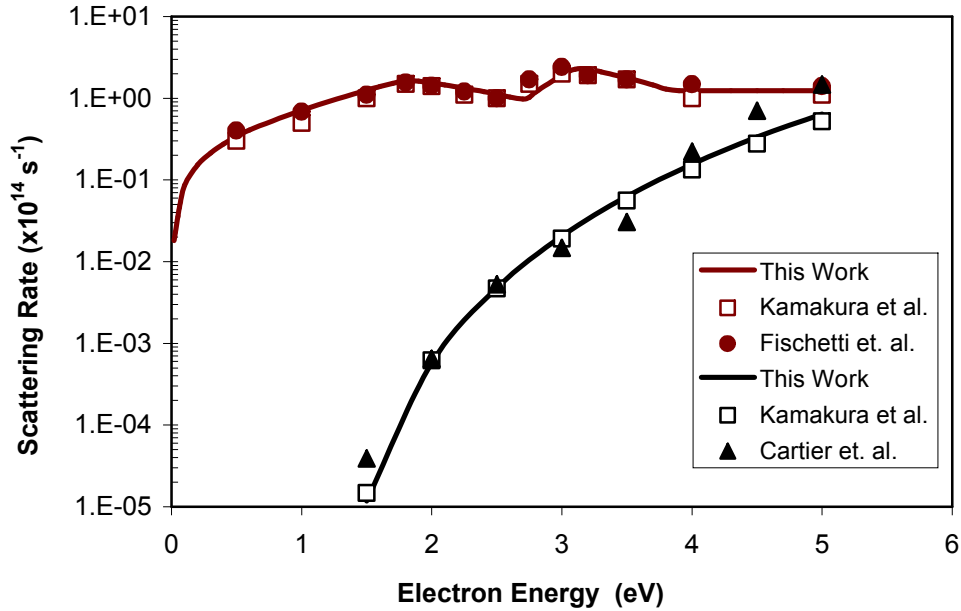


Figure 2.6: Total electron-phonon scattering rate and electron II rate as functions of energy. Rates from previous MC simulations are shown for comparison [9, 32-33]. Zero energy corresponds to the maximum point of the valance band edge.

After comparing many prominent MC models (figure 17 by Higman and Hess, figure 18 by Fischetti and Laux, and figure 20 by Kunikiyo *et al.* in Ref. [35]), the authors in Ref. [35] concluded that all MC programs sufficiently employing accurate conduction band model produce similar results for both electron-phonon scattering rate and electron energy distribution function in high homogeneous electric field. Figure 2.7 shows that this work's electron energy distribution agrees well with the results from full-band MC simulations (figure 17 by Higman and Hess, figure 18 by Fischetti and Laux, and figure 20 by Kunikiyo *et al.* in Ref. [35]), at the electric field equal to $3 \times 10^5 \text{V/cm}$.

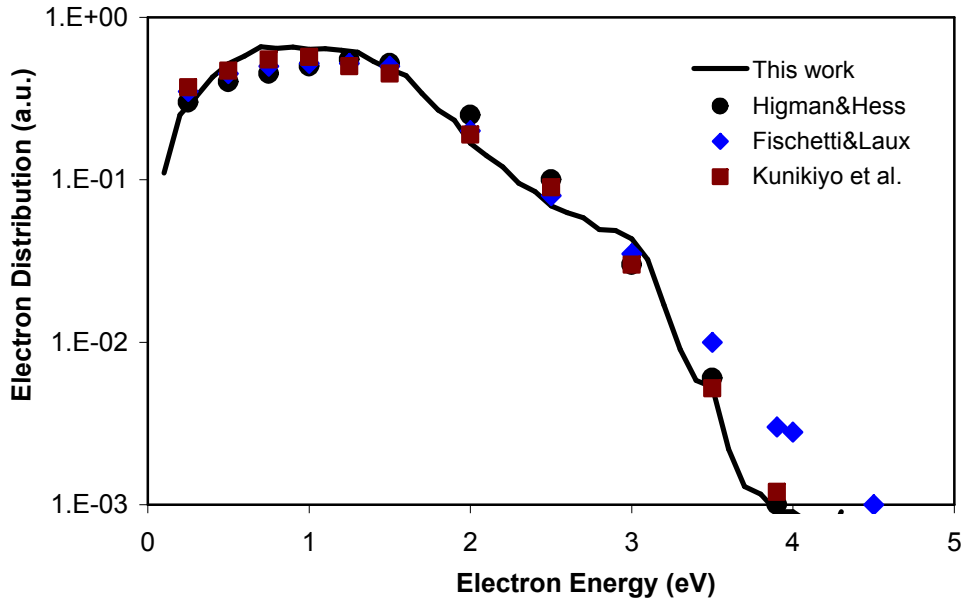


Figure 2.7: Electron distribution function at electric field = 300kV/cm in $\langle 100 \rangle$ crystallographic direction. Dots are MC results from other groups (figure 17 by Higman and Hess, figure 18 by Fischetti and Laux, and figure 20 by Kunikiyo *et al.* in Ref. [35]). Zero energy corresponds to the minimum point of the conduction band edge.

Average electron energy as a function of electric field is also simulated and compared with other MC simulations (figure 2.8). The average electron energy agrees with both the analytical band MC in Ref. [6] and the full band MC in Ref. [9] for electric field strength less than 10^5 V/cm. At electric field higher than 10^5 V/cm, this result is a little smaller than the full band MC result in Ref. [9]. This is due to the fact that the II rate used in Ref. [9] is smaller than the II rate derived in this work. The authors in Ref. [9] modified their II rate in their later work [32], and their new II rate agrees more closely with this work's.

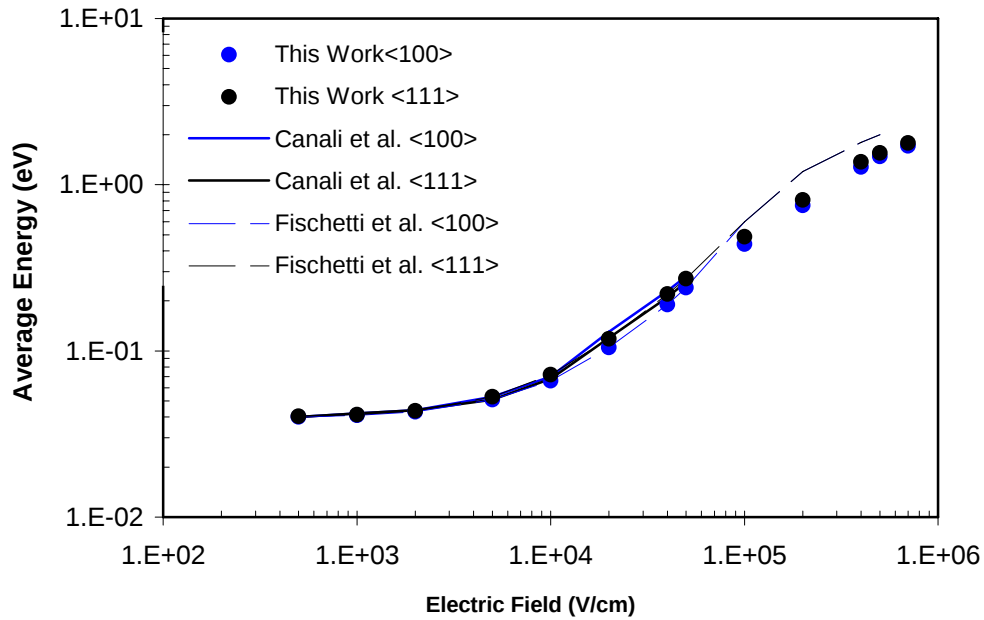


Figure 2.8: Electron average energy as a function of electric field. For comparison, results from previous MC simulations are shown (Canali *et al.* [6]: analytical band MC, Fischetti *et al.* [9]: full band MC).

Quantum yield is also simulated to further justify the accuracy of this MC program for electrons. Quantum yield is calculated by releasing electrons with a certain value of initial energy in zero electric field and counting the total number of II events the electrons experience until the energy is lower than the II threshold energy. Figure 2.9 shows a good agreement between this work's result and the experimental data in Ref. [36-38].

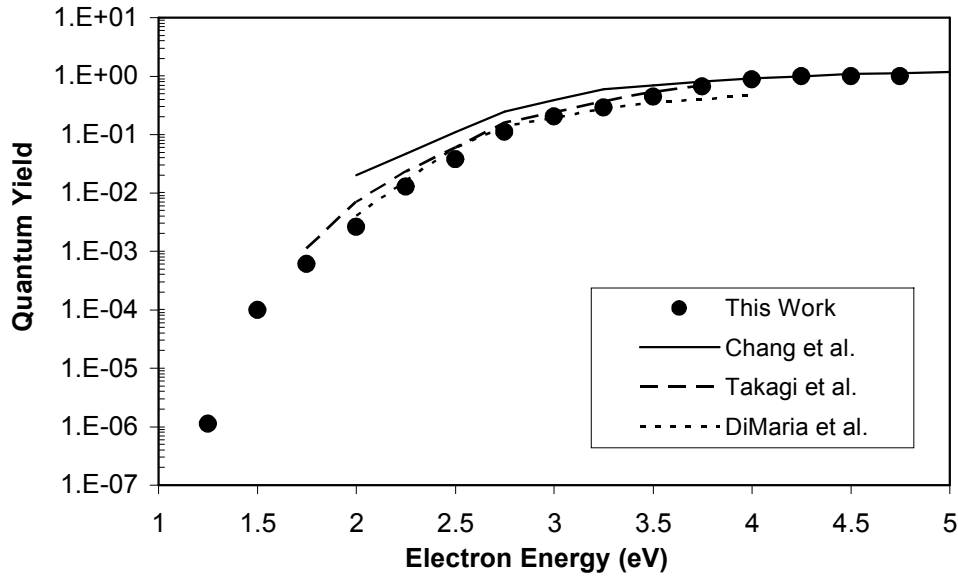


Figure 2.9: Experimental [36-38] and simulated quantum yield as a function of electron energy. Zero energy corresponds to the maximum point of the valance band edge.

2.3 Monte Carlo model for holes

Compared to electrons, MC simulations of hole transport in Si are still yet to be mature. Reports on the basic hole transport parameters such as drift velocity, average energy, and μ coefficients are still very different even among different full-band MC groups [39-42]. The average hole energy as a function of electric field is different, especially at high electric field, among Ref. [39, 41]. For example, at $E=3 \times 10^5 \text{V/cm}$, the average hole energy reported in Ref. [39] is 50% smaller than that in Ref. [41]. While the average hole energy agrees well among Ref. [40] and Ref. [41], their calculated results of hole saturation velocity are different. While the average hole energy in Ref. [39] agrees satisfactorily with the energy reported in Ref. [42], the simulated hole velocity shows a much more severe negative resistance for the electric field higher than 10^5V/cm in Ref. [42]. While there is a good agreement of the hole μ coefficients reported in Ref. [41] and Ref. [42], their calculated values of average hole energy are markedly different. Clearly, detailed descriptions of the valance band in the full-band MC programs do not ensure the accurate results. On the other hand, analytical band MC simulations have been shown to satisfactorily produce the experimental drift velocity and the diffusion coefficient [43-45].

In this work, a single warped parabolic analytical band model [43-44] is used:

$$E = (1 - \beta)ak^2 \{1 - [b^2 + c^2(\sin^4 \vartheta \sin^2 \varphi \cos^2 \varphi + \sin^2 \vartheta \cos^2 \vartheta)]^{0.5}\} \quad (2.27)$$

where β is a warping parameter; $a = \hbar^2 \frac{|A|}{2m_0}$; $b = \frac{|B|}{|A|}$; $c = \frac{|C|}{|A|}$; ϑ and φ are polar and azimuthal angles of $\mathbf{k}$ with respect to crystallographic axes.

The probability of the isotropic acoustic phonon scattering (P_{ac}) and the isotropic optical phonon scattering (P_{op}) are [43-44]:

$$P_{ac}(E) = \frac{k_B T_0 (E_1^0)^2 m_D^{3/2}}{\sqrt{2\pi} \hbar^4 \rho u^2} \sqrt{E} \quad (2.28a)$$

$$P_{op}(\varepsilon) = \frac{(D_t K)^2 m_D^{3/2}}{2\sqrt{2\pi} \hbar^3 \rho \omega_{op}} \left[\frac{N_{op}}{N_{op} + 1} \right] \sqrt{E \pm \hbar \omega_{op}} \quad (2.28b)$$

where E_1^0 is the acoustic deformation potential.

$$m_D^{3/2} = m_0^{3/2} \frac{1}{4\pi A^{3/2} (1-\beta)^{3/2}} \int_0^{2\pi} \int_0^\pi [1 - g(\vartheta, \varphi)]^{-3/2} \sin \vartheta d\varphi d\vartheta \quad (2.29)$$

$$\text{where } g(\vartheta, \varphi) = [b^2 + c^2 (\sin^4 \vartheta \sin^2 \varphi \cos^2 \varphi + \sin^2 \vartheta \cos^2 \vartheta)]^{0.5} \quad (2.29a)$$

Similar to electrons, phonon scattering rates at high energy are weighted according to the DOS (Appendix A). All of the phonon parameters and the valance band parameters are listed in table 2.4.

II is treated as an additional scattering besides phonon scatterings. The derived II rate is (see Appendix A):

$$P_{II} = 1.4 \times 10^{11} (\varepsilon - 1.2 eV)^2 + 3.1 \times 10^{11} (\varepsilon - 1.3 eV)^{4.72} \quad (2.30)$$

The II rate at low energy is adjusted in order for the MC program to produce drift velocity, average energy, and II coefficients in good agreement with the data from both experiments and full-band MC simulations.

Symbols	Descriptions	This Work	Ref. [43]
β	Warping parameter	0.2	0.69
A	VB parameter	-4.27	-4.27
B	VB parameter	-0.55	-0.63 ± 0.08
C	VB parameter	4.96	-
E_1^0	Acoustic deformation potential	2 (eV)	2.2 (eV)
$\hbar\omega_{op}$	Optical phonon energy	63.3(meV)	63.3(meV)
$D_r K$	Coupling constant	5.2×10^8 (eV/cm)	5.0×10^8 (eV/cm)

Table 2.4: Parameters used in the Monte Carlo program for holes.

Figure 2.10 shows the hole II rate and the total hole-phonon scattering rate taking into account of the DOS. Drift velocity agrees well with the experimental data [43] with the electric field in different crystallographic directions (figure 2.11). At high electric field, where the experimental drift velocity is not available, this work's result agrees satisfactorily with the result from the full-band MC simulations [40]. Hole average energy is plotted as a function of the applied electric field and agrees to a good extent with other authors' calculations [39-41] (figure 2.12). This work's calculated value of hole II coefficient is plotted in figure 2.13. The experimental hole II coefficients also vary among different groups [28-31]. At low electric field, this work's hole II coefficient agrees more with the data from Woods *et al.* [31]. At higher electric field, this work's result agrees more with the data from Grant and Overstraeten *et al.* [29-30].

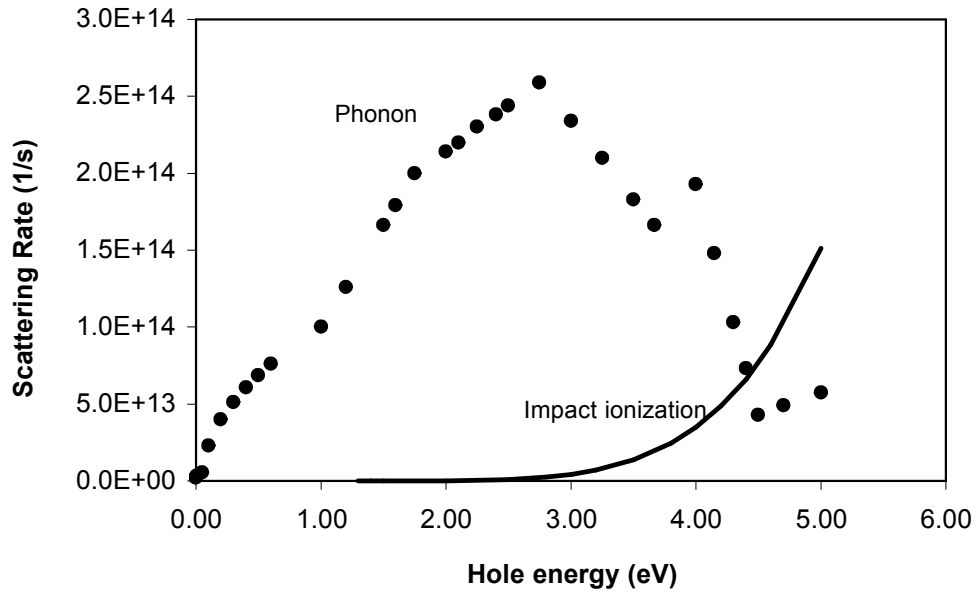


Figure 2.10: Hole II rate and total phonon scattering rate weighted by DOS.

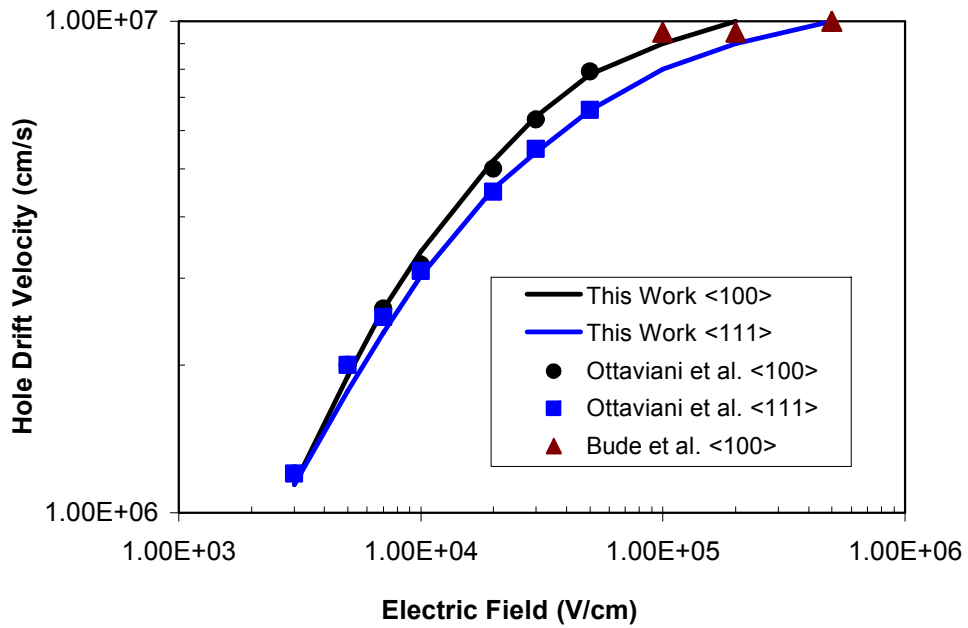


Figure 2.11: Hole drift velocity as a function of electric field along the <100> and <111> directions. Shown for comparison are the results of full-band MC simulations in Ref. [40] and the experimental results [43].

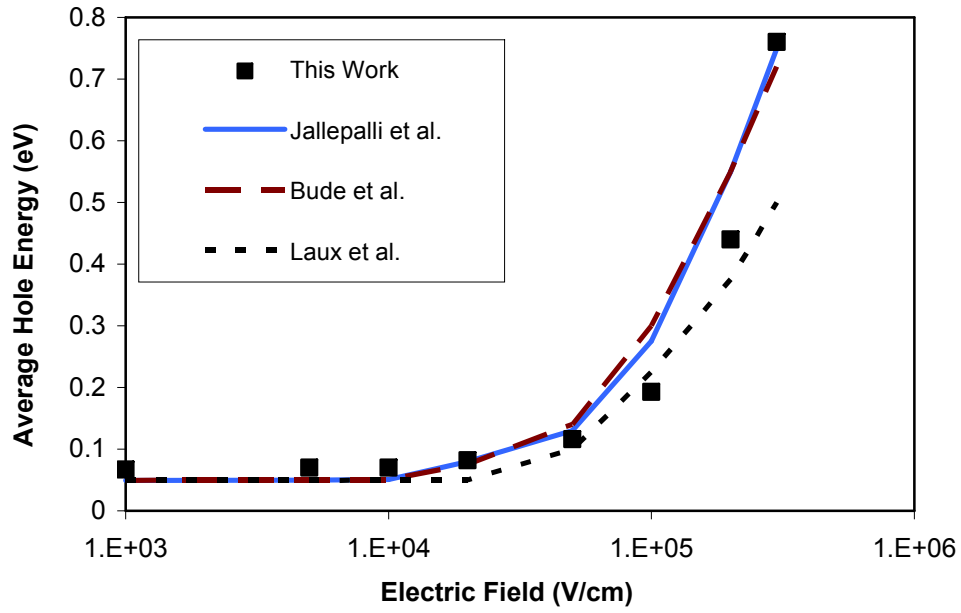


Figure 2.12: Hole average energy as a function of electric field. Results from previous MC simulations [39-41] are shown for comparison.

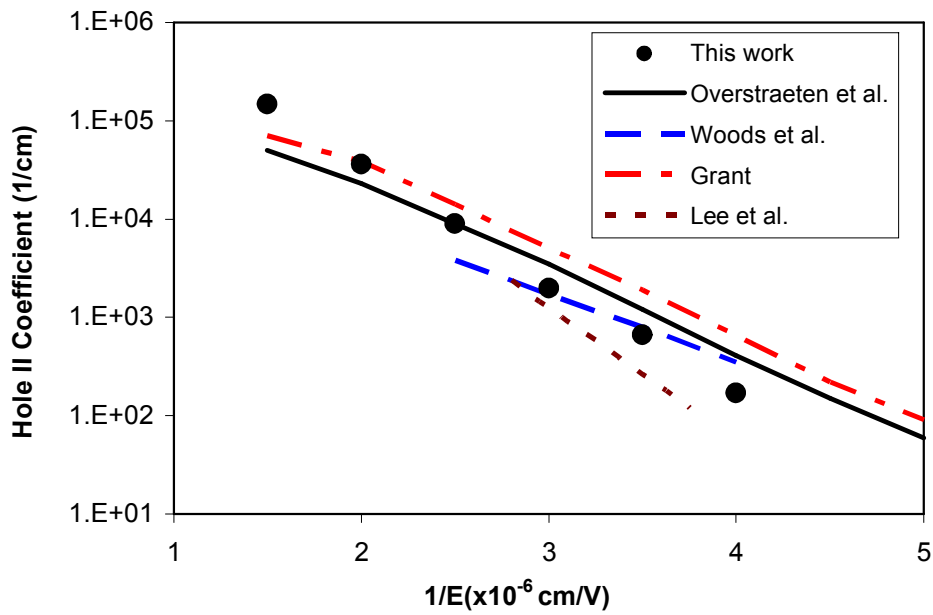


Figure 2.13: Experimental [28-31] and simulated hole II coefficients as a function of electric field.

2.4 Summary

A MC program for electrons is developed using the analytical elliptical, non-parabolic CB model. Six X valleys and eight L valleys are included to account for high-energy electrons (energy higher than II threshold energy ~ 1.1 eV). The electron-phonon scattering types involve acoustic intravalley, acoustic and optical X-X intervalley, and X-L intervalley scatterings. All of the phonon scattering types are isotropic. There are six phonons in X-X intervalley scattering (three f-type and three g-type) and four phonons in X-L intervalley scattering. The probability of phonon scattering is adjusted according to the DOS. The parameters of the MC program for electrons are carefully chosen and adjusted by extensively fitting both the theoretical and experimental transport parameters. This work's total electron-phonon scattering probability and II rate agree well with other groups' [9, 32-33]. Besides, the MC program for electrons is able to fit both the average energy and the energy distribution function to the results from Ref. [6, 9, 35]. In addition, with the same set of parameters, this work's calculated electron drift velocity, II coefficient, and quantum yield agree well with the experimental data [28-31, 36-38].

A MC program for holes is also developed using a single analytical warped parabolic band model to describe the valance band. The probability of the isotropic acoustic and optical phonon scattering is weighted by the hole DOS. The parameters for the MC program are calibrated by fitting the theoretical average energy [39-41], the experimental drift velocity [43], and the experimental II coefficients [28-31]. Compared to electrons, the MC program for holes is not as well-calibrated due to the complex valance band and due to the fact that both the theoretical and the experimental data are not as maturely-established.

Bibliography

- [1] T. Kurosawa, "Monte Carlo calculation of hot-electron problems," *Journal of Physical Society of Japan*, Vol. 21, pp. 424, 1966.
- [2] W. Fawcett, A.B. Boardman, and S. Swain, "Monte Carlo determination of electron transport properties in Gallium Arsenide," *Journal of Physics and Chemistry of Solids*, Vol. 31, pp. 1963, 1970.
- [3] K. Tomizawa, Y. Awano, N. Hashizume, and M. Kawashima, "Monte Carlo simulation of submicron GaAs $n^+-i(n)-n^+$ diode," *Solid State and Electron Devices*, Vol. 129, pp. 131, 1982.
- [4] Carlo Jacoboni and Lino Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Reviews of Modern Physics*, Vol. 55, pp. 645, 1983.
- [5] P.J. Price, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1979), Vol. 14, p. 249.
- [6] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta, "Electron drift velocity in silicon," *Physical Review B*, Vol. 12, pp. 2265, 1975.
- [7] C. Jacoboni, R. Minder and G. Majni, "Effects of band non-parabolicity on electron drift velocity in silicon above room temperature," *Journals of Physics and Chemistry of Solids*, Vol. 36, pp. 1129, 1975.
- [8] E. M. Conwell and M. O. Vassel, "High-Field Transport in n-Type GaAs," *Physical Review*, Vol. 166, pp. 797, 1968.
- [9] M.V. Fischetti and S.E. Laux, "Monte Carlo Analysis of Electron Transport in Small Semiconductor Devices Including Band-Structure and Space-Charge Effects," *Physical Review B*, Vol. 38, No.14 , pp. 9721, 1988.
- [10] James R. Chelikowsky and Marvin L. Cohen, "Electronic structure of silicon," *Physical Review B*, Vol. 10, pp. 5095, 1974.
- [11] Nobuyuki Sano, Takahiro Aoki, Masaaki Tomizawa, and Akira Yoshii, "Electron transport and impact ionization in Si," *Physical Reivew B*, Vol. 41, pp. 12122, 1990.
- [12] Convers Herring and Erich Vogt, "Transport and Deformation-Potential Theory for Many-Valley Semiconductors with Anisotropic Scattering," *Physical Review*, Vol. 101, No. 3, 1956.

- [13] C. Hamaguchi, *Basic Semiconductor Physics* (Springer, New York, 2001).
- [14] H. Mizuno, K. Taniguchi, and C. Hamaguchi, "Electron-transport simulation in silicon including anisotropic phonon scattering rate," *Physical Review B*, Vol. 48, pp. 1512, 1993.
- [15] C. Jacoboni and L. Reggiani, "Bulk hot-electron properties of cubic semiconductors," *Advances in Physics*, Vol. 28, pp. 493, 1979.
- [16] R. Brunetti, C. Jacoboni, F. Nava, L. Reggiani, G. Bosman, and R. J. J. Zijlstra, "Diffusion coefficient of electrons in silicon," *Journal of Applied Physics*, Vol. 52, pp. 6713, 1981.
- [17] J. Y. Tang and Karl Hess, "Impact ionization of electrons in silicon (steady state)," *Journal of Applied Physics*, Vol. 54, pp. 5139, 1983.
- [18] H.W. Streightwolf, "Intervalley Scattering Selection Rules for Si and Ge," *Physica Status Solidi B*, Vol. 37, K47 (1970).
- [19] M. Lax and J. L. Birman, "Intervalley scattering selection rules for Si and Ge," *Physica Status Solidi B*, Vol. 49, K153 (1972).
- [20] J.C. Portal, L. Eaves, S. Askenazy, and R.A. Stradling, Proceedings of the 12th International Conference on the Physics of Semiconductors, Stuttgart, edited by M. H. Pilkuhn (Teubner, Stuttgart, 1974), p. 259.
- [21] D.K. Ferry, "First-order optical and intervalley scattering in semiconductors," *Physical Review B*, Vol. 14, pp. 1605, 1976.
- [22] D. Long, "Scattering of Conduction Electrons by Lattice Vibrations in Silicon," *Physical Review*, Vol. 120, pp. 2024 (1960).
- [23] L. Reggiani and C. Calandra, "Bloch states mixing in Si conduction band," *Physics Letter A*, Vol. 43, pp. 339, 1973.
- [24] T. Yamada, J.-R Zhou, H. Miyata, and D. K. Ferry, "In-Plane Transport Properties of Si/Si_{1-x}Ge_x Structure and its FET Performance by Computer Simulation," *Transactions on Electron Devices*, Vol. 41, pp. 1513, 1994.
- [25] Eric Pop, Robert W. Dutton, and Kenneth E. Goodson, "Analytic band Monte Carlo model for electron transport in Si including acoustic and optical phonon dispersion," *Journal of Applied Physics*, Vol. 96, No. 9, 2004.
- [26] H.D. Rees, "Calculation of steady state distribution functions by exploiting stability," *Physics Letters A*, Vol. 26, pp. 416, 1968.

- [27] P.M. Smith, M. Inoue, and Jeffrey Frey, "Electron velocity in Si and GaAs at very high electric fields," *Applied Physics Letters*, Vol. 37, pp. 797, 1980.
- [28] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Holes and Electrons in Silicon," *Physical Review*, Vol. 134, No. 3A, 1964.
- [29] R. Van Overstraeten and H. De Man, "Measurement of the ionization rates in diffused silicon p-n junctions," *Solid-State Electronics*, Vol. 13, pp. 583-608, 1970.
- [30] W. N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields," *Solid-State Electronics*, Vol. 16, pp. 1189, 1973.
- [31] M. H. Woods, W. C. Johnson and M. A. Lampert, "Use of a Schottky Barrier to Measure Impact Ionization Coefficients in Semiconductors," *Solid-State Electronics*, Vol. 16, pp. 381, 1973.
- [32] E. Cartier, M. V. Fischetti, E. A. Eklund, and F. R. McFeely, "Impact ionization in silicon," *Applied Physics Letters*, Vol. 62, pp. 3339, 1993.
- [33] Y. Kamakura, H. Mizuno, M. Yamaji, M. Morifuji, K. Taniguchi, C. Hamaguchi, T. Kunikiyo, and M. Takenaka, "Impact Ionization Model for Full Band Monte Carlo Simulation," *Journal of Applied Physics*, Vol. 75, pp. 3500, 1994.
- [34] Keldysh L.V., "Kinetic Theory of Impact Ionization in Semiconductors," *Soviet Physics-JETP*, Vol. 37, pp 509, 1960.
- [35] A. Abramo, L. Baudry, R. Brunetti, R. Castagné, M. Charef, F. Dessenne, P. Dollfus, R. Dutton, W. L. Engl, R. Fauquembergue, C. Fiegna, M. V. Fischetti, S. Galdin, N. Goldsman, M. Hackel, C. Hamaguchi, K. Hess, K. Hennacy, P. Hesto, J. M. Higman, T. Iizuka, C. Jungemann, Y. Kamamura, H. Kosina, T. Kunikiyo, S. E. Laux, H. Lin, C. Maziar, H. Mizuno, H. J. Peifer, S. Ramaswamy, N. Sano, P. G. Scobohaci, S. Selberherr, M. Takenaka, T.-W. Tang, K. Taniguchi, J. L. Thobel, R. Thoma, K. Tomizawa, M. Tomizawa, T. Vogelsang, S.-L-Wang, X. Wang, C.-S. Yao, P.D. Yoder, and A. Yoshii, "A Comparison of Numerical Solutions of the Boltzmann Transport Equation for High-Energy Electron Transport Silicon," *Transactions on Electron Devices*, Vol. 41, pp. 1646, 1994.
- [36] C. Chang, C. Hu, and R. W. Brodersen, "Quantum yield of electron impact ionization in silicon," *Journal of Applied Physics*, Vol. 57, pp. 302, 1985.
- [37] D. J. DiMaria, T. N. Theis, J. R. Kirtley, F. L. Pesavento, D. W. Dong, and S. D. Brorson, "Electron heating in silicon dioxide and off-stoichiometric silicon dioxide films," *Journal of Applied Physics*, Vol. 57, pp. 1214, 1985.
- [38] S. Takagi and A. Toriumi, Extended Abstracts of *The Japan Society of Applied Physics*, Spring Meeting, 1993, Abstract No. 31-ZX-10 (in Japanese).

- [39] S. E. Laux, M. V. Fischetti and D. J. Frank, "Monte Carlo Analysis of Semiconductor Devices: The DAMOCLES Program", *IBM Journal of Research and Development*, Vol. 34, pp. 466-494, July 1990.
- [40] J. Bude and R.K. Smith, "Phase-space simplex Monte Carlo for semiconductor transport", *Semiconductor Science and Technology*, Vol. 9, pp. 840-843, 1994.
- [41] S. Jallepalli, M. Rashed, W.-K. Shih, C. M. Maziar, and A.F. Tasch, Jr., "A full-band Monte Carlo model for hole transport in silicon," *Journal of Applied Physics*, Vol. 81, No. 5, pp. 2250, 1997.
- [42] K. Roze, N. A. Bannov, K. W. Kim, W. C. Holton, and M. A. Littlejohn, "Temperature dependence of impact ionization coefficients in p-Si," *Journal of Applied Physics*, Vol. 83, pp. 4988, 1998.
- [43] G. Ottaviani, L. Reggiani, C. Canali, F. Nava, and A. Alberigi-Quaranta, "Hole drift velocity in silicon," *Physical Review B*, Vol. 12, No. 8, 1975.
- [44] Lino Reggiani and Rossella Brunetti, "Diffusion coefficient of holes in silicon by Monte Carlo simulation," *Journal of Applied Physics*, Vol. 59, pp. 1212, 1986.
- [45] Jim Dewey and M. A. Osman, "Monte Carlo study of hole transport in silicon," *Journal of Applied Physics*, Vol. 74, pp. 3219, 1993.

Chapter 3

A novel approach to Monte Carlo simulations of an impact ionization process in semiconductor devices

3.1 Introduction

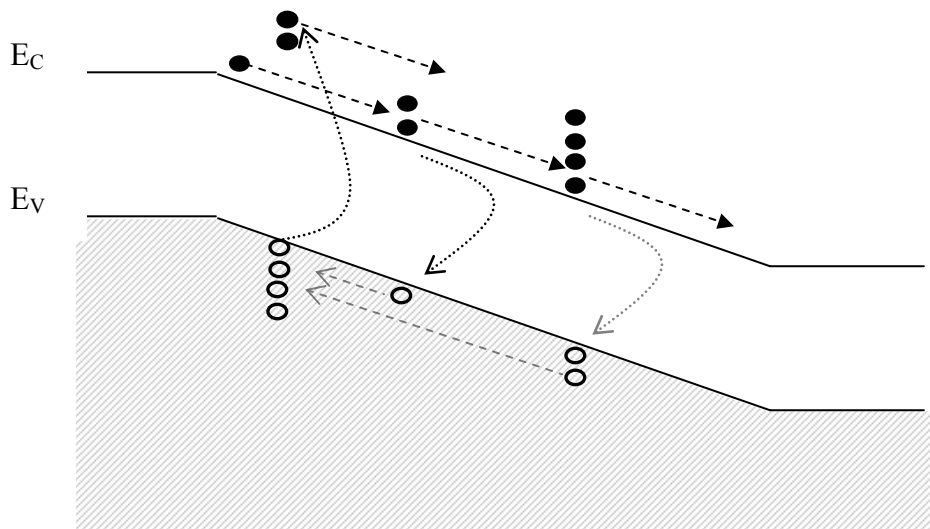


Figure 3.1: Impact ionization process in a high field region. Electrons in the conduction band impact-ionize and generate holes in the valence band. Secondary holes then impact-ionize and generate electrons in the conduction band.

An II process starts out with primary carriers, let say electrons, traveling downfield and gaining energy from the electric field. II can only occur after electrons gain energy higher than the II threshold energy (E_T). When they impact-ionize, the electron offspring join the original electrons in the new II process. The hole offspring transverse in the opposite direction and generate electrons by II. Those newly-generated electrons travel downfield, gain energy, and impact-ionize. At a high enough applied bias, the continuous loops of electrons and holes' II will cause avalanche breakdown. The interactions of electrons' and holes' II inside a high field region are shown in figure 3.1.

MC simulations are usually executed by following the flight trajectories of the electrons and holes. To simulate an II process, both MC programs for electrons and holes need to be run in such a way that the output of one program can be constantly fed into the other and vice versa. However, it is not practical, if not impossible, to keep track of the positions, momentum, and energy of the original electrons and their offspring and the offspring' offspring and so on in an II process, let alone constantly feeding their outputs to the other program's inputs. This complex inter-dependence of electrons' and holes' II can be decoupled by a positive feedback model in short devices. Consequently, with a single run of MC programs for electrons and holes subsequently, not only can the total II current gain in a device be determined, but the knowledge of the II current gain due to each type of carriers is also available. The knowledge of the contribution from each type of carriers to the total current gain is useful for device designers. To the best of my knowledge, this is the first time an II process, involving both electrons and holes, is modeled and formulated as a positive feedback loop and applied to MC simulations.

3.2 Positive feedback model

Figure 3.2 shows a positive feedback loop modeling II in a high field region. M_n is the electron current gain of the high field region, and M_p is the hole current gain of the high field region. The positive loop gain (M_n , M_p) captures the essential mechanism of the entire II process inside the high field region.

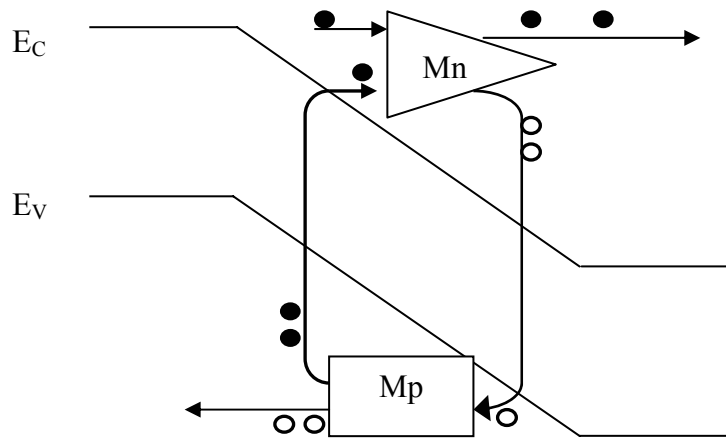


Figure 3.2: A simplified process of II in a high field region. M_n is the II gain due to electrons, and M_p is the II gain due to holes.

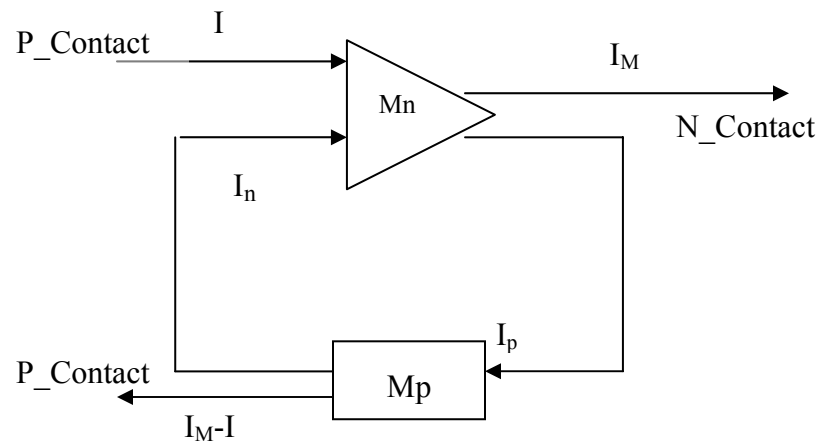


Figure 3.3: Block diagram of the positive feedback model for the II process in a p-n junction.

Derivation of the II gain with this model is trivial and typical of any positive feedback loop. Figure 3.3 is a block diagram describing electron and hole II in a p-n junction. I is the initiating electron current; I_p is the secondary hole current after the gain stage M_n . I_n is the secondary electron current after the gain stage M_p . I_M is the terminal current at the contacts. The total current gain can be formulated by the following steps:

$$I_M = M_n(I + I_n) \quad (3.1)$$

$$I_p = (M_n - 1)(I + I_n) \quad (3.2)$$

$$I_n = (M_p - 1)I_p \quad (3.3)$$

From equations (3.1), (3.2), and (3.3), the current at the contacts I_M can be easily derived as

$$I_M = \frac{M_n}{1 - (M_n - 1)(M_p - 1)} I \quad (3.4)$$

The current gain is the ratio of the contact current with the initiating current:

$$MF_n = \frac{I_M}{I} = \frac{M_n}{1 - (M_n - 1)(M_p - 1)} \quad (3.5)$$

Equation (3.5) is derived for the electron-initiating II process, but it is equally valid for a hole-initiating II process as well. The current gain due to initiating holes is

$$MF_p = \frac{I_M}{I} = \frac{M_p}{1 - (M_n - 1)(M_p - 1)} \quad (3.5a)$$

For simple cases where $M_n=1$, equation (3.5) reduces to $MF_n=1$, regardless of the value of M_p . For cases where the feedback gain $M_p=1$, equation (3.5) becomes $MF_n= M_n$. It is obvious from equation (3.5) that $(M_n - 1)(M_p - 1)=1$ at avalanche breakdown. For cases that $M_n=M_p=2$,

avalanche breakdown occurs and equation (3.5) yields $MF_n \rightarrow \infty$. In silicon, Mn is usually higher than Mp [1-4]. Therefore, avalanche breakdown in Si devices occurs only when $Mn > 2$. That means every initiating electron, on average, needs to experience II at least once before avalanche breakdown occurs. In addition, equations (3.5) and (3.5a) show that at avalanche breakdown, the condition $(M_n - 1)(M_p - 1) = 1$ holds true for both the electron-initiating and hole-initiating II process. This suggests that the breakdown voltage is the same for both the electron-initiating and hole-initiating II processes in a high field region. This was experimentally proven to be true [5]. Equations (3.5) and (3.5a) show that the current gain from the electron-initiating II process is always higher than the current gain from the hole-initiating II process within the same high field region. This was also confirmed by the experimental results in Ref. [2].

So far, only qualitative reasoning of an II process is provided, and the model's validity needs to be verified. This can be accomplished by two steps: (1) introducing a well-established formulation of II current gain, and (2) comparing this formulation of II current gain to equation (3.5).

1. Derivation of a well-established formulation of II current gain This derivation can be found in different textbooks (for example, see Ref. [6]), and is shown here for the purpose of clear demonstration. Figure 3.4 shows a high field region $[0, W_D]$ in which electrons drift from left to right ($0 \rightarrow W_D$) and holes from right to left ($W_D \rightarrow 0$). Electrons are the initiating carriers of the II process in this derivation.

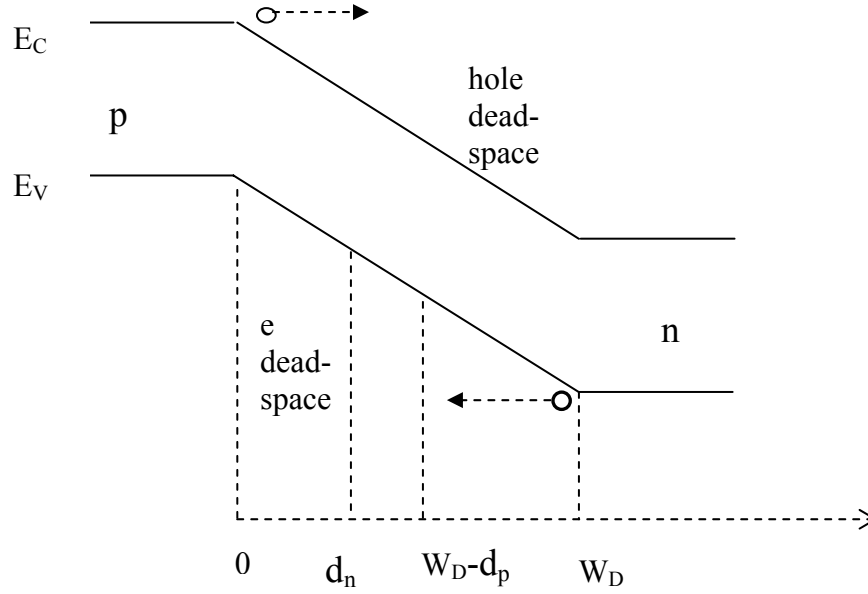


Figure 3.4: Band diagram of a p-n junction in reverse bias. Electrons cannot cause II from $[0, d_n]$, and holes cannot cause II from $[W_D - d_p, W_D]$; d_n and d_p are the “so-called” dead-space distances of electrons and holes, respectively.

The incremental electron current at position x is equal to the number of electron-hole pairs generated per second by both electrons and holes impact-ionizing in the distance dx ,

$$dI_n = I_n \alpha dx + I_p \beta dx \quad (3.6)$$

where α and β are II coefficients (cm^{-1}) of electrons and holes, respectively. An II coefficient is defined as the total number of electron-hole pairs generated per unit distance traveled.

In steady state, the total current $I (= I_n + I_p)$ is constant, and equation (3.6) becomes:

$$\frac{dI_n}{dx} - (\alpha - \beta)I_n = \beta I \quad (3.7)$$

The measured current is $I = I_n(W_D)$, and the total current gain is $MF_n = I/I_n(0)$. Equation (3.7) has the solution of the form

$$I_n(x) = I \left\{ \frac{1}{MF_n} + \int_0^x \alpha \exp\left(\int_0^x \beta - \alpha dx'\right) dx' \right\} / \exp\left(\int_0^x \beta - \alpha dx'\right) dx \quad (3.8)$$

From equation (3.8), it is easy to derive the total current gain from the initiating electron current:

$$1 - \frac{1}{MF_n} = \int_0^{W_D} \alpha \exp\left(\int_0^x \beta - \alpha dx'\right) dx \quad (3.9)$$

2. Comparison of equation (3.5) to equation (3.9) Expanding the right hand side of equation (3.9) gives

$$1 - \frac{1}{MF_n} = \int_0^{W_D} \beta \exp\left(\int_0^x \beta - \alpha dx'\right) dx - \int_0^{W_D} (\beta - \alpha) \exp\left(\int_0^x \beta - \alpha dx'\right) dx \quad (3.10)$$

Let

$$u = \exp\left(\int_0^x \beta - \alpha dx'\right) \quad (3.11)$$

and take the derivative of equation (3.11) with respect to position x gives

$$\frac{du}{dx} = (\beta - \alpha) \exp\left(\int_0^x \beta - \alpha dx'\right) \quad (3.11a)$$

So, the second term on the right hand side of equation (3.10) can be simplified as:

$$\int_0^{W_D} (\beta - \alpha) \exp\left(\int_0^x \beta - \alpha dx'\right) dx = \exp\left(\int_0^x \beta - \alpha dx'\right) \Big|_0^{W_D} = \exp\left(\int_0^{W_D} \beta - \alpha dx\right) - 1 \quad (3.12)$$

or

$$\int_0^{W_D} (\beta - \alpha) \exp\left(\int_0^x \beta - \alpha dx'\right) dx = \frac{\exp\left(\int_0^{W_D} \beta dx\right)}{\exp\left(\int_0^{W_D} \alpha dx\right)} - 1 \quad (3.12a)$$

Electrons with low initial energy need to travel a certain distance to gain enough energy from the electric field before they can cause impact ionization. This distance is normally referred to as

dead-space. Within this dead-space distance $[0, d_n]$, electron II coefficient is zero (figure 3.4).

Similarly, holes cannot cause II within $[W_D - d_p, W_D]$. Equation (3.12a) becomes

$$\int_0^{W_D} (\beta - \alpha) \exp\left(\int_0^x \beta - \alpha dx'\right) dx = \frac{\exp\left(\int_0^{W_D - d_p} \beta dx\right) - 1}{\exp\left(\int_{d_n}^{W_D} \alpha dx\right)} - 1 \quad (3.13)$$

and the first term on the right hand side of equation (3.10) can be rewritten as

$$\int_0^{W_D} \beta \exp\left(\int_0^x \beta - \alpha dx'\right) dx = \int_0^{W_D - d_p} \beta \exp\left(\int_0^x \beta - \alpha dx'\right) dx \quad (3.14)$$

For short devices, where $(d_n + d_p)$ is approximately equal to or larger than the depletion width

W_D , electrons cannot cause II within the region $[0, W_D - d_p]$, and equation (3.14) becomes

$$\int_0^{W_D} \beta \exp\left(\int_0^x \beta - \alpha dx'\right) dx = \int_0^{W_D - d_p} \beta \exp\left(\int_0^x \beta dx'\right) dx \quad (3.15)$$

Performing integration on the right hand side of the equation (3.15) gives

$$\int_0^{W_D} \beta \exp\left(\int_0^x \beta - \alpha dx'\right) dx = \exp\left(\int_0^{W_D - d_p} \beta dx\right) - 1 \quad (3.15a)$$

Substituting equations (3.13) and (3.15a) into equation (3.10) gives

$$1 - \frac{1}{MF_n} = \exp\left(\int_0^{W_D - d_p} \beta dx\right) - \frac{\exp\left(\int_0^{W_D - d_p} \beta dx\right) - 1}{\exp\left(\int_{d_n}^{W_D} \alpha dx\right)} \quad (3.16)$$

Simple algebra is performed on equation (3.16) to yield

$$MF_n = \frac{\exp\left(\int_{d_n}^{W_D} \alpha dx\right)}{1 - \left(\exp\left(\int_{d_n}^{W_D} \alpha dx\right) - 1\right) \left(\exp\left(\int_0^{W_D - d_p} \beta dx\right) - 1\right)} \quad (3.17)$$

Equation (3.17) resembles the form of the positive feedback model (equation 3.5). The following section will prove that equation (3.17) is, in fact, a long form of equation (3.5). In the positive feedback model formulated by equation (3.5), Mn is defined as the current gain due to electrons' II in the high field region (ignoring the holes' II) and is derived as follows. The incremental electron current at position x is equal to the number of electron-hole pairs generated per second by electrons:

$$dJ_n(x) = J_n(x)\alpha \quad (3.18)$$

The differential equation (3.18) yields the following equation:

$$M_n = \frac{J_n(W_D)}{J_n(0)} = \exp\left(\int_0^{W_D} \alpha dx\right) \quad (3.19)$$

Taking into account the dead-space distance, equation (3.19) becomes:

$$M_n = \frac{J_n(W_D)}{J_n(0)} = \exp\left(\int_{d_n}^{W_D} \alpha dx\right) \quad (3.20)$$

Similarly, Mp in equation (3.5) is defined as the current gain due to holes' II in the high field region (ignoring the electrons' II).

$$M_p = \frac{J_p(0)}{J_p(W_D)} = \exp\left(\int_0^{W_D - d_p} \beta dx\right) \quad (3.21)$$

Substituting equations (3.20) and (3.21) into equation (3.17), the total current gain due to electron-initiating II becomes:

$$MF_n = \frac{M_n}{1 - (M_n - 1)(M_p - 1)} \quad (3.22)$$

Equation (3.22) is identical to equation (3.5). This completes the proof that in short devices, the positive feedback model is a valid model for an II process involving both electrons' and holes' II in a high field region.

3.3 Implementation of the positive feedback model in Monte Carlo simulations

To simulate II in a p-n junction with electrons as the initiating carriers, first the MC program simulates II due to electrons alone to find Mn , and then the MC program simulates holes' II to find Mp . The results of electrons' and holes' II current gain will be substituted into equation (3.5) to find the total current gain due to electron-initiating II.

MC calculation of Mn A large number of initial electrons Ne (from 10^3 to 10^4) are released subsequently from the p+ side of the diode. The program follows the trajectory of each of the initial electrons, and the next electron is not simulated until the current one exits the depletion region. The initial electrons are in thermal equilibrium with the lattice and have random wave vector directions.

Mn is calculated by (1) summing up all the electrons (initial electrons and secondary electrons generated by II) that exit the depletion width and enter the n+-region, and (2) dividing the sum in (1) by the initial electrons Ne . It is trivial for the program to keep track of the total number of II events experienced by each initiating electron during the time it spends in the depletion width. Complications arise at high bias, where II due to secondary electrons (created by the initial electrons through II) becomes important. In addition, II due to the electrons generated by those secondary electrons may also be important, and so on. Therefore, the popular method, in which the secondary carriers are omitted [7-9], cannot be applied here. Another method is letting the MC program continue to simulate the secondary electrons after finishing simulating all the initial electrons [10]. Not only is this method confusing and complicated in coding the program, but it also causes a huge drag on the run-time.

In this work, a new and more efficient approach is developed. Each initial electron is assigned an initial weight $W=1$. This weight W doubles each time the electron encounters an II. The energy of the two resulting electrons is a function of the initial electron's energy and was calculated in Ref. [11]. The resulting two electrons after an II event are averaged by their energy and treated as one electron. The MC program continues simulating this new average electron whose weight is double of the initiating electron (figure 3.5). This method not only allows the program to simulate a fixed number of initial electrons but also correctly accounts for II from all subsequent secondary electrons. The program sums up the weights of all electrons exiting the high field region (Sum_W), and Mn is calculated as:

$$M_n = \frac{Sum_W}{Ne} \quad (3.23)$$

The resulting hole of an II event has the same weight as the initiating electron. The resulting holes' positions, weights, and energy are recorded and later fed into the MC program for holes.

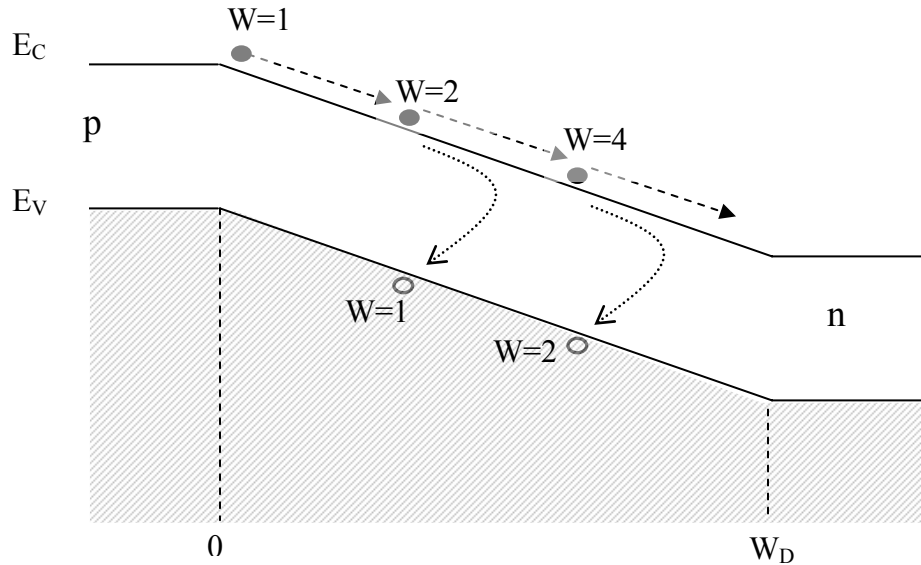


Figure 3.5: Band diagram explaining how the MC program counts the number of II events by initiating electrons in a p-n diode.

MC calculation of M_p

Unlike primary electrons whose original positions are at the edge of p+ depletion region, holes were generated at different positions along the depletion width. So, to calculate M_p , the MC program needs to simulate II current gain $M_p(x)$ due to holes at a location x for all x within the depletion width. At location x , the number of holes (created by the electrons' II in the previous simulations), represented by the weight $W(x)$, is retrieved from the record along with their energy. The number of holes at any position x is usually too small for the MC program to ensure the statistical accuracy. Therefore, those holes are represented by a large number of holes, $N_p(x)$, in the MC simulations. The initial energy of the representative holes is the hole average energy at x . The simulation method is the same as the method used for electrons. The current gain $M_p(x)$ from these $N_p(x)$ is calculated by the same formula used for electrons

$$M_p(x) = \frac{\text{Sum } W(\text{exit})}{N_p(x)} \quad (3.24)$$

After the program finishes simulating all holes from 0 to W_D , M_p is calculated by taking the average of $M_p(x)$ for all x within the depletion region as

$$M_p = \frac{\int_0^{W_D} W(x) M_p(x) dx}{\int_0^{W_D} W(x) dx} \quad (3.25)$$

Having calculated M_n and M_p , the current gain can now be easily calculated by equation (3.5).

Matlab pseudo-code to calculate M_n :

$Sum_W = Ne$

For $i = 1:Ne$

$Wi = 1$

While (electron not exit high field region)

Electron accelerated by the electric field

Electron experiencing a scattering event

If (impact ionization occurs)

$Sum_W = Sum_W + Wi$

$Wi = 2 * Wi$

end If

Electron's state after scattering

end While

end For

$$M_n = \frac{Sum_W}{Ne}$$

Matlab pseudo-code to calculate M_p :

For $x=0:W_D$

$W(x)$ =weight input from electron MC simulation's result

$Sum_W=Np$

For $i=1:Np$

$Wi=1$

While (hole not exit high field region)

Hole accelerated by the electric field

Hole experiencing a scattering event

If (impact ionization occurs)

$Sum_W=Sum_W + Wi$

$Wi=2*Wi$

end If

Hole's state after scattering

end While

end For

$M_p(x)=Sum_W/Np$

end For

$$M_p = \frac{\int_0^{W_D} W(x)M_p(x)dx}{\int_0^{W_D} W(x)dx}$$

3.4 An efficient Monte Carlo algorithm to simulate spatial transient transport in semiconductor devices

In semiconductor devices, electric field is usually not constant but rather varying across devices' structures. To include the spatially-varying electric field in MC simulations, Jacoboni and Lugli [12] divide a device into meshes of equal sizes. The electric field is constant within each mesh. During free flights, the positions of the charged carriers (let say electrons) need to be checked against mesh boundaries so that the correct electric field is updated. It is simple for the case that a stochastically-generated free-flight time (fFT) is small enough for the electrons to remain in the same mesh. However, if the stochastically-generated fFT is longer than the time it takes to transverse a mesh, the simulations are more involving. First, the MC program needs to calculate the required time for the electrons to reach the next mesh boundary. Then the program determines the states of the electrons at the mesh boundary. The electrons will then be simulated in the next mesh (with the new electric field and the starting position is the mesh boundary) for the remaining fFT. If the remaining fFT is still larger than the required time to reach the next mesh boundary, this process will continue for the rest of the fFT. Figure 3.6a shows the flow chart describing the steps necessary for a MC program to simulate charge transport in spatially-varying electric field.

- First, the spatial directions of the electrons are identified in order to determine the indexes of the next mesh boundary.
- Second, the position of the next mesh boundary is retrieved from the lookup table that has been set up before starting the MC program.

- Third, the MC program calculates the time (t_{mesh}) it takes the electrons to reach this mesh boundary.
- Fourth, a comparison step is done to extract the smaller value among the remaining fFT and t_{mesh} to be the current fFT.
- Fifth, the electrons' new states and new electric field (newEField) are updated.

The whole five-step cycle will continue while the remaining fFT is larger than 0. This five-step cycle is particularly more time-consuming and calculation-burdening in 2-D and 3-D meshes.

In this work, a different and more efficient algorithm is developed. Instead of dividing the devices into spatial meshes, the stochastically-generated fFT is divided into equal temporal meshes t_{step} . The electric field is constant within each temporal mesh. Figure 3.6b shows the implementation of the temporal mesh method in MC simulations. Obviously, for the temporal mesh method, a MC program can eliminate extra steps required to determine fFT, during which the electric field is treated as constant. This saves a huge amount of simulation time considering many charge carriers to be simulated and many fFT steps to be stochastically generated.

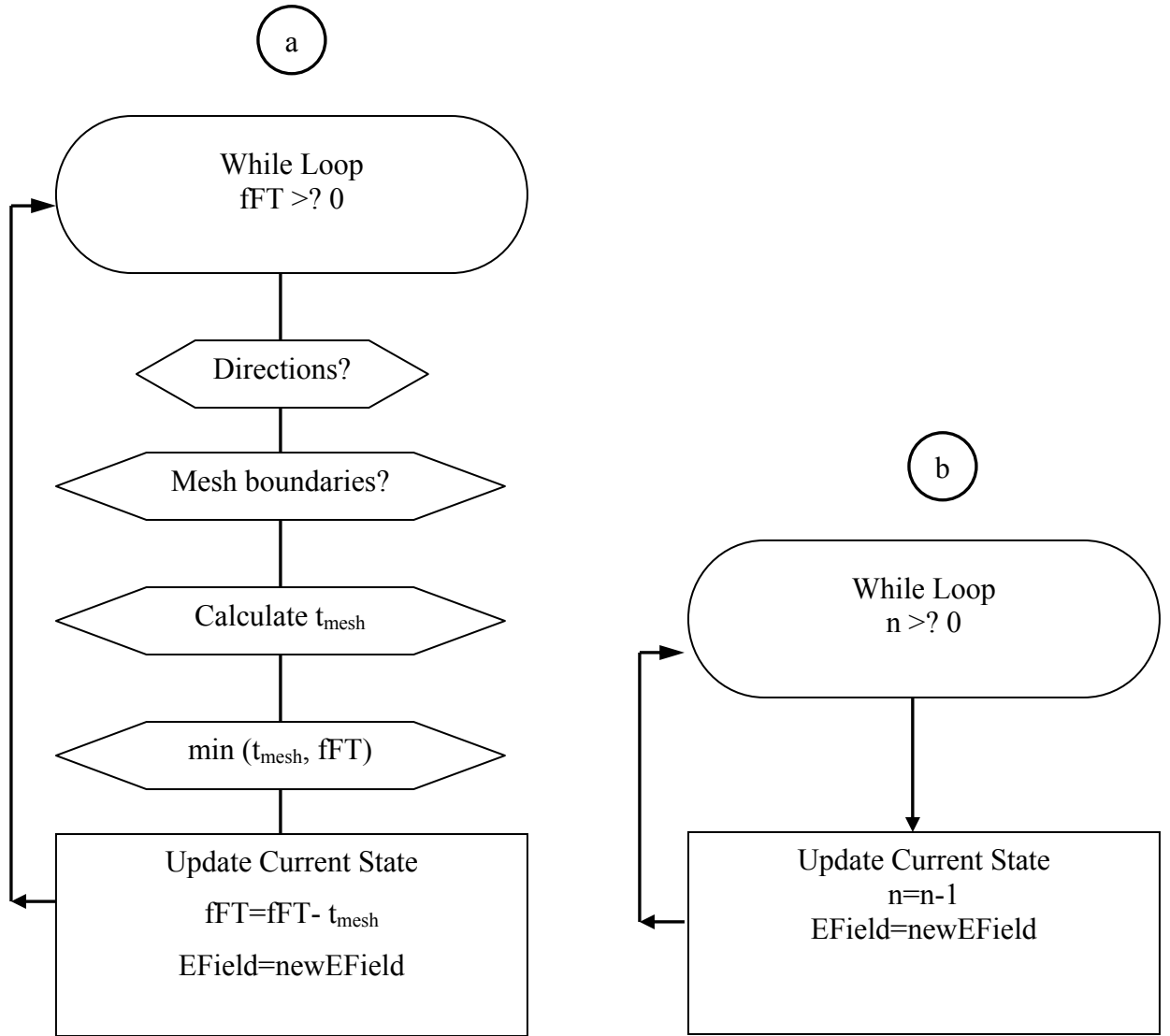


Figure 3.6: Flowcharts of spatial mesh (a) and temporal mesh (b) methods in spatial transient MC simulations. For temporal mesh method, n is calculated before the program executes the internal loop and $n = fFT / t_{step}$.

3.5 Summary

An II process, for the first time, is successfully modeled and formulated as a positive feedback model and applied to MC simulations. The main reason for such simple modeling of the complex electron-hole II loop is the existence of electrons' and holes' dead-space distances. In short devices, electrons' and holes' dead-space distances overlap. Consequently, holes cannot have II in the regions where electrons can, and vice versa. Therefore, II due to electrons and II due to holes can be decoupled from one another in short devices' high field regions. The validity of the model is confirmed by proving that the II current gain equation in the positive feedback model is, in fact, identical to the traditionally-used and well-known II formula [6].

A new algorithm is developed, enabling MC programs to simulate II from secondary carriers while maintaining the fixed number of simulated carriers. This is accomplished by assigning a variable W to each primary carrier. The variable W is initially equal to 1, and it doubles each time the primary carrier experiences an II event. The II current gain is determined by simply calculating the average W of all the primary carriers at the end of the high field region.

An efficient MC algorithm is developed to study spatial-transient transport in semiconductor devices. Rather than using the traditional spatial-mesh method, this work introduces a temporal-mesh method that eliminates the extra time-consuming steps but still preserves the accuracy. It is tremendously useful that accurate results can be obtained from the temporal-mesh algorithm in combination with analytical band MC simulations. A huge amount of time and computing resources can be saved, enabling adequate MC programs to be run on PCs and potentially offering easy access to MC programs among university students.

Bibliography

- [1] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Holes and Electrons in Silicon," *Physical Review*, Vol. 134, No. 3A, 1964.
- [2] R. Van Overstraeten and H. De Man, "Measurement of the ionization rates in diffused silicon p-n junctions," *Solid-State Electronics*, Vol. 13, pp. 583-608, 1970.
- [3] W. N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields," *Solid-State Electronics*, Vol. 16, pp. 1189, 1973.
- [4] M. H. Woods, W. C. Johnson and M. A. Lampert, "Use of a Schottky Barrier to Measure Impact Ionization Coefficients in Semiconductors," *Solid-State Electronics*, Vol. 16, pp. 381, 1973.
- [5] S.L. Miller, "Ionization rates for holes and electrons in silicon," *Physical Review B*, Vol. 105, pp. 1246, 1957.
- [6] S.M. Sze, "Physics of Semiconductor Devices," John Wiley & Sons, 1981.
- [7] P.A. Lebowitz and P.J. Price, "Hybrid method for hot electron calculations," *Solid State Communications*, Vol. 9, pp. 1221, 1971.
- [8] R.C. Curby and D.K. Ferry, "Impact ionization in narrow gap semiconductors," *Physica Status Solidi A*, Vol. 15, pp. 319, 1973.
- [9] H. Shichijo, K. Hess, and G.E. Stillman, "Simulation of high-field transport in GaAs using a Monte Carlo method and pseudopotential band structures," *Applied Physics Letters*, Vol. 38, pp. 89, 1981.
- [10] D.S. Ong, K.F. Li, S.A. Plimmer, G.J. Rees, J.P.R. David, and P.N. Robson, "Full band Monte Carlo modeling of impact ionization, avalanche multiplication, and noise in submicron GaAs p⁺-i-n⁺ diodes," *Journal of Applied Physics*, Vol. 87, pp. 7885, 2000.
- [11] Y. Kamakura, H. Mizuno, M. Yamaji, M. Morifuji, K. Taniguchi, C. Hamaguchi, T. Kunikiyo, and M. Takenaka, "Impact Ionization Model for Full Band Monte Carlo Simulation," *Journal of Applied Physics*, Vol. 75, pp. 3500, 1994.
- [12] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation (Computational Microelectronics)*, Springer-Verlag Wien New York, 1989.

Chapter 4:

New perspectives on physics and modeling of impact ionization

4.1 Introduction

Modeling II in devices has been of great interests. Among the most widely-used models are the local electric field model [1-2], the delay model [3-6], and the energy model [7-9]. All of these three models are only applicable to limited devices' structures and limited operating regimes. The main causes for the failures of the three models are investigated in details in this chapter. Microscopic properties of II need to be established before a more complete model can be developed. This can be accomplished with the help of MC simulations.

Short p-i-n diodes are chosen for the II study in this chapter. Understanding deeper physical insights of II in p-i-n structures is extremely useful because of their well-recognized applications in avalanche photodiodes. Besides, p-i-n diodes resemble the base-collector junctions, and, consequently, investigating II in p-i-n diodes enables better understanding and modeling of breakdown conditions in BJTs and HBTs. In addition, the simple p-i-n structures and their trivial electric field profiles make p-i-n diodes an ideal candidate to investigate new physical insights without the confusion caused by parasitic effects. Figure 4.1 shows the cross-section of a p-i-n diode and its electric field profile.

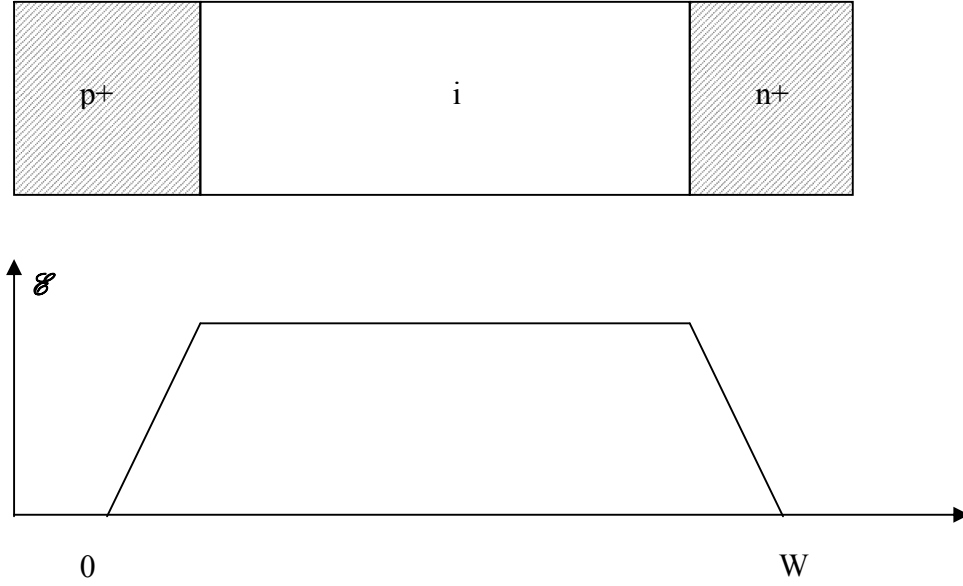


Figure 4.1: A simplified p-i-n diode's cross-section and its electric field profile. The electric field is constant in the i-region and varies according to Poisson's equation in the p^+ and n^+ regions.

The experimental current gains of p-i-n diodes reported in Ref. [9] are used to analyze different II models. These experimental results are also studied using MC simulations in this work. In the experiment [9], the photo-generation method was used to generate electrons in the p^+ region. The electrons diffused to the p^+ -i junction, accelerated in the high field region, and caused II. Agarwal *et al.* [9] measured the current gain MF_n (due to electron-initiating II) by taking the difference of the measured current with light-on and light-off conditions

$$MF_n = \frac{I_{light}(V) - I_{dark}(V)}{I_{light}(0)} \quad (4.1)$$

The effects of tunneling current were inherently eliminated since tunneling current did not change with the light's on-off conditions. Agarwal *et al.* [9] also oscillated the light source at 10 Hz and measured the current with a lock-in amplifier. The gain MF_n was calculated by dividing the measured ac-current at bias (V) by the ac-current at very low bias ($V \sim 0$):

$$MF_n = \frac{I_{ac}(V)}{I_{ac}(0)} \quad (4.2)$$

The tunneling-current component was not modulated with the frequency of the light and rejected by the lock-in amplifier. The results from the two methods were close [9]. The careful measurement setup guaranteed that the current gain was purely due to II mechanism.

In this study, n^+ and p^+ dopant profiles are treated as uniform and abrupt with $1 \times 10^{19} \text{cm}^{-3}$ doping concentration for 124nm, 210nm, and 290nm diodes. For shorter diodes whose i -widths are 77nm, 31nm, exact n^+ and p^+ dopant profiles [9] are used. Throughout this chapter, electrons are treated as traveling from left $x=0$ to right $x=W$ and holes traveling from right $x=W$ to left $x=0$, unless specified otherwise. The electric field profile is assumed unchanged in the presence of the electron-hole pairs generated by II. Current gain due to the initiating electrons MF_n is defined as:

$$MF_n = \frac{J_n(W)}{J_n(0)} \quad (4.3)$$

4.2 Analysis of models for II in semiconductor devices

4.2.1 Local electric field model

The II current gain due to primary electrons moving from $x=0$ to $x=W$ is (equation (3.9) derived in Chapter 3):

$$1 - \frac{1}{MF_n} = \int_0^W \alpha \exp\left(\int_0^x \beta - \alpha dx'\right) dx \quad (4.4)$$

where $\alpha(\mathcal{E}) = A_n \exp(-B_n/\mathcal{E})$, and $\beta(\mathcal{E}) = A_p \exp(-B_p/\mathcal{E})$.

The local electric field model treats the II coefficients (α , β) in equation (4.4) as pure functions of the local electric field. The validity of this model is demonstrated by calculating the current gain of p-i-n diodes using the experimental values of the electron and hole II coefficients reported in Ref. [10]. Figure 4.2 shows the calculated results of the II current gain in comparison with the experimentally measured values [9]. Clearly, this model overestimates the current gain, especially for shorter diodes. The reason for the overestimation is explained as follows. As shown in chapter 3, equation (4.4) is derived from the current continuity equation:

$$dI_n(x) = I_n(x)\alpha dx + I_p(x)\beta dx \quad (4.5)$$

Equation (4.5) assumes that all electrons and holes at position x have the same II probability $\alpha(x, \mathcal{E})$ and $\beta(x, \mathcal{E})$, respectively. The local electric field model makes a further assumption that α and β only depend on the local electric field, implying that mobile carriers can experience II at any position in the high field region. This assumption is the cause of the overestimation. Only carriers that have energy higher than the threshold energy E_T can cause II, and carriers with very low initial energy need to travel at least a distance of $E_T/q\mathcal{E}$ to gain the threshold energy from the electric field. Clearly, within the distance $(0, E_T/q\mathcal{E})$, mobile carriers have zero II probability.

Consequently, the local electric field model is not valid for short devices where the distance $E_T/q\mathcal{E}$ consumes a large portion of the devices' lengths.

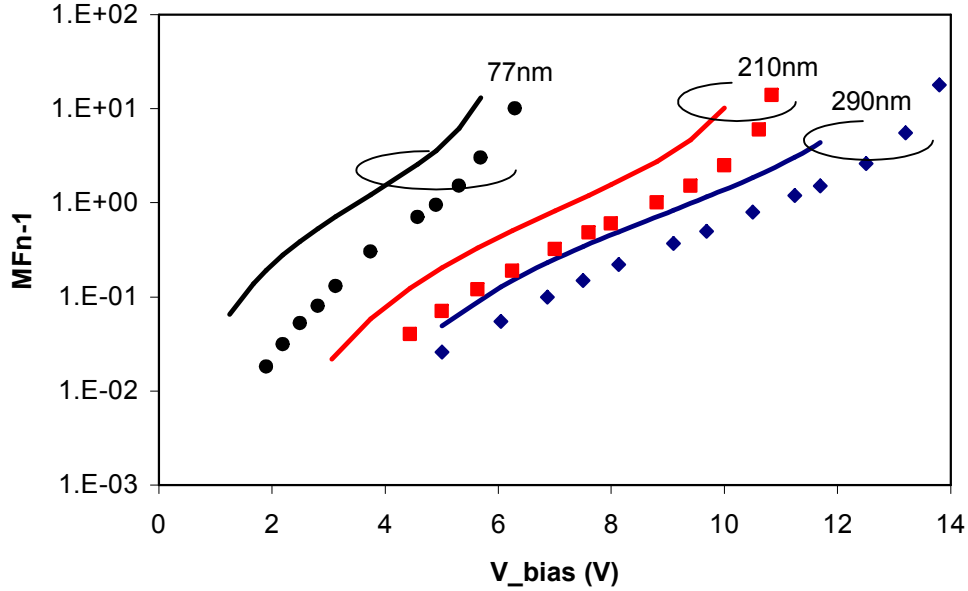


Figure 4.2: II current gain of p-i-n diodes versus bias. Lines are theoretical results using the local electric field model, and dots are experimental data [9].

4.2.2 Delay model

To account for the non-local effect of II, modifications to the current continuity equation (equation 4.5) were implemented [3-6]:

$$dI_n(x) = I_n(x - d_n)\alpha(x)dx + I_p(x + d_p)\beta(x)dx \quad (4.6)$$

where d_n and d_p are the electron and hole dead-space distances, respectively. Dead-space is defined as the distance that electrons or holes, starting out at very low energy, need to travel to gain enough energy before they can cause II. Equation (4.6) basically makes certain that the carriers, generated by II within a dead-space distance of position x , cannot cause II at x . That is because those carriers have not accumulated enough energy higher than the threshold energy

when they reach position x . The biggest issue of the delay model is the strict requirement that only $n(x-dn)$ and $p(x+dp)$ can cause II at x . First, this model cannot be solved analytically. Computer simulations are needed to solve equation (4.6) by following each carrier's trajectory to keep track of the II events. Second, this model does not rigorously account for the fact that dn and dp are variables because newly generated carriers have non-zero initial energy. In fact, newly generated carriers' initial energy depends on the initiating carriers' energy, according to the momentum and energy conservation law [11-12]. Third and most importantly, equation (4.6) only ensures that carriers, generated by II within a dead-space distance of position x , cannot cause II at x . However, equation (4.6) does not take into the account of primary carriers that cause II within a dead-space distance of position x . For example, when some of the electrons $n(x-dn)$ at position $(x-dn)$ experience II before reaching x , the number of electrons that can cause II at x is smaller than $n(x-dn)$. In conclusion, this model cannot be justified in predicting the II current gain in semiconductor devices.

4.2.3 Energy model

Another approach of dealing with the dead-space effect is introducing the effective electric field (ϵ_{eff}) [7-9]. The electron and hole II coefficients in the current gain equation (4.4) are treated as pure functions of the **effective** electric field (not the real, local electric field). The effective electric field can be derived as follows. Starting from the following equations of the steady-state hydrodynamic model for electrons

$$div\vec{S} = \vec{E} \cdot \vec{J} - \frac{w - w_0}{\tau_w} n - w_{th} (G(n, T) - R(n, T)) \quad (4.7)$$

$$\vec{S} = -\kappa(T) \overrightarrow{grad} T - \frac{\vec{J}}{q} (k_B T + w) \quad (4.8)$$

where $\mathbf{S}$ is energy flux, $\mathbf{E}$ is the electric field, $\mathbf{J}$ is electron current density, n is electron density, τ_w is energy relaxation time, κ is thermal conductivity, G and R are generation and recombination rates, respectively. Electron average energy is $w = \frac{3}{2}k_B T + \frac{1}{2}m^* v^2$, and $w_{th} = \frac{3}{2}k_B T$ is average thermal energy, and $w_0 = \frac{3}{2}k_B T_L$ where T_L is the lattice temperature.

In the energy model, several assumptions are made [13-14]:

- Heat flux, generation and recombination are negligible.
- Drift energy is a negligibly small fraction of the kinetic energy.
- Energy relaxation time is constant.

Taking the divergence of equation (4.8) and substituting $\text{div}\vec{S}$ into equation (4.7) give

$$nv \cdot \text{div}(w) + nvk_B \cdot \text{div}(T) = qnvE - \frac{w - w_0}{\tau_w} n \quad (4.9)$$

Since the drift energy is neglected, $w = \frac{3}{2}k_B T$; consequently, $vk_B \cdot \text{div}(T) = \frac{2}{3}v \cdot \text{div}(w)$. Equation

(4.9) becomes

$$\frac{5}{3}v \cdot d(w - w_0) + \frac{w - w_0}{\tau_w} = qvE \quad (4.10)$$

Substituting the energy relaxation length [13] $\lambda_w = \frac{5}{3}v\tau_w$ into equation (4.10) yields

$$d(w - w_0) + \frac{w - w_0}{\lambda_w} = \frac{3}{5}qE \quad (4.11)$$

The solution of the first order differential equation (4.11) is

$$w(x) - w(0) = \frac{\int_0^x \frac{3}{5}qE \exp\left(\int_0^u \frac{1}{\lambda_w} d\xi\right) du}{\exp\left(\int_0^x \frac{1}{\lambda_w} d\xi\right)} \quad (4.12)$$

or

$$w(x) - w(0) = \frac{\int_0^x \frac{3}{5} q E \exp\left(\frac{u}{\lambda_w}\right) du}{\exp\left(\frac{x}{\lambda_w}\right)} \quad (4.13)$$

Bringing the denominator into the integral yields

$$w(x) - w(0) = \frac{3}{5} q \int_0^x E \exp\left(\frac{u-x}{\lambda_w}\right) du \quad (4.14)$$

Equation (4.14) describes the average energy of electrons and holes as a function of the electric field. The energy model introduces the effective electric field ε_{eff} such that the average energy can be written as

$$w(x) - w(0) = \frac{3}{5} q \varepsilon_{eff}(x) \lambda_w \quad (4.15)$$

From equations (4.14) and (4.15), the effective electric field is

$$\varepsilon_{eff}(x) = \frac{1}{\lambda_w} \int_0^x E \exp\left(\frac{u-x}{\lambda_w}\right) du \quad (4.16)$$

Dead-space effect is inherently included in equation (4.16). As electrons start traveling downfield, their average energy increases from the initial thermal energy, and ε_{eff} increases from the low initial value. Consequently, at the beginning of their drift in the high field region, electrons have zero II probability since the electron II coefficient is an exponential function of the effective electric field. The higher the real electric field is, the faster ε_{eff} rises, and the shorter the dead-space distance becomes. As the electrons travel farther, ε_{eff} approaches the real electric field, and the electron II coefficient is in equilibrium with the real electric field.

It is worth stressing that the energy model is derived assuming negligible generation of electron-hole pairs in the high field region. Therefore, the energy model is not applicable at high

biases where generation of electron-hole pairs due to II there is sufficiently high. This explains the difficulty Agarwal *et al.* [9] had when fitting the experimental data at high bias using this model.

4.3 MC investigations of spatial-transient II

4.3.1 MC calculations of current gain in short p-i-n diodes

The II current gain of p-i-n diodes is calculated using MC simulations and the positive feedback model developed in chapter 3. The time step in the temporal-mesh method is carefully chosen such that the distance over which the electric field remains constant is less than 1% of the devices' widths. This ensures the correct simulation of the spatially-varying electric field in the n^+ and p^+ depletion regions. The initial electrons are released at the edge of p^+ -i region and have the distribution function in equilibrium with the lattice temperature. The MC programs calculate the II current gain due to electrons M_n (ignoring holes' II) and due to holes M_p (ignoring electrons' II), and the total current gain is calculated by equation (3.5) derived in chapter 3:

$$MF_n = \frac{M_n}{1 - (M_n - 1)(M_p - 1)}$$

Figure 4.3 shows the simulated results in comparison with the measured data [9]. With the exception of the 31nm p-i-n diode, the calculated results agree well with the experimental data. For the 31nm p-i-n diode, the simulated result is much smaller than the measured values at all biases.

At low biases, where the total current gain is less than 1.5, the contribution from holes' II is negligible. However, at higher bias, holes' II becomes significant, and the positive feedback mechanism sharply increases the current gain (figure 4.3). Figure 4.4 shows that, for shorter diodes, avalanche breakdown occurs at lower values of M_n and higher values of M_p . Therefore, holes' II is more important in shorter diodes. At avalanche breakdown, both M_n and M_p approach the value of 2 with the scaling of p-i-n diodes (figure 4.4).

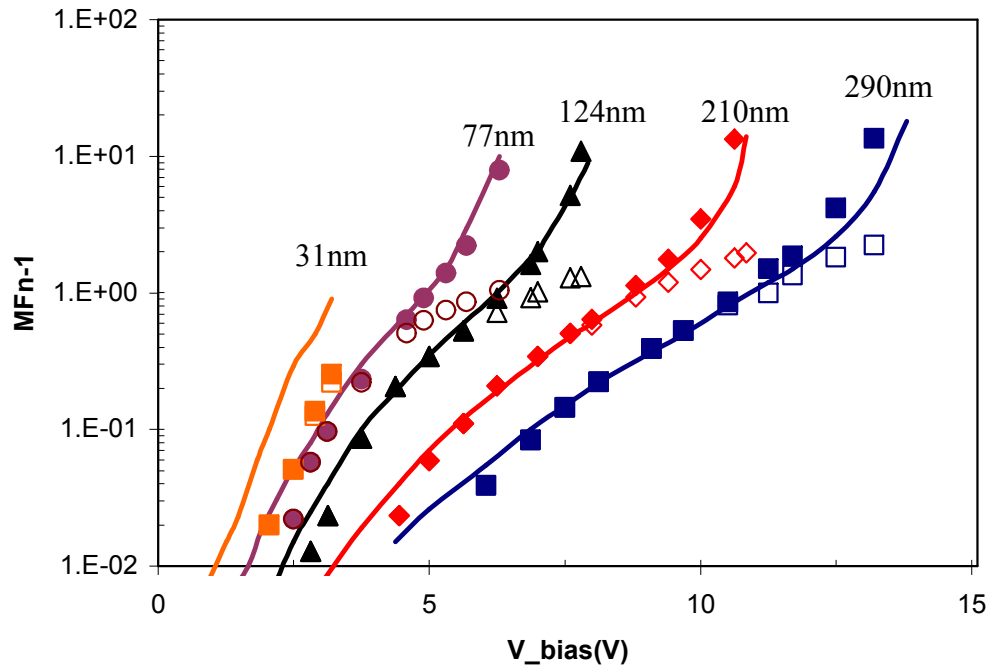


Figure 4.3: II current gain of p-i-n diodes versus reverse bias. Lines are the measured data [9]. Filled and empty dots are the simulated results with and without hole feed back, respectively.

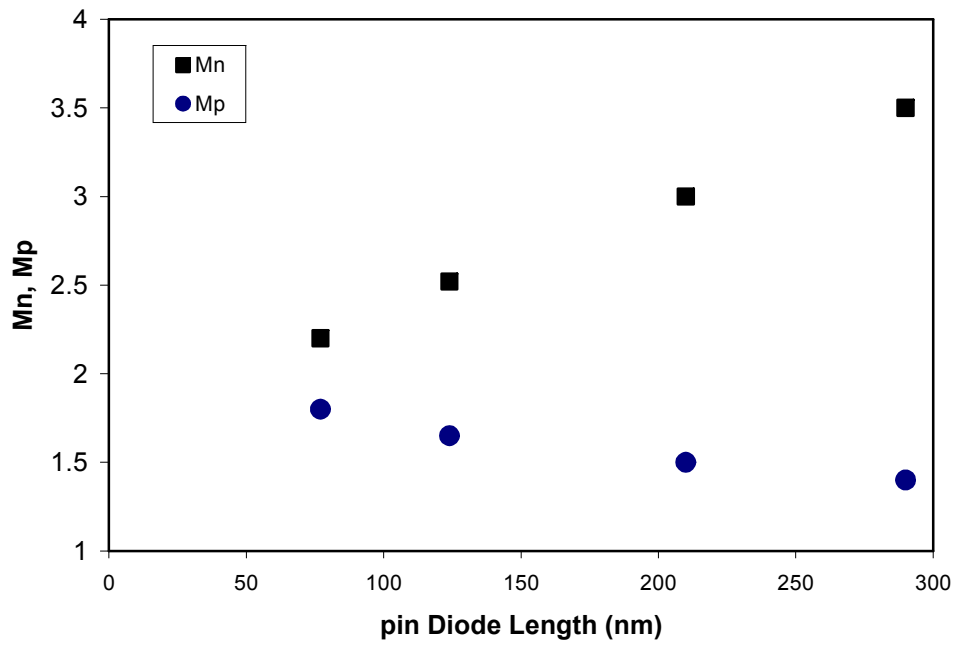


Figure 4.4: II current gain of electron and hole versus the intrinsic length of p-i-n diodes at avalanche breakdown condition.

The avalanche breakdown voltage, as expected, is smaller for shorter diodes (figure 4.5). On the contrary, at breakdown, the electric field in the i-region increases for shorter diodes. This is due to the fact that MFn depends on both the electric field and the diode's length. After experiencing an II, electrons lose energy. These electrons need to travel a certain distance to gain enough energy in order to cause other II events. The distance available for the electrons to travel after the first II event is shorter for shorter diodes. Therefore, to gain the same energy at shorter distance in shorter diodes, the electric field must be higher. It is interesting that the breakdown voltage (V_{BR}) can be best fit as a linear function of diodes' lengths (L) as shown in figure 4.5. The best-fit line yields the following equation

$$V_{BR}(V) = 0.032L(nm) + 4.41 \quad (4.17)$$

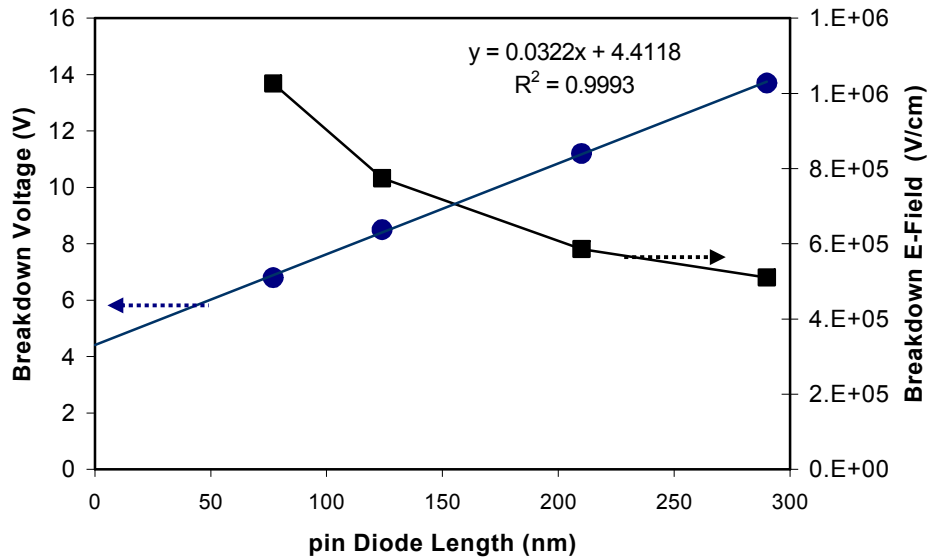


Figure 4.5: Bias voltages and the i-region's electric field at avalanche breakdown conditions for diodes of different lengths.

Equation (4.17) is used to predict the breakdown voltage of p-i-n diodes fabricated by a different group [15]. Table 4.1 shows that the theoretical breakdown voltages are in a good agreement

with the measured results reported in Ref. [15]. Equation (4.17) also suggests that as p-i-n diodes' lengths are scaled to 0nm (becoming heavily doped p-n diodes), avalanche breakdown voltage cannot be lower than 4.41 V. This number is more justified by the fact that II coefficients tend to saturate at high electric field [10]. This minimum 4.41V of the breakdown voltage is close to the 4.5V value observed for I-MOS shorter than 50nm [16] and close to the breakdown voltage of 4.0V measured experimentally in heavily doped, abrupt p-n junctions [17]. In fact, no experimental reports of II avalanche breakdown voltage smaller than 4.0V have come to my knowledge.

i- length (nm)	Theoretical V_{BR} (V)	Measured V_{BR} (V)
800	30.2	30
290	13.7	15
140	8.9	9
100	7.6	7.5

Table 4.1: Breakdown voltage of p-i-n diodes obtained from the theoretical calculation (equation 4.17) and from the experimental data [15].

4.3.2 Franz-Keldysh effect on impact ionization

At low biases, simulated II current gain is smaller than the experimental data for all p-i-n diodes of different lengths (figure 4.3). For the 31nm p-i-n diode, the calculated current gain is smaller than the experimental data at all biases. Agarwal *et al.* [9] suggested that unwanted hole generation during the measurements caused the abnormally-high current gain at low bias. This is not a satisfactory explanation due to the large difference between the calculated and measured II current gain and due to the fact that current gain is higher for an electron-initiating II process

than hole. The main cause for such discrepancy at low bias and in short diodes can be explained as follows.

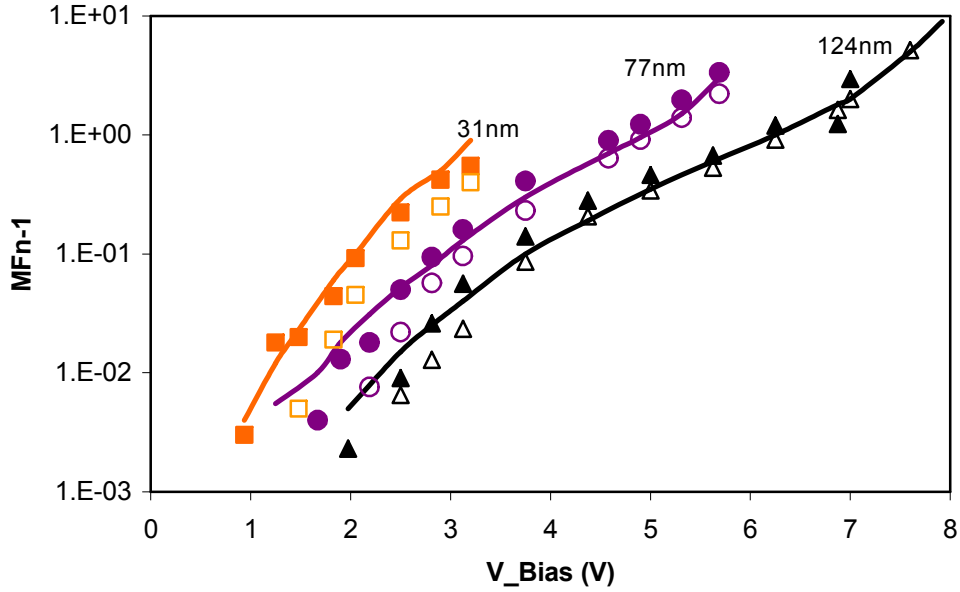


Figure 4.6: II current gain in p-i-n diodes versus reverse bias. Lines are the measured data [9]. Filled and empty dots are the results of the MC simulations with and without Franz-Keldysh effect on the electron II rate, respectively.

II is not an instantaneous process. During an II event, the primary carrier still gains extra energy from the electric field. Consequently, carriers whose initial energy is smaller than the threshold energy E_T can also cause II. In other words, effective E_T becomes smaller at higher electric field. This is often referred to as the intracollisional field effect or the effective two-particle Franz-Keldysh effect. Franz-Keldysh effect strongly influences II rate at energy near E_T but becomes negligible at higher energy [18-19]. The reduction of effective E_T and the increase of II rate at low energy have been numerically calculated in Ref. [18-19]. At energy smaller than $E_T(\mathcal{E}=0)$, the electron II rate calculated in Ref. [19] can be approximated as

$$P_{II} = 2.2 \times 10^{13} (E - E_T(\mathcal{E}))^7 \quad (4.18)$$

where $E_T(\mathcal{E})$ is expressed as

$$E_T(0)eV - E_T(\mathcal{E})eV = 0.19\mathcal{E}(MV/cm) + 0.12 \text{ for } 0.1MV/cm \leq \mathcal{E} \leq 1MV/cm \quad (4.18a)$$

Equations (4.18) and (4.18a) are used for the electron II rate at low energy ($E < E_G = 1.12\text{eV}$) in the MC program. Figure 4.6 shows that the MC results with Franz-Keldysh effect are in good agreement with the experimental data at low biases. Franz-Keldysh effect clearly plays a significant role in II current gain, especially at low gain level and in short diodes. Therefore, this effect needs to be included in designing devices where small gain value can influence the devices' operations. For example, in BJTs with gain $\beta=100$, the open-base breakdown voltage corresponds to $MFn-1=0.01$ [9].

The improvement of the theoretical calculations with Franz-Keldysh effect for the 31nm p-i-n diode is the clearest. In fact, due to the II dead-space effect, Franz-Keldysh effect is more important for shorter diodes. Charge carriers need to travel a certain distance to gain energy larger than E_T before being able to cause II. Franz-Keldysh effect reduces the effective E_T , and, as a result, the effective dead-space distance becomes smaller. Therefore, in a shorter diode, where the dead-space consumes a significant portion of the diode's length, Franz-Keldysh effect increases the II current gain more significantly.

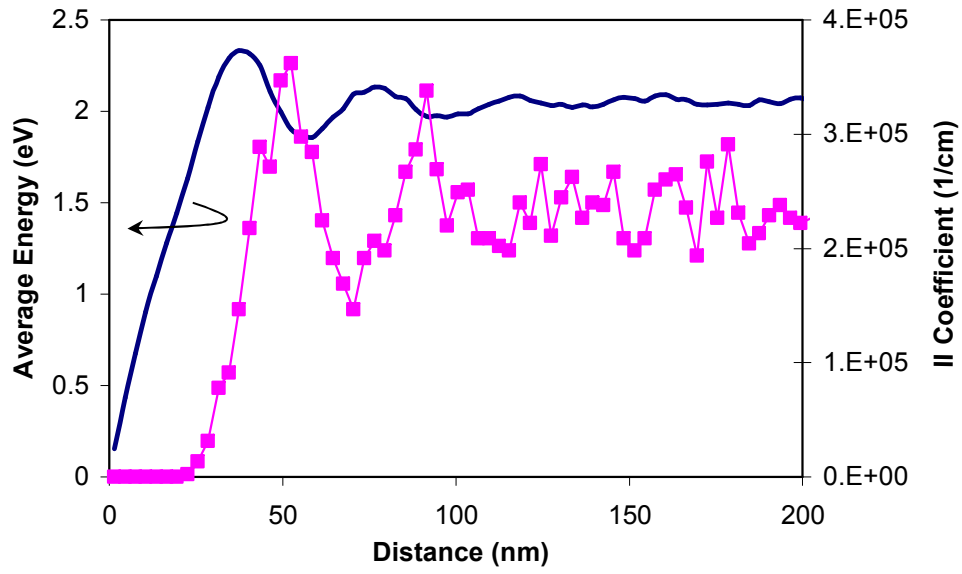
To the best of my knowledge, this is the first time Franz-Keldysh effect is successfully shown to have a strong influence on the experimental II current gain.

4.3.3 Spatial-transient values of II coefficients

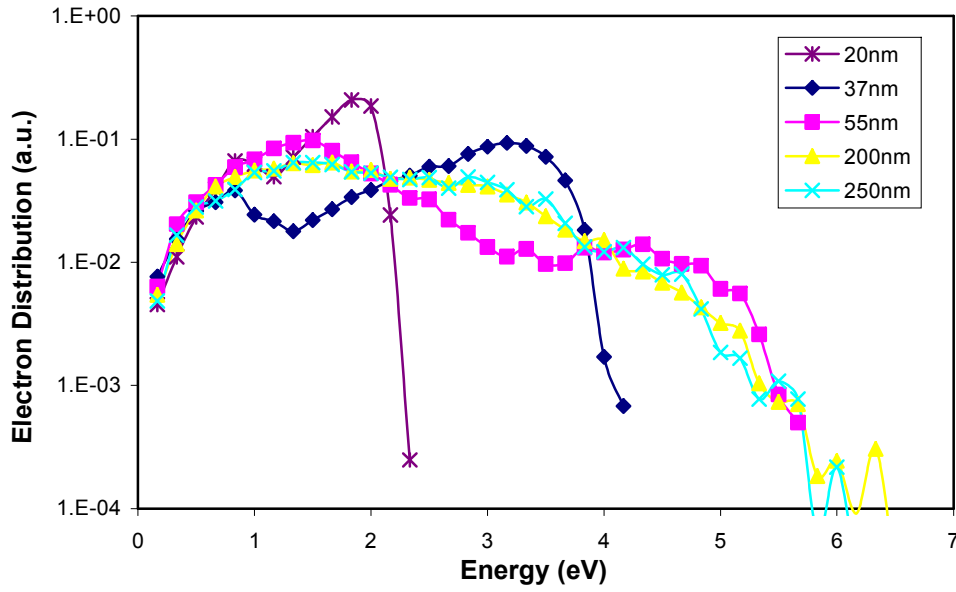
It was previously shown that the non-local effect of the II coefficients causes the limited applicability of different II models. Investigations of electron and hole dead-space are conducted

in this section using MC simulations. A large number of electrons are released at $x=0$ with the initial average energy of $3/2k_B T_L$. As the electrons travel in a high field region, they pick up energy from the electric field, and, at the same time, they lose a small fraction of energy mainly due to intervalley phonon scatterings. At high enough energy, the electrons lose a significant amount of energy through II events. The electron II coefficient is calculated by counting the average number of electron-hole pairs generated by an electron per unit distance. Figure 4.7a shows the spatial variation of the electron average energy and the II coefficient in homogeneous electric field of 10^6 V/cm. Clearly, there exists a dead-space distance where electrons have not accumulated enough energy to cause II. At high enough energy, II occurs, causing the average energy to decrease. Then, the electrons continue gaining energy from the electric field and losing energy to subsequent II events. The electron average energy and II coefficient are eventually in equilibrium with the electric field at the distance of about 100nm. The electron energy distribution function is calculated by the MC program to explain the spatial-transient behaviors (figure 4.7b). Electrons initially start out with low average energy at $x=0$ nm and maintain the uniformity in energy at $x=20$ nm. Some of the electrons gain enough energy and experience their first II events, spreading the distribution function (37nm curve). As the electrons continue drifting downfield, more II events occur, and the distribution function spreads further (55nm curve). At farther distance, the distribution function is in equilibrium with the electric field (200nm and 250nm curves). This explains the observations that (1) the fluctuation of the electron average energy and II coefficient decrease with distance, and (2) the average energy and the II coefficient become steady at about 100nm.

The electron spatial-transient II coefficient in different homogeneous electric field strengths is plotted in figure 4.8a.

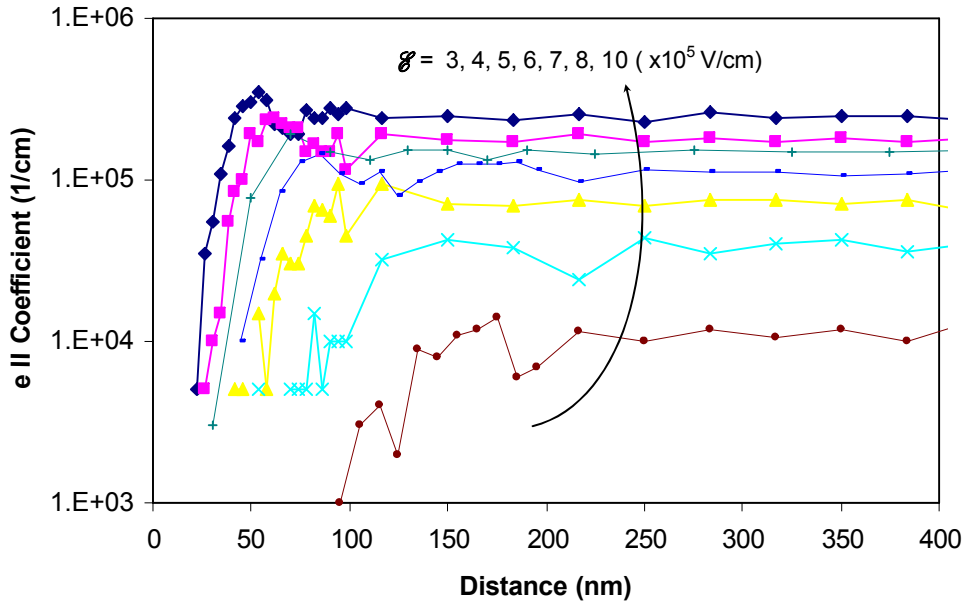


(a)

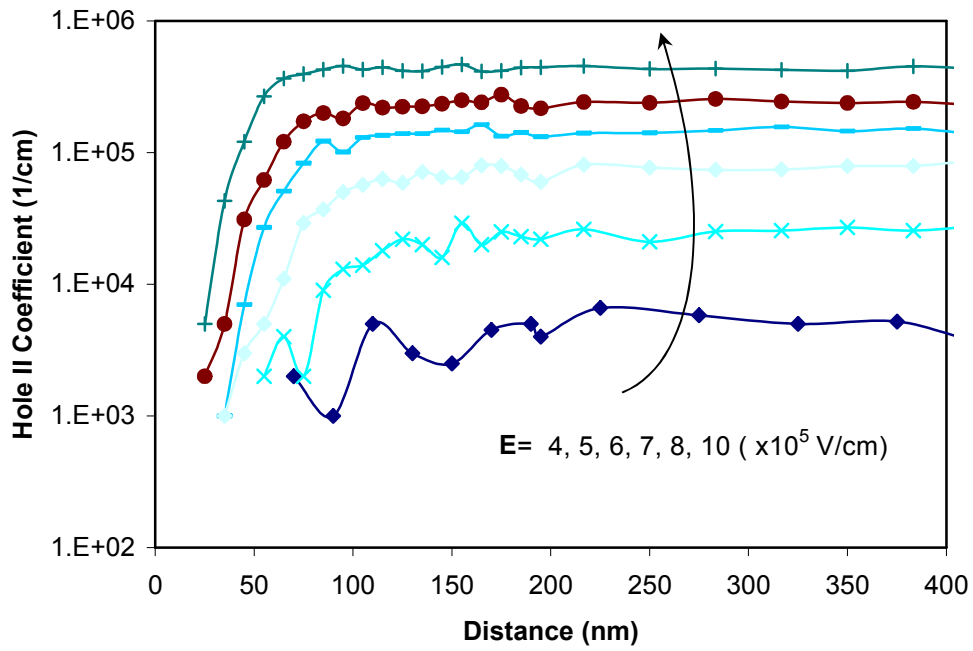


(b)

Figure 4.7: (a) Electron average energy and II coefficient versus energy at $\mathcal{E}=10^6\text{V/cm}$. (b) Distribution functions at different positions versus energy in a homogeneous field $\mathcal{E}=10^6\text{V/cm}$. 20nm: before the first peak in average energy is reached. 37nm and 55nm: the first peak and trough of the average energy, respectively. 200nm and 250nm: electron energy is in equilibrium with the electric field.



(a)



(b)

Figure 4.8: Electron (a) and hole (b) spatial-transient II coefficients in different homogeneous electric field.

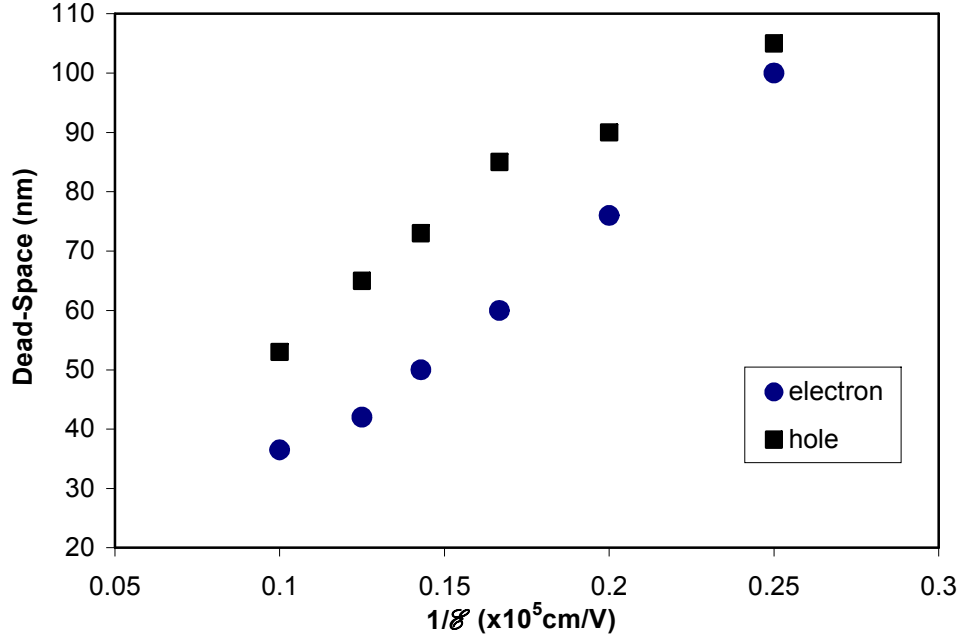


Figure 4.9: Electron's and hole's dead-space versus electric field.

The hole spatial-transient II coefficient is also calculated using MC simulations and plotted in figure 4.8b. As expected, the behaviors of the hole II coefficient are similar to electrons'. There exists a dead-space distance in which holes cannot cause II. The II coefficient is in equilibrium with the electric field when holes travel far enough.

Both electron and hole dead-space distances become shorter at higher electric field, and they are plotted as a function of the electric field in figure 4.9. Dead-space is estimated from $x=0$ to the position where the II coefficient is about half of the steady state value. The dead-space is shorter for electrons than for holes at all electric field values (figure 4.9). This is consistent with reports that the II threshold energy is larger for holes than for electrons [17, 20]. It is worth noting that if the dead-space distances are to be calculated by $dn,p=E_{Tn,p}/\mathcal{E}$, then $E_{Tn,p}$ are not constant. This partly explains the various values of the threshold energy reported in literatures [10, 21-22].

4.3.4 Evaluations of the experimentally extracted II coefficients

II coefficients were measured by different groups using the photo-multiplication method [10, 21, 23-24]. The experimentally-extracted II coefficients were very different from one another (figure 4.10). For electrons traveling from left to right (0 to W_D) and holes from right to left (W_D to 0) inside depletion regions, the following equations were employed by all of these groups to extract the II coefficients:

$$1 - \frac{1}{MF_n} = \int_0^{W_D} \alpha \exp\left(\int_0^x \beta - \alpha dx'\right) dx \quad (4.19a)$$

$$1 - \frac{1}{MF_p} = \int_0^{W_D} \beta \exp\left(\int_x^{W_D} \alpha - \beta dx'\right) dx \quad (4.19b)$$

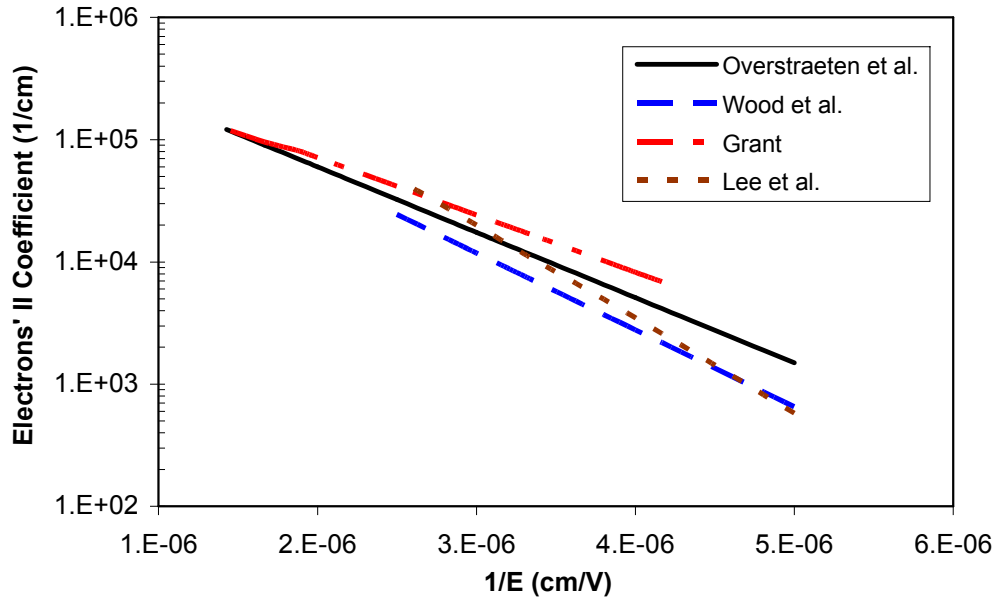


Figure 4.10: Experimental II coefficients extracted in Ref. [10, 21, 23-24].

Equation (4.19a) describes the II current gain due to the initiating electrons (which is the same as equation (3.9) derived in chapter 3). Equation (4.19b) is the holes' equivalent counterpart of equation (4.19a).

The validity of the result in Ref. [23] was not well-justified since the authors did not include the dead-space effect in their calculations. Lee *et al.*'s [23] data had the highest slope of all the results (figure 4.10), which was a direct consequence of ignoring the dead-space effect [10].

Figure 4.11 shows different types of diodes and their simplified electric field profiles in the experiments of Ref. [10, 21, 24]. A dead-space distance d_n was included in their calculations of the II coefficients [10, 21, 24]. The maximum electric field in the region excluding d_n was defined as effective maximum electric field $\mathcal{E}_{\max_eff}$ (figure 4.11). The values of the II coefficients were extracted based on $\mathcal{E}_{\max_eff}$ [10, 21, 24]. Therefore, the correct interpretations of $\mathcal{E}_{\max_eff}$ directly affected the accuracy of the extracted II coefficients. For example, for the same measured current gain of a diode, if $\mathcal{E}_{\max_eff}$ was interpreted to be smaller, then $\alpha(\mathcal{E})$ was extracted to be higher [10, 21, 24]. Obviously, $\mathcal{E}_{\max_eff}$ was smaller for larger d_n , and d_n was larger for larger E_T . Woods *et al.* [24] underestimated the dead-space distance by using the following simple formula (equation B2 in Ref. [24])

$$d_n = \frac{E_T}{q\mathcal{E}_{\max}} \quad (4.20a)$$

A more accurate formula was used by Overstraeten *et al.* [10] and Grant [21]:

$$E_T = q \int_0^{d_n} \mathcal{E}(x) dx \quad (4.20b)$$

Besides, the II threshold energy E_T of 1.65eV was used in Woods *et al.*'s calculations [24] while 1.8eV was used in Overstraeten *et al.*'s [10] and Grant's [21]. Therefore, dn was interpreted to be smaller and $\mathcal{E}_{\max_eff}$ higher in Woods *et al.* [24]. As a result, their reported II coefficient was smaller than both Overstraeten *et al.*'s [10] and Grant's [21] (figure 4.10).

More importantly, as shown earlier, the threshold energy E_T is not a fixed value, and it varies with the electric field. However, Overstraeten *et al.* [10], Grant [21], and Woods *et al.* [24] used fixed values of the II threshold energy. The wrong value of E_T directly affects the value of $\mathcal{E}_{\max_eff}$ in abrupt, one-sided junctions (figure 4.11). This is the case for the experiments in Ref. [21, 24]. However, for double-drift diodes (used in Ref. [10]), the value of E_T may not affect $\mathcal{E}_{\max_eff}$, which is, in fact, equal to the real maximum electric field in the depletion region (figure 4.11). Consequently, in terms of accounting for the spatial-transient effects of II, results from Ref. [10] were more reliable than the results in Ref. [21, 23-24]. In fact, most of Grant's experimental data [21] were from one-sided p^+-n diodes, and the II coefficient extracted from those diodes was higher than Overstraeten *et al.*'s [10]. However, the data extracted from the double-drift diodes (figure 13 in Ref. [21]) agreed more closely with Overstraeten *et al.*'s [10].

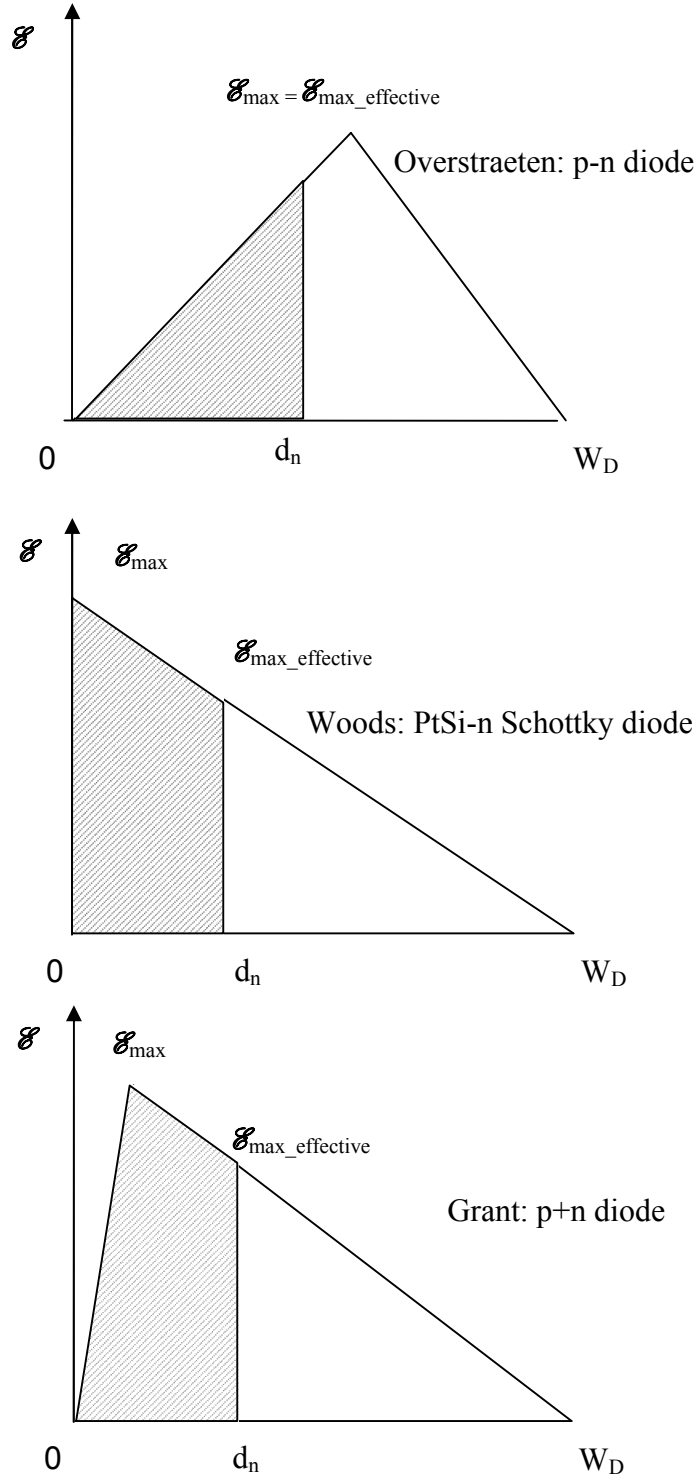


Figure 4.11: Typical electric field profiles of diodes in the photo-multiplication experiments in Ref. [10, 21, 24]. Electrons move from $x=0$ to $x=W_D$, and electrons' II can only occur from $x=d_n$ to $x=W_D$.

4.4 A proposal of simple and sufficient models for II in semiconductor devices

4.4.1 Pseudo-local electric field model

Figure 4.8 shows that the II coefficient is in equilibrium with the local electric field after a dead-space distance. Therefore, the local electric field model discussed in section 4.2.1 is still valid as long as the initial spatial-transient effect is accounted for. This work proposes a simpler and more efficient approach than both the delay model and the energy model. Taking the dead-space into account, the electron and hole II coefficients are defined as (electrons travel from left to right and holes from right to left)

$$\alpha(x) = \begin{cases} 0 & \text{for } x \leq d_n \\ \alpha(\mathcal{E}(x)) & \text{for } x > d_n \end{cases} \quad (4.21a)$$

$$\beta(x) = \begin{cases} 0 & \text{for } W - d_p \leq x \leq W \\ \beta(\mathcal{E}(x)) & \text{for } x < W - d_p \end{cases} \quad (4.21b)$$

where d_n and d_p are the dead-space distances for electrons and holes, respectively, and are functions of the electric field (figure 4.9). $\alpha(\mathcal{E})$ and $\beta(\mathcal{E})$ are pure functions of the local electric field.

The II current gain due to electrons and current gain due to holes can be solved by applying equations (4.21a, 4.21b) into equation (4.19a) and into equation (4.19b), respectively. The II current gain due to initiating electrons and holes can be easily solved numerically. Alternatively, it can also be solved analytically. The following are the complete four cases of the dead-space distances in comparison with the width of the high field region. The list is arranged in the order of decreasing dead-space distance relatively to the depletion width.

1. $d_n \geq W_D$ (both electrons and holes cannot cause II in the entire high field region)

$$1 - \frac{1}{MF_n} = 0 \quad (4.22)$$

2. $0 \leq d_n \leq W_D \leq d_p$ (holes cannot cause II in the entire high field region)

$$1 - \frac{1}{MF_n} = \int_{d_n}^{W_D} \alpha \exp\left[\int_{d_n}^x -\alpha dx'\right] dx \quad (4.23)$$

3. $0 \leq W_D - d_p \leq d_n \leq W_D$ (there is no joined region where electrons and holes can cause II)

$$1 - \frac{1}{MF_n} = \int_{d_n}^{W_D} \alpha \exp\left[\left(\int_0^{W_D-d_p} \beta dx' + \int_{d_n}^x -\alpha dx'\right)\right] dx \quad (4.24)$$

4. $0 \leq d_n \leq W_D - d_p \leq W_D$ (there is a joined region where both electrons and holes can cause II)

$$1 - \frac{1}{MF_n} = \int_0^{W_D} \alpha \exp\left(\int_0^x \beta - \alpha dx'\right) dx = I_1 + I_2 \quad (4.25)$$

where

$$I_1 = \int_{d_n}^{W_D-d_p} \alpha \exp\left[\left(\int_0^{d_n} \beta dx' + \int_{d_n}^x \beta - \alpha dx'\right)\right] dx \quad (4.25a)$$

$$I_2 = \int_{W_D-d_p}^{W_D} \alpha \exp\left[\left(\int_0^{d_n} \beta dx' + \int_{d_n}^{W_D-d_p} \beta - \alpha dx' + \int_{W_D-d_p}^x -\alpha dx'\right)\right] dx \quad (4.25b)$$

Equations (4.22-4.25) describe a complete set of II current gain due to initiating electrons. These equations are also equivalently valid for initiating holes by substituting dn by dp . It is worth

noting again that the II coefficients $\alpha(\mathcal{E})$ and $\beta(\mathcal{E})$ in equations (4.22-4.25) are **pure** functions of the local electric field.

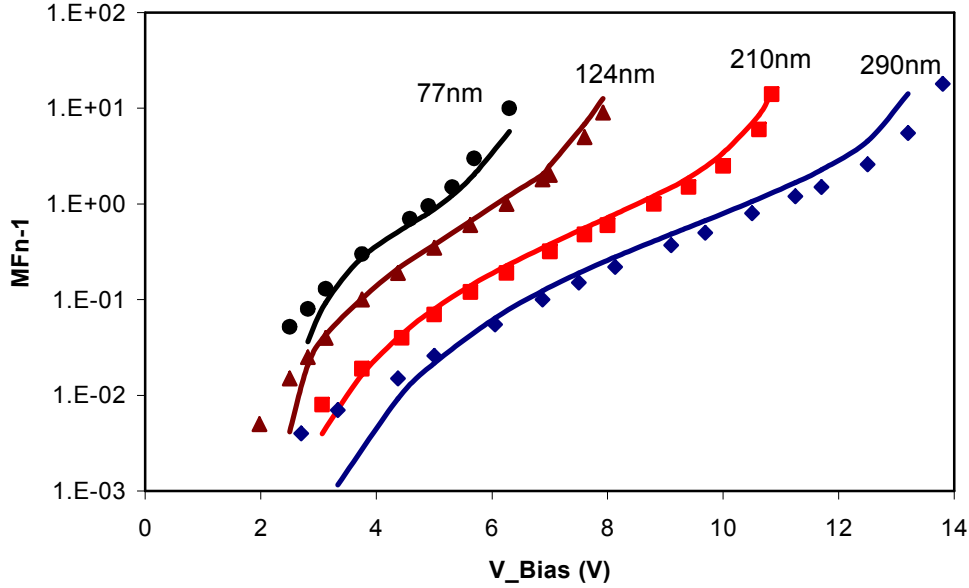


Figure 4.12: II current gain in p-i-n diodes versus bias. Lines are calculated results using the pseudo-local electric field model. Dots are the experimental data [9].

Figure 4.12 shows the current gain calculated by the pseudo-local electric model. The II coefficients $\alpha(\mathcal{E})$ and $\beta(\mathcal{E})$ are taken from the experimental results reported in Ref. [10]. The electron and hole dead-space distances are calculated from the MC simulations (figure 4.9). The theoretical and the experimental results agree well with each other. The good fit confirms that the dead-space effect needs to be accounted for in short devices. This model is sufficient in predicting the II current gain in devices of different lengths and at high gain regimes.

4.4.2 Positive feedback model

Another model that offers easier analytical solutions to the II current gain is the positive feedback model. This model was analytically proven to be valid in chapter 3, and the total current gain (MF_n) with electrons as the initiating carriers for II can be expressed as

$$MF_n = \frac{M_n}{1 - (M_n - 1)(M_p - 1)} \quad (4.26)$$

where

$$M_n = \exp\left(\int_{d_n}^W \alpha dx\right) \quad (4.26a)$$

$$M_p = \exp\left(\int_{d_p}^W \beta dx\right) \quad (4.26b)$$

The electron and hole II coefficients in equations (4.26a,b) are pure functions of the local electric field. Figure 4.13 shows that the positive feedback model is equivalent to the pseudo-local electric field model. The two models are even more equivalent for shorter devices, where the electron and hole dead-space distances overlap (case 3 in section 4.4.1).

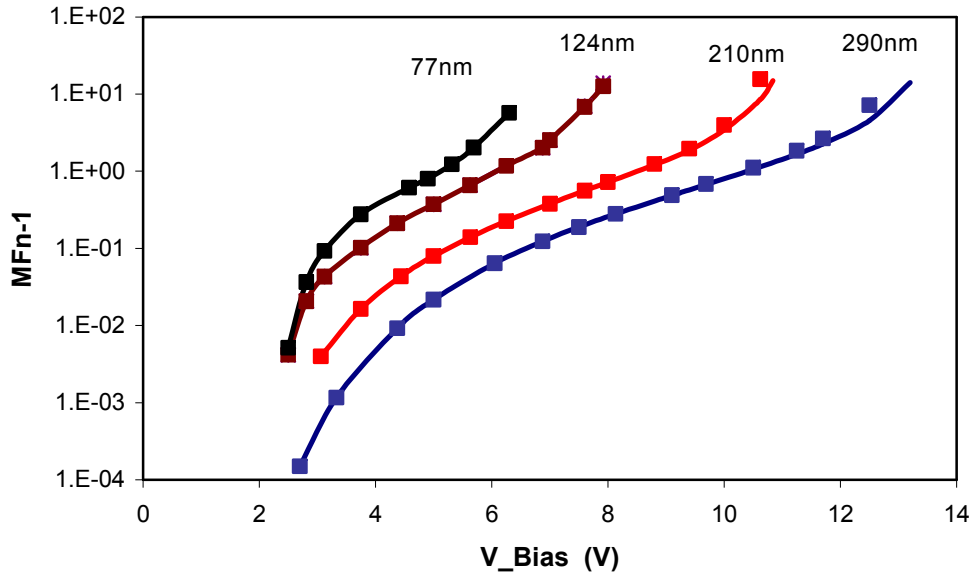


Figure 4.13: II current gain in p-i-n diodes versus bias. Lines and dots are the theoretical calculations using the pseudo-local electric field model and positive feedback model, respectively.

4.5 Summary

The three most widely-used models of II in semiconductor devices, the local electric field model, the delay model, and the energy model, are analyzed in this chapter. All of the three models have one common cause for their incomplete accuracy, which is the dead-space effect of II . This non-local effect is then studied using MC simulations for both electrons and holes.

The validity of the MC simulations is established by producing good fits to the experimental current gain of short p-i-n diodes. Several note-worthy results from the MC simulations are observed. For p-i-n diodes with electrons as the initiating carriers, the contribution of holes' II to the total current gain is more important at higher bias and more important in shorter diodes. This is due to the dead-space effect and the fact that the hole II coefficient increases faster than electron as the electric field increases. Even though the avalanche breakdown voltage is lower for shorter diodes, the electric field in the i-region is higher. For all diodes simulated in this chapter, avalanche breakdown occurs only when the II current gain due to electrons alone Mn (without holes) is larger than 2, as correctly predicted in chapter 3. It can also be predicted that, regardless of diodes' lengths, the minimum avalanche breakdown voltage is larger than 4.4 V, which is confirmed by different experimental data.

At low bias, Franz-Keldysh effect is important. This is especially true for short devices, where the dead-space consumes a large percentage of the devices' lengths. MC simulations including Franz-Keldysh effect improves the theoretical fitting with the experimental current gain for p-i-n diodes.

The electron and hole spatial-transient II coefficients are simulated by the MC programs in homogenous electric field. After the initial dead-space distances, the II coefficients are in

equilibrium with the electric field. Based on the simulated results of the spatial-transient II coefficients, the experimental data of the II coefficients reported by different groups are evaluated. The data reported in Ref. [10] is the most reliable in terms of accounting for the non-local effect of II. In addition, based on the simulated results of the spatial-transient II coefficients, more accurate models for II in devices are proposed. Simple modifications to the local electric field model correctly predict the current gain in short p-i-n diodes. The analytical formula of the positive feedback model is also confirmed to be an accurate model to describe II current gain in short devices.

Bibliography

- [1] S.A. Plimmer, J.P.R. David, D.C. Herbert, T.-W. Lee, G.J. Rees, P.A. Houston, R. Grey, P.N. Robson, A.W. Higgs, and D.R. Wight, "Investigation of impact ionization in thin GaAs diodes," *Transactions on Electron Devices*, Vol. 43, pp. 1066, 1996.
- [2] R.J. McIntyre, "A new look at impact ionization—Part I: A theory of gain, noise, breakdown probability, and frequency response," *Transactions on Electron Devices*, Vol. 46, pp. 1623, 1999.
- [3] Wieslaw A. Lukaszek, Aldert Van Der Ziel, and Eugene R. Chenette, "Investigation of the transition from tunneling to impact ionization multiplication in Silicon p-n junctions," *Solid-State Electronics*, Vol. 19, pp. 57-71, 1976.
- [4] A. Di Carlo and P. Lugli, "Dead-space effects under near-breakdown conditions in AlGaAs/GaAs HBT's", *Electron Device Letters*, Vol. 14, pp. 103, 1993.
- [5] S.P. Wilson, S. Brand, and R.A. Abram, "Avalanche multiplication properties of GaAs calculated from spatially transient ionization coefficients," *Solid-State Electronics*, Vol. 38, pp. 2095, 1995.
- [6] C. Canali, P. Pavan, A. Di Carlo, P. Lugli, R. Malik, M. Manfredi, A. Neviani, L. Vendrame, E. Zanoni, and G. Zandler, "Experimental and Monte Carlo analysis of impact-ionization in AlGaAs/GaAs HBT's," *Transactions on Electron Devices*, Vol. 43, pp. 1769, 1996.
- [7] E.F. Crabbé, J.M.C. Stork, G. Baccarani, M.V. Fischetti, and S.E. Laux, "The impact of non-equilibrium transport on breakdown and transit time in bipolar transistors," *International Electron Devices Meeting 1990. Technical Digest. IEEE*. 1990, pp. 463.
- [8] P. Palestri, L. Selmi, G.A.M. Hurkx, J.W. Slotboom, and E. Sangiorgi, "Energy dependent electron and hole impact ionization in Si bipolar transistors," *International Electron Devices Meeting 1998. Technical Digest. IEEE*. 1998, pp.885.
- [9] P. Agarwal, M.J. Goosens, V. Zieren, E. Aksen, and J.W. Slotboom, "Impact Ionization in Thin Silicon Diodes," *Electron Device Letters*, Vol. 25, pp. 807, 2004.
- [10] R. Van Overstraeten and H. De Man, "Measurement of the ionization rates in diffused silicon p-n junctions," *Solid-State Electronics*, Vol. 13, pp. 583-608, 1970.
- [11] Y. Kamakura, H. Mizuno, M. Yamaji, M. Morifuji, K. Taniguchi, C. Hamaguchi, T. Kunikiyo, and M. Takenaka, "Impact Ionization Model for Full Band Monte Carlo Simulation," *Journal of Applied Physics*, Vol. 75, pp. 3500, 1994.

- [12] T. Kunikiyo, M. Takenaka, M. Morifuji, K. Taniguchi, and C. Hamaguchi, "A model of impact ionization due to the primary hole in silicon for full band Monte Carlo simulation," *Journal of Applied Physics*, Vol. 79, pp. 7718, 1996.
- [13] J. W. Slotboom, G. Streutker, M. J. Dort, P.H. Woerlee, A. Pruijboom, D.J. Gravesteijn, "Non-local impact ionization in silicon devices," *International Electron Devices Meeting 1991. Technical Digest. IEEE*. 1991, pp 127.
- [14] D. Chang and J.G. Fossum, "Simplified energy-balance model for pragmatic multi-dimensional device simulation," *Solid-State Electronics*, Vol. 41, pp. 1795, 1997.
- [15] D.J. Massey, J. P. R. David, C. H. Tan, B. K. Ng, G. J. Rees, D. J. Robbins, and D. C. Herbert, "Impact ionization in submicron silicon devices," *Journal of Applied Physics*, Vol. 95, pp. 5931, 2004.
- [16] Andrea Savio, Stephane Monfray, Clément Charbuillet, and Thomas Skotnicki, "On the limitations of silicon for I-MOS integration," *Transaction on Electron Devices*, Vol. 56, pp. 1110, 2009.
- [17] S.L. Miller, "Ionization rates for holes and electrons in silicon," *Physical Review B*, Vol. 105, pp. 1246, 1957.
- [18] J. Bude, K. Hess, and G.J. Iafrate, "Impact ionization: beyond the Golden Rule," *Semiconductor Science and Technology*, Vol. 7, pp. B506, 1992.
- [19] Wolfgang Quade, Eckehard Schöll, Fausto Rossi, and Carlo Jacoboni, "Quantum theory of impact ionization in coherent high-field semiconductor transport," *Physical Review B*, Vol. 50, pp. 7398, 1994.
- [20] Y. Okuto and C. R. Crowell, "Energy-conservation considerations in the characterization of impact ionization in semiconductors," *Physical Review B*, Vol. 6, pp. 3076, 1972.
- [21] W. N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields," *Solid-State Electronics*, Vol. 16, pp. 1189, 1973.
- [22] A. Spinelli, A. Pacelli, and A.L. Lacaita, "Dead space approximation for impact ionization in silicon," *Applied Physics Letters*, Vol. 69, pp. 3707, 1996.
- [23] C. A. Lee, R. A. Logan, R. L. Batdorf, J. J. Kleimack, and W. Wiegmann, "Ionization Rates of Holes and Electrons in Silicon," *Physical Review*, Vol. 134, No. 3A, 1964.
- [24] M. H. Woods, W. C. Johnson and M. A. Lampert, "Use of a Schottky Barrier to Measure Impact Ionization Coefficients in Semiconductors," *Solid-State Electronics*, Vol. 16, pp. 381, 1973.

Chapter 5

Study of mixed tunneling and impact ionization in short p-n diodes

5.1 Introduction

Terahertz frequency range (0.1THz-10THz) has attracted increasing interests for its promising applications in communications, imaging, sensing, and spectroscopy [1]. Impact ionization avalanche transit time (IMPATT) diodes produce sufficient power in the millimeter and sub-millimeter frequency range. Si IMPATT diodes were shown to achieve oscillation in a fundamental mode up to 341GHz [2] and in a harmonic mode up to 423 GHz [3]. The upper limit of Si IMPATT (around 400 GHz) is due to several factors including series resistance, the saturation of the impact ionization rate with respect to electric field, and diffusion-aided spreading of the injected current pulse [4]. Si tunnel transit time (TUNNETT) diodes can operate at a much higher frequency than IMPATT [4]. TUNNETT diodes also have superior noise performance than IMPATT [4-5]. However, RF power output of TUNNETT is smaller than that of IMPATT [5].

It was shown that mixed tunneling and avalanche transit time (MITATT) diodes are capable of generating useful power in the terahertz frequency regime [6-7]. Unfortunately, basic principles of mixed tunneling and II have not been well established. Elta *et al.* [5] demonstrated

that the dc models of mixed tunneling and avalanche current developed in Ref. [8-9] were not consistent with the experimental results [11]. Elta *et al.* [5] proposed a new model including the dead-space effect. In their model [5], they assumed that II due to the thermally-generated carriers dominates II due to the tunneling-generated carriers. This assumption is not correct and will be shown later in this chapter by means of MC simulations. The dc characteristics of the mixed tunneling and II current in short p-n diodes are investigated in this chapter.

5.2 Modeling mixed tunneling and II current

5.2.1 Tunneling current

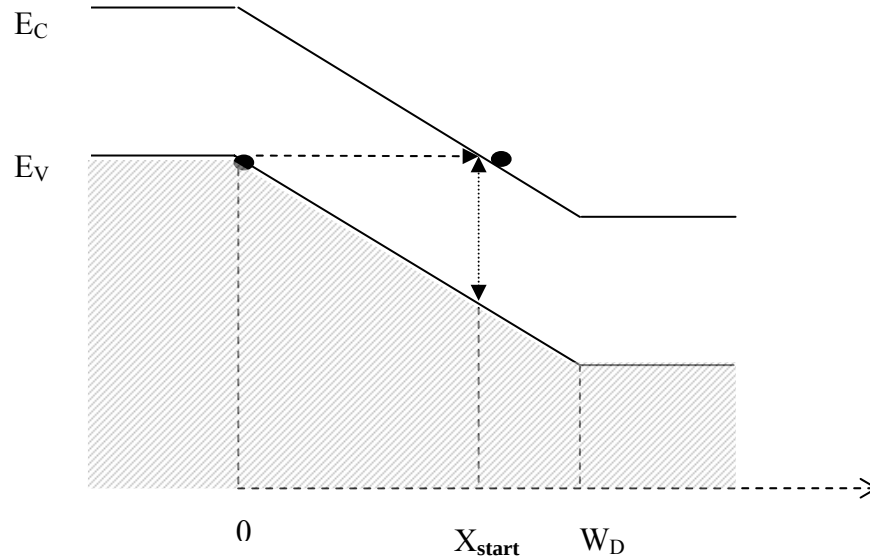


Figure 5.1: Band diagram of a p-n diode showing a typical tunneling process. A valance band electron tunnels from $x=0$ to the conduction band at position $x=X_{start}$.

Figure 5.1 shows a simplified band diagram of a p-n diode in a reverse bias. Electrons in the valance band tunnel to the conduction band and travel toward the n-region. The tunneling current can be modeled as the generation current inside the high field region. The field-dependent tunneling generation rate, derived by E.O. Kane [10], has the form:

$$G_T(\mathcal{E}) = A_T \mathcal{E}^2 \exp\left(-\frac{B_T}{\mathcal{E}}\right) \quad (5.1)$$

where the tunneling rate coefficients A_T and B_T are:

$$A_T = \frac{1}{2\pi^2} \left(\frac{q}{E_G}\right)^2 \left(\frac{E_G}{2\pi\hbar}\right) \left[\frac{2M_r^* E_G}{\hbar^2}\right]^{1/2} \quad (5.1a)$$

$$B_T = \frac{\pi}{4} \left(\frac{E_G}{q}\right) \left[\frac{2M_r^* E_G}{\hbar^2}\right]^{1/2} \quad (5.1b)$$

where E_G : band gap energy, and $\hbar$: reduced Planck constant, and

$$M_r^* = \frac{M_c^* M_v^*}{M_c^* + M_v^*} \quad (5.1c)$$

where M_c^* and M_v^* are the conduction and valance band effective mass, respectively.

Tunneling current density in a p-n junction can be calculated by summing the tunneling generation rate in the depletion width:

$$J_T = q \int_0^{W_D} A_T \mathcal{E}^2 e^{-\frac{B_T}{\mathcal{E}}} d\mathcal{E} \quad (5.2)$$

Similar to an II process, there also exists a dead-space distance for tunneling within the depletion width. Figure 5.1 shows that from $x=0$ to $x=x_{\text{start}}$, there are no electrons that can tunnel from the valance band edge (VBE) to the conduction band edge (CBE). So the tunneling current density is adjusted as:

$$J_T = q \int_{x_{\text{start}}}^{W_D} A_T \mathcal{E}^2 e^{-\frac{B_T}{\mathcal{E}}} d\mathcal{E} \quad (5.3)$$

E.O. Kane [10] derived the tunneling generation rate in constant electric field. In the depletion region of a p-n diode, the electric field varies according to Poisson's equation. So, the electric field in equation (5.3) is taken as the average electric field along the tunneling distance. For example, for electrons generated at the $x=x_{start}$ position, the effective electric field is calculated as

$$\mathcal{E}_{eff} = \frac{\int_0^{x_{start}} \mathcal{E} dx}{x_{start}} \quad (5.4)$$

Applying equation (5.4) to equation (5.3) yields

$$J_T = q \int_{x_{start}}^{W_D} A_T \mathcal{E}_{eff}^2 e^{\frac{-B_T}{\mathcal{E}_{eff}}} d\mathcal{E}_{eff} \quad (5.5)$$

Equation (5.5) is used to fit the tunneling current of diodes IN751, IN749, and IN746, reported in Ref. [11]. All of the three samples' doping concentrations are assumed to have the box profile, with the p+-doping concentration of $1 \times 10^{19} \text{ cm}^{-3}$. Table 5.1 summarizes the effective n-doping concentrations and areas for the three diodes. The effective n-doping concentrations for the three diodes are within the range of the experimental data [11]. The areas of the three samples stay well within the experimental variation due to the wet etching process [11]. The tunneling rate coefficients A_T and B_T are also listed in table 5.1. Figure 5.2 shows that the tunneling model works well at biases where tunneling is dominant for all diodes. At high bias, the measured current is higher than the theoretical values due to the II current gain, which has not been accounted for.

	IN746	IN749	IN751
$A_T (\times 10^{22} \text{V}^{-2} \text{m}^{-1} \text{s}^{-1})$	3.53	3.53	3.53
$B_T (\times 10^7 \text{V/cm})$	2.86	2.86	2.86
Area ($\times 10^{-4} \text{cm}^2$)	4.78	4.0	4.2
$N_D (\times 10^{18} \text{cm}^{-3})$	4.73	2.97	2.15
$N_A (\times 10^{19} \text{cm}^{-3})$	1	1	1

Table 5.1: Summary of fitting parameters in calculating the tunneling current of diodes IN746, IN749, and IN751 in Ref. [11].

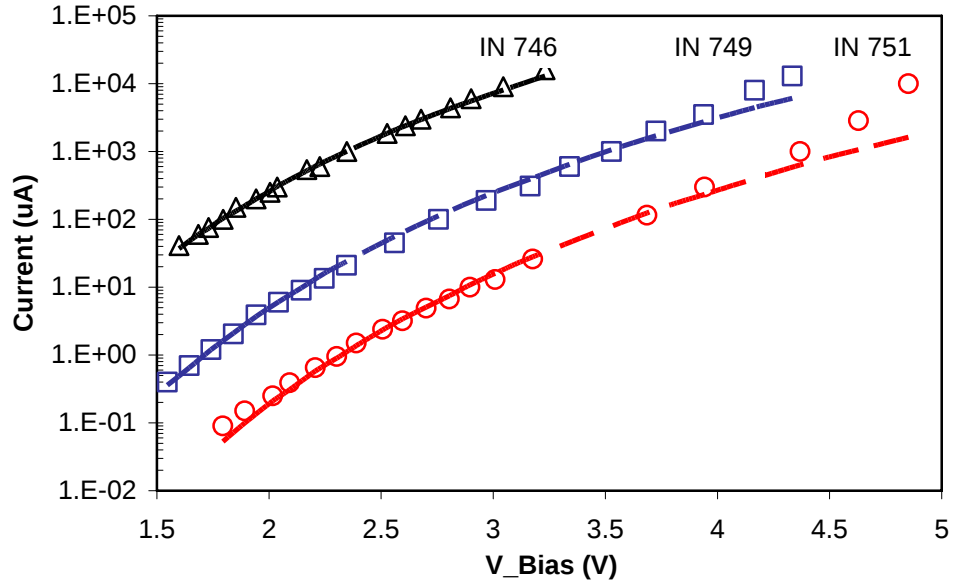


Figure 5.2: Experimental [11] (dots) and simulated current (lines) of p-n diodes versus reverse bias. The theoretical current is calculated as the tunneling generation rate with the effective electric field and including the dead-space effect.

5.2.2 Modeling an II process initiated by the tunneling-generated electrons

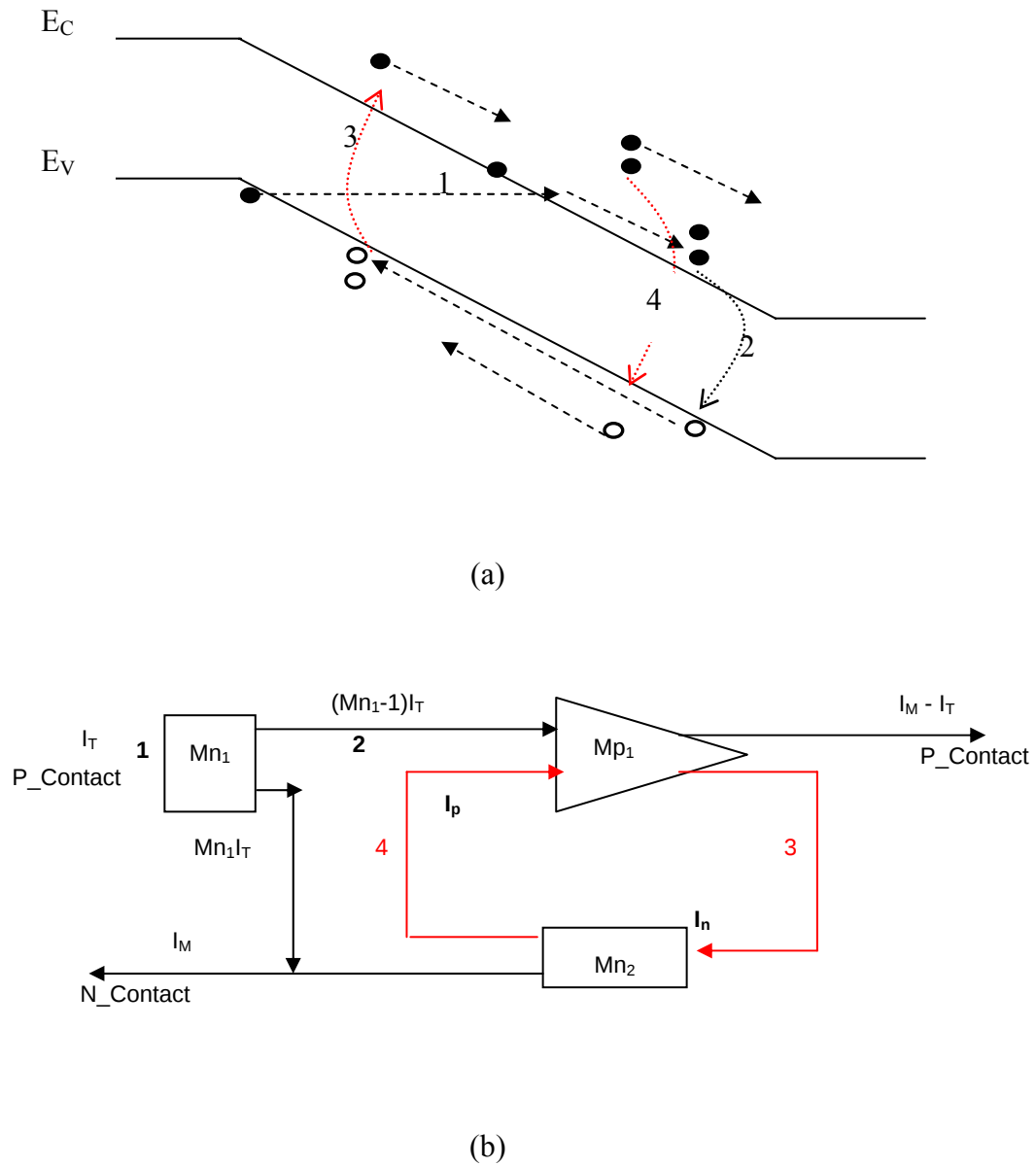


Figure 5.3: (a) Energy band diagram of a p-n junction describing a mixed tunneling and II process. (b) Block diagram modeling the mixed tunneling and II process.

Figure 5.3a shows a typical mixed tunneling and II process in a p-n diode in reverse bias.

(1) Electrons first tunnel from VBE to CBE. (2) These conduction band electrons then travel across the depletion region, gain energy, and generate e-h pairs by II. (3) Newly generated holes travel in the opposite direction and generate e-h pairs by II. (4) The newly-generated secondary electrons, in turn, travel down field and cause II.

The cycle of the mixed tunneling and II process is described by the block diagram in figure 5.3b. Compared to modeling pure II caused by the diffusing electrons from the p^+ region (described in chapter 3), processes (1) and (2) are added before the positive feedback loop (3) and (4) (figure 5.3a). The additional processes are necessary because of the existence of the tunneling dead-space. The available distance for the tunneling-generated electrons to travel and gain energy from the electric field is shorter than the entire depletion width (figure 5.3a). So, the II current gain M_{n1} from those tunneling-generated electrons is smaller than the current gain M_{n2} from the electrons traveling the entire depletion width. Therefore, the final current gain will be underestimated if M_{n1} and M_{p1} are plugged directly into the positive feed back formula.

From the block diagram (figure 5.3b), it is trivial to derive the final terminal current:

$$I = \left[\frac{M_{p1}}{1 - (M_{p1} - 1) \times (M_{n2} - 1)} \right] \times (M_{n1} - 1) \times I_T + I_T \quad (5.6)$$

or the current gain due to the II process initiated by the tunneling-generated electrons:

$$MF_T = \frac{M_{p1}(M_{n1} - 1)}{1 - (M_{p1} - 1) \times (M_{n2} - 1)} + 1 \quad (5.6a)$$

In cases that the tunneling-generated electrons do not gain enough energy to cause II, $M_{n1}=1$; consequently, equation (5.6) results in the measured current $I=I_T$. In cases that the II current gain

due to the tunneling-generated electrons is equal to the gain due to the electrons whose II distance is not restricted by the tunneling dead-space, $M_{n1}=M_{n2}=M_n$, equation (5.6) can be simplified as

$$I = \left[\frac{M_n}{1 - (M_{p1} - 1) \times (M_n - 1)} \right] \times I_T \quad (5.7)$$

Equation (5.7) is equivalent to the formula derived in chapter 3 for II current gain due to diffusing electrons.

5.3 MC simulations of mixed tunneling and impact ionization current

5.3.1 Impact ionization due to tunneling-generated electrons

Equation (5.6a) is used to calculate the total current gain due to II of the tunneling-generated electrons. MC simulations are run to calculate the stage gains M_{n1} , M_{p1} , and M_{n2} . M_{n1} is basically the average II gain due to a tunneling-generated electron (ignoring II caused by secondary holes) during its drifting in the depletion region. This stage gain is calculated by summing up all the II events caused by all of the tunneling-generated electrons and dividing the sum by those initiating electrons.

Because the tunneling-generated electrons have different original positions along the depletion width, the MC program first divides the depletion width into many small equal meshes. Each mesh contains the same large number of initial electrons. In this MC simulation, the depletion width is divided into 100 equidistance meshes, each of which contains 500 initial electrons. A finer mesh resolution and a larger number of initial electrons have also been tried, and the results are almost identical. All electrons in the same mesh are simulated before the MC program continues to the next mesh's electrons. The initial conduction band electrons (tunneling-generated electrons) in all meshes are assumed to have zero energy and have the weight of 1. Each time an electron experiences an II, its weight doubles. When the electrons originated from mesh i exit the depletion width, their weights are recorded. The II current gain due to the electrons in mesh i is calculated as:

$$M_{n1}(i) = \frac{\sum W_{e_exit}}{N_{e_in_Mesh(i)}} \quad (5.8)$$

Each mesh i carries a different weight $W_{Mesh}(i)$, according to the tunneling generation rate $G_T(i)$ at i . The II current gain M_{n1} due to the tunneling-generated electrons is calculated as:

$$M_{n1} = \frac{\sum_i M_{n1}(i)W_{Mesh}(i)}{\sum_i W_{Mesh}(i)} \quad (5.9)$$

The stage gains, M_{p1} and M_{n2} , in the positive feedback loop are calculated by the same method as described in chapter 3 with the secondary holes as the initiating carriers. The total current gain due to mixed tunneling and II is then calculated by equation (5.6a).

The mixed tunneling and II current in samples IN749 and IN751 in Ref. [11] is calculated by the MC algorithm described above. Table 5.2 summarizes the II stage gain of the tunneling-generated electrons (M_{n1}) and the II stage gain of the secondary electrons (M_{n2}) in sample IN751. Clearly, the stage gain M_{n2} , as expected, is higher than M_{n1} due to the existence of the tunneling dead-space.

Bias(V)	M_{n1}	M_{n2}
3.94	1.06	1.82
4.37	1.12	1.91
4.63	1.17	1.97
4.85	1.27	2.04

Table 5.2: Comparison of current gain between II due to the tunneling-generated electrons and II due to of the secondary electrons in sample IN751.

Figure 5.4 shows that the mixed tunneling and II current calculated by the MC simulations is in good agreement with the measured data [11]. This confirms the validity of the three separate gain stages in modeling the mixed tunneling and II current.

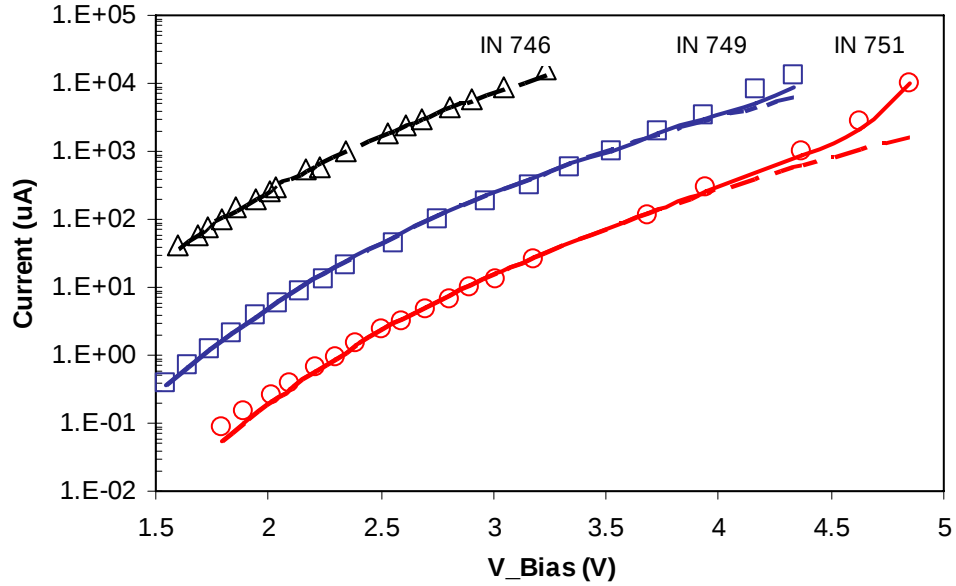


Figure 5.4: Mixed tunneling and II current versus reverse bias in different p-n diodes. Dots are the measured values [11]. Broken lines are the simulated tunneling current. Solid lines are the results of MC calculations of mixed tunneling and II.

Determining the voltage and current at the onset of II is important for transit-time device designers to set the appropriate dc bias. For example, if the bias is much lower than the onset of II, the device may operate in TUNNETT mode. On the other hand, if the bias is much higher than the onset of II, the device may operate in IMPATT mode. The voltages at which the II process starts to play a more dominant role in determining the total current are about 4V and 4.3V for IN749 and IN751 diodes, respectively (figure 5.4). These numbers are close to the 4.1V (IN749) and 4.4V (IN751) reported in Ref. [11] through noise measurements. Figure 5.2 also shows that II starts playing a more important role at a higher current level for more heavily-doped p-n diodes. This implies that, at the same II current gain, the tunneling current is higher in more heavily-doped diodes. This is consistent with reports that breakdown is dominated by tunneling current in more heavily doped p-n junctions [11-12]. This can be explained as follows.

II depends on both the electric field and the distance available for the carriers to travel and gain energy. Consequently, even though the avalanche breakdown voltage is lower for shorter diodes, the electric field is higher due to the II dead-space effect (as observed in chapter 4). The tunneling generation rate is an exponential function of the electric field (equation 3.5). Therefore, at avalanche breakdown, the tunneling current is exponentially higher in more heavily-doped diodes.

5.3.2 Impact ionization due to diffusing current

The current in a reverse-biased p-n junction is composed of tunneling current and diffusion current. Diffusing carriers are not affected by the tunneling dead-space effect. Therefore, the available II distance and II current gain of the diffusing carriers are larger than those of the tunneling-generated carriers. MC simulations are run to calculate the II gain due to the diffusing holes M_p and the feedback gain of the secondary electrons M_n . Table 5.3 summarizes each stage gain in the positive feedback model due to the primary diffusion holes (M_p , M_n) and due to tunneling-generated electrons (M_{n1} , M_{p1} , M_{n2}). For all biases larger than 3.9V of sample IN751, $M_n \approx M_{n2}$ and $M_p \approx M_{p1}$. The gain of the positive feed back loop (M_{p1} , M_{n2}) is close to that of the positive feedback loop (M_p , M_n). The total II current gain due to the tunneling-generated electrons, as expected, is smaller than the gain due to the diffusing carriers.

The magnitude of the diffusing current in sample IN751 can be easily calculated as the reverse saturation current of a p-n diode [12].

$$I_S = A \frac{qD_p p_{n0}}{L_p} \quad (5.10)$$

where A is area = $4.2 \times 10^{-4} \text{cm}^2$; D_p is diffusion coefficient $D_p = kT/q(\mu) \sim 5.2 \text{cm}^2/\text{s}$ [12]; p_{n0} is the thermal-equilibrium hole concentration in the n^+ region $p_{n0} = n_i^2 / N_D$; and L_p is the diffusion length $L_p = \sqrt{D_p \tau_p}$, where τ_p is the minority carrier life time; $\tau_p = 0.3 \mu\text{s}$ with $N = 2.15 \times 10^{18} \text{cm}^{-3}$ [13]. The diffusion current of sample IN751 is about $3 \times 10^{-17} \text{A}$, which is many orders of magnitude smaller than the tunneling current (figure 5.2). Therefore, the contribution of the diffusion current and its II in diode IN751 is negligible compared to the tunneling current and its II. This is even more valid for more heavily-doped p-n diodes, since the tunneling current is more dominant over the diffusion current. This result invalidates the assumption in Ref. [5] that the contribution of II due to tunneling-generated electrons is small compared to that of thermally-generated carriers.

Bias(V)	M_{nl}	M_{pl}	M_{n2}	M_p	M_n	(M_{nl}, M_{pl}, M_{n2})	(M_p, M_n)
3.94	1.06	1.35	1.82	1.28	1.76	1.11	1.63
4.37	1.12	1.55	1.91	1.51	1.92	1.37	2.84
4.63	1.17	1.71	1.97	1.61	1.95	1.96	3.83
4.85	1.27	1.87	2.04	1.8	1.98	6.30	8.33

Table 5.3: Comparison of current gain between II due to primary diffusion current and II due to primary tunneling current in sample IN751.

5.4 Summary

Tunneling current is modeled as the generation current and derived by Kane [1]. Similar to II, there exists a distance inside the depletion region where tunneling generation cannot occur. Using the effective electric field in Kane's tunneling generation model, the measured current of different diodes in Ref. [11] is well fit. II initiated by the tunneling electrons is calculated based on the positive feedback model. Due to the existence of the tunneling dead-space, an additional gain stage is added, and its output acts as the input to the positive feedback loop. It is also demonstrated that the diffusion current and its II current gain are negligible compared to the total current in a short p-n diode. In addition, breakdown is more dominated by the tunneling mechanism than by the II mechanism in more heavily-doped diodes.

Bibliography

- [1] P.H. Siegel, "THz technology: an overview, " in "Terahertz sensing technology, Vol. 1: Electronic devices and advanced systems technology," edited by D.L. Woolard, W.R. Loerop, M.S. Shur, World Scientific, 2003
- [2] L.S. Bowman and C.A. Burrus, Jr., "Pulse-driven silicon p-n junction avalanche oscillators for the 0.9 to 20mm band," *Transactions on Electron Devices*, Vol. 14, pp. 411, 1967.
- [3] M. Ohmori, T. Ishibashi, and S. Ono, "Dependency of the highest harmonic oscillation frequency on junction diameter of IMPATT diodes," *Transactions on Electron Devices*, Vol. 24, pp. 1323, 1977.
- [4] M.E. Elta and G.I. Haddad, "High-frequency limitations of IMPATT, MITTAT, and TUNNETT mode devices," *Transactions on Microwave Theory and Techniques*, Vol. 27, pp. 442, 1979.
- [5] M.E. Elta and G.I. Haddad, "Mixed tunneling and avalanche mechanisms in p-n junctions and their effects on microwave transit-time devices," *Transactions on Electron Devices*, Vol. 25, pp. 694, 1978.
- [6] X. Bi, J.R. East, U. Ravaioli, G.I. Haddad, "Analysis and design of Si terahertz transit-time diodes," *Solid-State Electronics*, Vol. 50, pp. 889, 2006.
- [7] X. Bi, J.R. East, U. Ravaioli, G.I. Haddad, "A Monte Carlo study of Si/SiGe MITATT diodes for terahertz power generation" *Solid-State Electronics*, Vol. 52, pp. 689, 2008.
- [8] M.S. Tyagi, "Zener and avalanche breakdown in silicon alloyed p-n junctions," *Solid-State Electronics*, Vol. 11, pp. 99, 1968.
- [9] S.P. Kwok and G.I. Haddad, "Effects of tunneling on an IMPATT oscillator," *Journal of Applied Physics*, Vol. 43, pp. 3824, 1972.
- [10] E. O. Kane, "Zener tunneling in semiconductors," *Journal of Physics and Chemistry of Solids*, Vol. 12, pp. 181-188, 1959.
- [11] Wieslaw A. Lukaszek, Aldert Van Der Ziel, and Eugene R. Chenette, "Investigation of the transition from tunneling to impact ionization multiplication in Silicon p-n junctions," *Solid-State Electronics*, Vol. 19, pp. 57-71, 1976.
- [12] See S.M. Sze, "Physics of Semiconductor Devices," John Wiley & Sons, 1981.

- [13] .E. Law, E. Solley, M. Liang, D.E. Burk, “Self-consistent model of minority-carrier lifetime, diffusion length, and mobility,” *Electron Device Letters*, Vol. 12, pp. 401, 1991.

Chapter 6

Investigating paradoxical behaviors of impact ionization and modeling substrate current in deep sub-micrometer MOSFETs

6.1 Introduction

As temperature increases, carrier-phonon scattering increases, and fewer carriers can gain energy higher than the impact ionization threshold energy (E_T). Consequently, the number of electron-hole pairs generated by II per unit distance decreases with increasing temperature. This has been widely proven by both experimental data and MC simulations [1-5.] However, as MOSFETs were scaled down to the sub-micrometer regime, II current was experimentally shown to increase with temperature, especially at low biases [6-9]. This counter-intuitive phenomenon has caught wide attention partly because the hot carrier effect and thus the device reliability are monitored by the II substrate current [10]. By MC simulations, Mastrapasqua *et al.* [7] attributed this paradoxical behavior to the memory effect. Mastrapasqua *et al.* [7] observed that the distribution function of carriers in the pinch-off region has the high-energy tail whose effective temperature is that of the lattice. This high-energy thermal tail represents the memory of the initial thermal distribution injected into the pinch-off region. So, at higher lattice

temperature, there are more carriers having energy higher than the II threshold energy, and higher II current is observed. By MC simulations, Sano *et al.* [11] showed that both memory effect and temperature-dependent bandgap effect are the major factors in determining the II current in sub-0.1 μm devices. Also by MC simulations, Fischetti *et al.* [12] argued that the dynamically-screened Coulomb interactions are the dominant force on the II current at biases higher than the bandgap energy. Meanwhile, at biases close to or lower than the bandgap energy, the temperature-dependent bandgap effect dominates. Overall, these explanations were mostly reasoned by qualitative comparisons using MC simulations. It is important to develop an analytical formula for II current modeling and device simulations. Eitan *et al.* [13] derived an analytical expression for the II current at a low-bias regime. The derivation was not well-founded since it over-simplified the distribution function. Besides, Eitan *et al.* [13] used the electron II coefficient α as a pure function of the local electric field, which is not accurately applicable to deep sub-micrometer devices [5, 14]. This work will show, for the first time, a complete derivation of the analytical formula for II current in low supply bias regime.

6.2 Investigating and modeling low voltage impact ionization in deep sub-micrometer MOSFETs

In this work, an analytical formula for II current is derived for p-MOSFETs. The primary current in the channel is hole current. Holes travel from the source junction, gain energy in the channel, and experience II near the drain junction. The analytical result obtained here is equally valid for n-MOSFETs. Thornber [15] derived the steady state current spatial gradient through Boltzmann transport equation:

$$dI_n = I_n(x)\alpha(x)dx + I_p(x)\beta(x)dx \quad (6.1)$$

where α (cm^{-1}), β (cm^{-1}) are II coefficients of electrons and holes, respectively.

At low supply voltage in p-MOSFETs, especially sub-bandgap voltage, electron current in the channel is negligible. In addition, the secondary holes' contribution to the channel's hole current is also negligible. So, equation (6.1), integrated over the channel length (0, W), is simplified as

$$I_n = I_p \int_0^W \beta(x)dx \quad (6.2)$$

The impact ionization factor M is

$$M = \frac{I_n}{I_p} = \int_0^W \beta(x)dx \quad (6.3)$$

At sub-bandgap bias, the macroscopic II coefficients that are experimentally extracted cannot be used since they represent more or less the values of carriers with average energy. The microscopic value of the hole II coefficient β (cm^{-1}) is defined as [15]

$$\beta(x) = \frac{\sum_{VB} \int_{BZ} R_{II}(\mathbf{k}) f(x, \mathbf{k}) d^3k}{J_p(x)} \quad (6.4)$$

where VB: initial conduction and valance band; $R_{II}(\mathbf{k})$: II rate (1/s).

For carriers whose energy is much higher than the optical phonon energy, the momentum relaxation time is much smaller than the energy relaxation time, so it is reasonable to average the II rate over wave vector directions [16-17.] This is justified by MC calculations in Ref. [18.] The distribution function $f(\mathbf{k})$ is also isotropic and approximated by $f(E)$ [18.] So, equation (6.4) becomes

$$\beta(x) = \frac{\int_{E_T}^{\infty} R_{II}(E) f(x, E) g(E) dE}{J_p(x)} \quad (6.5)$$

where $g(E)$:density of states. $R_{II}(E)$, at low energy, generally has the formula [19]:

$$R_{II}(E) = R_0 (E - E_T)^a \quad (6.5a)$$

where R_0 is a constant.

Equation (6.5) becomes

$$\beta(x) = \frac{\int_{E_T}^{\infty} R_0 (E - E_T)^a p(x, E) dE}{J_p(x)} \quad (6.6)$$

where $p(x, E) = f(x, E)g(E)$ is the number of holes at position x which have energy E.

For typical p-MOSFETs operating in the velocity saturation regime, the lateral electric field increases linearly near the source and increases exponentially near the drain [20]. Equations (6.7a) and (6.7b) represent a typical lateral electric field profile for 0.1 μ m p-MOSFET with the drain source bias $V_S = 0.8$ V.

$$\varepsilon = 5 \times 10^4 \frac{x}{0.05 \mu m} \left(\frac{V}{cm} \right) \text{ for } x \leq 0.05 \mu m \quad (6.7a)$$

$$\varepsilon = 5 \times 10^4 \exp\left(\frac{x - 0.05 \mu m}{0.029 \mu m}\right) \left(\frac{V}{cm} \right) \text{ for } 0.05 \mu m < x \leq 0.1 \mu m \quad (6.7b)$$

The electric field profile described in equations (6.7a) and (6.7b) is plotted in figure 6.1. It will become clear later that the analytical formula for II current is not very sensitive to details of the electric profile along the device. This is not surprising because, at low supply bias, especially when $qV_S < E_T$, II can only occur by the carriers belonging to the energy tail of the distribution function. Therefore, the energy supply available for the carriers is far more important than the electric field profile in determining the II current.

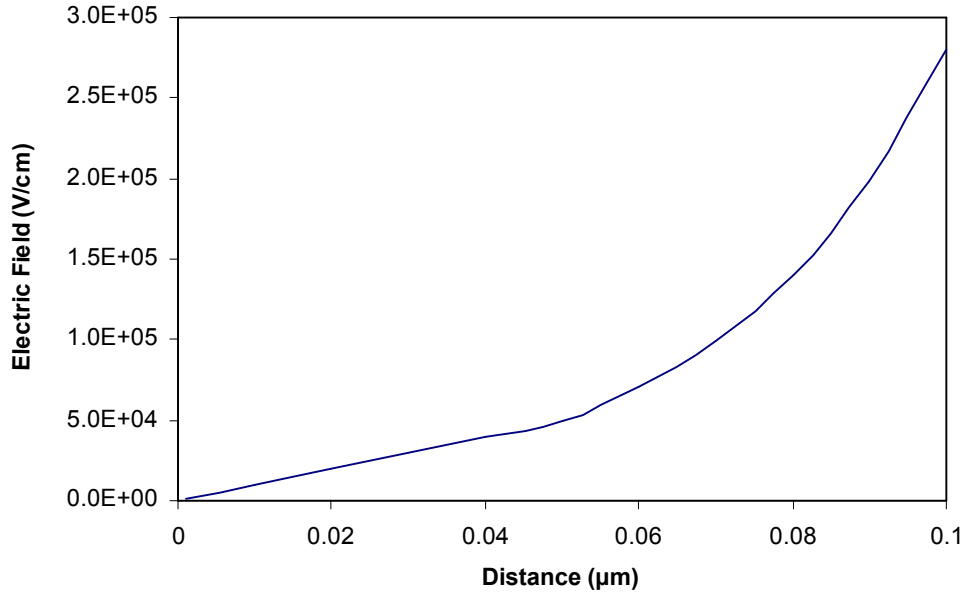


Figure 6.1: Typical p-MOSFET channel's lateral electric field in the velocity saturation mode. Channel length is $0.1 \mu m$, and the source-drain bias voltage is 0.8 V . The electric field increases linearly near the source ($x=0$) and increases exponentially near the drain ($x=0.1$).

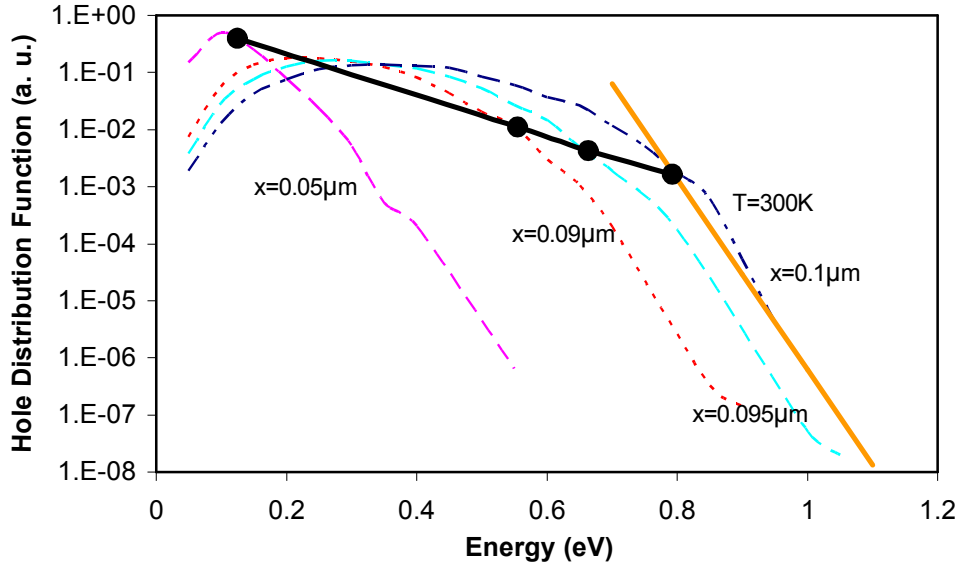


Figure 6.2: Hole distribution as a function of energy at different channel positions in a $0.1\mu\text{m}$ p-MOSFET with 0.8V source-drain bias voltage at 300K lattice temperature. The high-energy tails of distribution curves have the effective temperature equal to the lattice temperature. Filled dots are the values of holes at position x whose energy is qVx .

The hole distribution function (hdf) along the channel is calculated by MC simulations. The distribution function of high-energy holes is of the most interest since these high-energy holes are responsible for II at low biases. However, the number of the high-energy holes, particularly those with energy higher than the supply bias, is small and hence the variance of the high-energy tail of the distribution curve is excessive. The rare-state algorithm developed by Phillips and Price [21] is implemented in this MC program. The state of a hole is categorized into common states and rare states. Holes are in rare states when their energy is higher than a certain pre-set value. When a hole enters a rare state, its state (S) is stored. This hole will later be simulated by a large number of representative holes whose initial states are identical to S . The hole's transport parameters are the average values of the representative holes. Figure 6.2 shows the hdf at different positions along the $0.1\mu\text{m}$ channel under the electric field profile of equations

(6.7a, 6.7b) at lattice temperature $T_L=300\text{K}$. There are two important characteristics of the hdf in figure 6.2. First, the high-energy tail of the hdf is Maxwellian with the lattice temperature T_L as the effective temperature. Second, $p(x, qV_x)$ is an exponential function of qV_x with the effective temperature T . These observations are confirmed by other reports [7, 22]. The hdf can be formulated as follows:

$$p(x, E) = p(x, qV_x) \exp\left(\frac{qV_x - E}{k_B T_L}\right) \text{ for } E \geq qV_x \quad (6.8a)$$

$$p(x, qV_x) = p(W, qV_S) \exp\left(\frac{qV_S - qV_x}{k_B T}\right) \quad (6.8b)$$

where V_x and V_S are the potentials at positions x and W . The hole II coefficient at the position x can now be written as

$$\beta(x) = \frac{R_0 p(x, qV_x) \int_{E_T}^{\infty} (E - E_T)^a \exp\left(\frac{qV_x - E}{k_B T_L}\right) dE}{J_p(x)} \quad (6.9)$$

From equations (6.3) and (6.9), current gain can be given by

$$M = \frac{R_0 (k_B T_L)^{a+1} \left(\int_0^{\infty} u^a \exp(-u) du \right)}{J_p} \int_0^W p(x, qV_x) \exp\left(\frac{qV_x - E_T}{k_B T_L}\right) dx \quad (6.10)$$

From equations (6.8b) and (6.10), the current gain is given by

$$M = \frac{R_0 p(W, qV_S) (k_B T_L)^{a+1} \left(\int_0^{\infty} u^a \exp(-u) du \right) \exp\left(\frac{qV_S}{k_B T} - \frac{E_T}{k_B T_L}\right)}{J_p} \int_0^W \exp\left(\frac{qV_x}{k_B T_L} - \frac{qV_x}{k_B T}\right) dx \quad (6.11)$$

In the low bias regime, II mostly occurs near the channel-drain junction, where the maximum

electric field locates. So, the maximum electric field is used in equation (6.11) for simplification without sacrificing much accuracy. Taking into the account of the source-channel barrier height (ϕ_B), equation (6.11) becomes

$$M \approx \frac{R_0 p(W, qV_S) (k_B T_L)^{a+1} \left(\int_0^\infty u^a \exp(-u) du \right) \exp\left(\frac{qV_S + q\phi_B - E_T}{k_B T_L}\right)}{\epsilon_{\max} \left(\frac{q}{k_B T_L} - \frac{q}{k_B T} \right) J_P} \quad (6.12)$$

A simpler formula is much easier for device modeling. This can be achieved by eliminating the parameters that are weak functions of the applied bias. The terms outside of the exponential function are weak functions of V_S . Besides, T is usually much greater than T_L (figure 6.2). So, equation (6.12) can be simplified as

$$M \approx M_0 (T_L)^{a+2} \exp\left(\frac{qV_S + q\phi_B - E_T}{k_B T_L}\right) \quad (6.13)$$

where M_0 is treated as a fitting constant.

Equation (6.13) is valid as long as the carriers in the thermal tail of the distribution function are dominantly responsible for the II current. The dependence of the II current gain on the supply voltage for scaled MOSFETs at low bias is $\log(M) \propto V_S$. However, at high bias, a majority of the carriers capable of causing II are not located in the high-energy tail of the distribution function. Therefore, II is more accurately modeled using average energy, which is a function of the electric field. This explains the well-known relationship [10], $\log(M) \propto -\mathcal{E}_{\max}^{-1}$, which has been used to predict the substrate current in MOSFETs at bias much higher than the II threshold energy.

The analytical formula (Eq. 6.13) is used to fit the experimental data reported in Ref. [9]. At gate bias $V_G=0.4$ V and 0.8 V, the source-channel barrier heights are taken to be 0.37 eV and

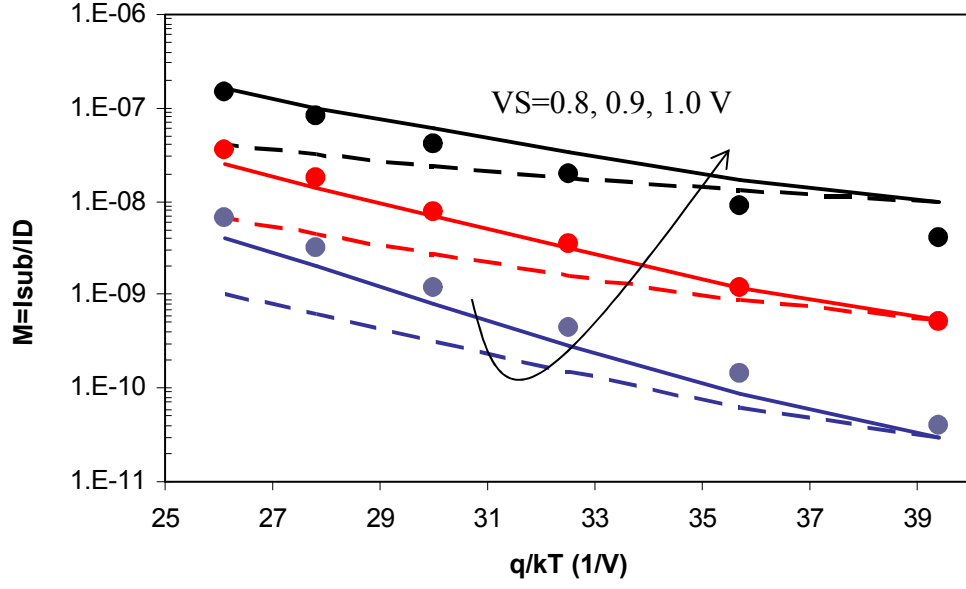
0.29 eV, respectively. The drain induced barrier lowering effect is taken to be 0.15 V/V. The II threshold energy depends on T_L by the relationship (like the assumption in Ref. [23])

$$E_T = E_T(300K) \frac{E_G(T)}{E_G(300K)} \quad (6.14)$$

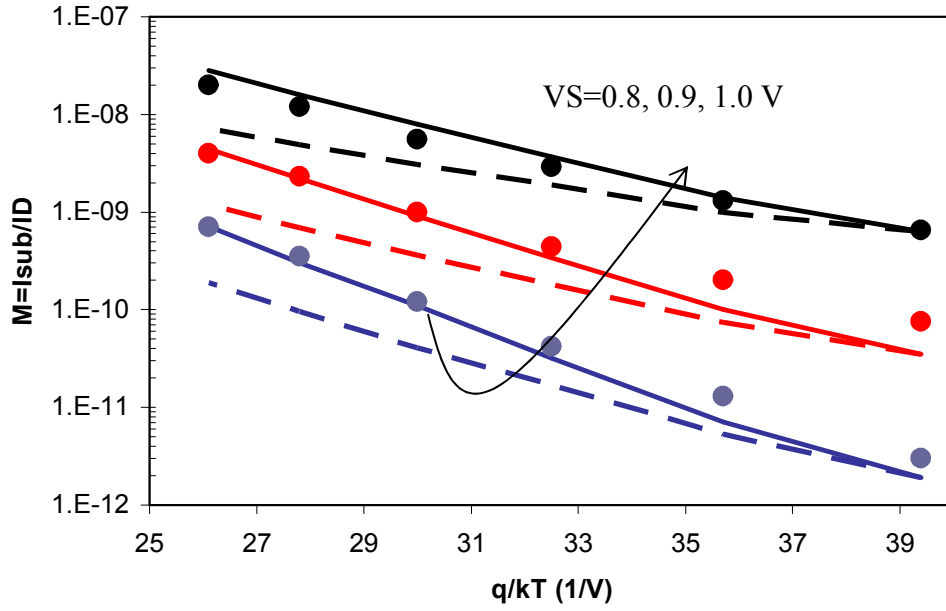
where $E_T(300K)=1.2\text{eV}$; E_G is the bandgap energy.

The theoretical calculations match the measured data well (figure 6.3), using M_0 as the fitting parameter. The temperature-dependent bandgap plays a significant role in determining the II current.

Two important observations regarding the substrate current in n-MOSFETs were demonstrated in Ref. [8]. The first is that M depends more strongly on V_G in the sub-threshold regime. The second is that, in the sub-threshold region, as the gate bias increases, the II current gain (I_{sub}/I_D) decreases for $T \geq 300\text{K}$ and increases for $T < 300\text{K}$. These two observations can be explained as follows. As V_G increases, the source-channel barrier height is reduced, so carriers gain less energy traveling across the channel. As a result, M is reduced. In MOSFETs, the source-channel barrier height is more strongly modulated by V_G in the sub-threshold regime than in inversion, so M depends more strongly on V_G in the sub-threshold regime (equation 6.13). At temperature above 300K, the dynamically-screened Coulomb interactions are negligible in determining II current at low supply voltage. However, as the temperature reduces, the high-energy tail of the carrier distribution function becomes more abrupt. Thus, the high-energy tail broadening effect of the dynamically-screened Coulomb interactions becomes more significant. Therefore, at low temperature, II current is also influenced by the dynamically-screened Coulomb interactions.



(a)



(b)

Figure 6.3: Experimental (dots) and theoretical (lines) values of the multiplication factor versus the lattice temperature at different source-drain biases with the gate bias at $V_G = 0.4V$ (a) and $V_G = 0.8V$ (b). The solid lines and dash lines show the results of the theoretical calculations with and without taking into account the temperature-dependent bandgap, respectively.

6.3 Summary

As supply bias in deep sub-micrometer MOSFETs is reduced, the following paradoxical behaviors have been observed:

- II current is detected even as voltage is scaled down below the bandgap energy.
- The substrate current increases with increasing temperature. The increase is more significant at lower supply voltage.
- In the sub-threshold region, as the gate bias increases, the II current gain (I_{sub}/I_D) decreases for $T \geq 300K$ and increases for $T < 300K$.
- At bias below the bandgap energy, $\log(I_{sub}/I_D) \propto (V_{supply})$.

The above observations cannot be explained by the conventional II current formula in Ref. [10]. In this chapter, a new analytical formula is successfully derived. At low bias in deep, sub-micrometer MOSFETs, II is caused mainly by the carriers belonging to the thermal energy tail of the channel carriers' distribution. The experimental data can be fit and explained by this newly developed model. The temperature-dependent bandgap energy plays an important role on the temperature-dependent substrate current. At temperature lower than 300K, the dynamically-screened Coulomb interactions become important as they broaden the carriers' thermal tails.

Bibliography

- [1] C.R. Crowell and S.M. Sze, "Temperature dependence of avalanche multiplication in semiconductors," *Applied Physics Letters*, Vol. 9, pp. 242, 1966.
- [2] W.N. Grant, "Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields," *Solid-State Electronics*, Vol. 16, pp. 1189, 1973.
- [3] W. Maes, K. De Meyer and R. Van Overstraeten, "Impact ionization in silicon: a review and update," *Solid-State Electronics*, Vol. 33, pp. 705 1990.
- [4] Ranbir Singh and B. J. Baliga, "Analysis and optimization of power MOSFETs for cryogenic operation," *Solid-State Electronics*, Vol. 36, pp. 1203, 1993.
- [5]: M Ershov and V Ryzhii, "Temperature dependence of the electron impact ionization coefficient in silicon," *Semiconductor Science and Technology*, Vol. 10, pp. 138, 1995.
- [6] M. Koyanagi, T. Matsumoto, M. Tsuno, T. Shimatani, Y. Yoshida, and H. Watanabe, "Impact ionization phenomenon in 0.1 μ m MOSFET at low temperature and low voltage," *International Electron Devices Meeting*, pp. 341-344, 1993.
- [7] M. Mastrapasqua, J. Bude, M. Pinto, L. Manchanda, and K. F. Lee, "Temperature and channel-length dependence of impact ionization in p-channel MOSFETs," in *1994 SymD. VLSI Tech. Dig. Tech. Papers*, p. 125-126.
- [8] K. G. Anil, S. Mahapatra, and I. Eisele, "Experimental verification of the nature of the high energy tail in the electron energy distribution in n-channel MOSFETs," *Electron Device Letters*, Vol. 22, pp. 478, 2001.
- [9] Pin Su, Ken_Ichi Goto, Toshihiro Sugii, and Chenming Hu, "A Thermal Activation View of Low Voltage Impact Ionization in MOSFETs," *Electron Device Letter*, Vol. 23, pp. 550, 2002.
- [10] C. Hu, S. Tam, F. Hsu, P. Ko, T. Chan, and K. Terill, "Hot-electron-induced MOSFET degradation – model, monitor, and improvement," *Transactions on Electron Devices*, Vol. 32, pp. 375, 1985.
- [11] Nobuyuki Sano, Masaaki Tomizawa, and Akira Yoshii, "Temperature dependence of hot carrier effects in short-channel Si-MOSFET's," *Transactions on Electron Devices*, Vol. 42, pp. 2211, 1995.
- [12] M. V. Fischetti and S. E. Laux, "Monte Carlo study of sub-band-gap impact ionization in small silicon field-effect transistors," *International Electron Devices Meeting*, pp. 305-308, 1995.

- [13] B. Eitan, D. Frohman-Bentchkowsky, and J. Shappir, "Impact ionization at very low voltages in silicon," *Journal of Applied Physics*, Vol. 53, pp. 1244, 1982.
- [14] E.F. Crabbe, J.M.C Stork, G. Baccarani, M.V. Fischetti, and S.E. Laux, "The impact of non-equilibrium transport on breakdown and transit time in bipolar transistors," *International Electron Devices Meeting*, pp. 463-469, 1990.
- [15] K.K. Thornber, "Applications of scaling to problems in high-field electronic transport," *Journal of Applied Physics*, Vol. 52, pp. 279, 1981.
- [16] B K Ridley, "Lucky-drift mechanism for impact ionization in semiconductors," *Journal of Physics C: Solid State Physics*, Vol. 15, pp. 3373, 1983.
- [17] Alan R Beattie, "Impact ionization rate and soft energy thresholds for anisotropic parabolic band structures," *Semiconductor Science and Technology*, Vol. 3, pp. 48, 1988.
- [18] Jeff D. Bude and Marco Mastrapasqua, "Impact ionization and distribution functions in sub-micron nMOSFET technologies," *Electron Device Letters*, Vol. 16, pp. 439, 1995.
- [19] Keldysh L. V., "Kinetic Theory of Impact Ionization in Semiconductors," *Soviet Physics-JETP*, Vol. 37, pp 506, 1960.
- [20] Kartikeya Mayaram, Jack C. Lee, and Cheming Hu, "A model for the electric field in lightly doped drain structures," *Transactions on Electron Devices*, Vol. 34, pp. 1509, 1987.
- [21] A. Phillips, Jr. and P. J. Price, "Monte Carlo calculations on hot electron energy tails," *Applied Physics Letter*, Vol. 30, pp. 528, 1977.
- [22] Andrea Lacaita, "Why the effective temperature of the hot electron tail approaches the lattice temperature," *Applied Physics Letter*, Vol. 59, pp. 1623, 1991.
- [23] J. Y. Tang and Karl Hess, "Impact ionization of electrons in silicon (steady state)," *Journal of Applied Physics*, Vol. 54, pp. 5139, 1983.

Chapter 7

Conclusion

7.1 Summary

Monte Carlo simulations were developed to study impact ionization for both electrons and holes. Analytical band models were used in the MC programs, and the parameters were carefully calibrated to ensure the simulations' accuracy. This work's MC programs were able to fit a wide range of experimental charge transport results in bulk silicon such as drift velocity, II rate, II coefficients, and quantum yield. Besides, this work was able to fit previous full-band MC results of average energy, phonon scattering rates, and energy distribution function.

To help enable simulations of an II process involving both electrons and holes, a positive feedback model was designed. Taking advantage of the II dead-space effect, this model decoupled the II of electrons and holes. The analytical formula of the model was proven accurate, especially for II in short devices. This model suggested that the breakdown voltage is the same for both the electron-initiating and hole-initiating II processes in the same high field region. In addition, the positive feedback model showed that current gain from the electron-initiating II process is always higher than current gain from the hole-initiating II process within the same high field region. An efficient algorithm was also developed for spatial-transient MC

simulations using temporal meshes instead of the traditional spatial meshes. Besides, II due to the secondary carriers was counted by assigning weight to each initiating carrier. Each time a carrier impact-ionized, its weight doubled, representing both the original carrier and the secondary carrier.

The spatial-transient effects of II were identified to be the main cause for the failures of different well-known II models for semiconductor devices. Microscopic parameters of II were investigated by the MC programs by simulating the current gain in short p-i-n diodes and in homogeneous electric field. The theoretical results of the current gain of p-i-n diodes fit well to the experimental data for various diodes' lengths, further confirming the validity of the MC simulations. This work showed that, at low biases, Franz-Keldysh effect was important and could not be neglected, particularly for shorter devices. With Franz-Keldysh effect taken into account, carriers could experience II at energy even lower than the bandgap energy. From the MC simulations, it was clear that the effective II threshold energy varies with the electric field's strengths. This partly explained the difference in the reported values of the effective threshold energy. The spatial-transient II coefficients were also calculated by means of MC simulations. From the physical findings of the MC simulations, the experimentally extracted values of II coefficients were evaluated. One of the main factors responsible for the disagreements of the experimental II coefficient values was the dead-space effect. It was then concluded that experiments that employed double-drift p-n junctions were more reliable to yield accurate II coefficients in terms of the spatial-transient effects. A pseudo-local electric field model was then proposed and proven to be sufficient in predicting the II current gain in short devices.

Mixed tunneling and II current in short p-n diodes was simulated by MC simulations. Tunneling current was treated as generation current. Similar to II, tunneling-generated electrons

also experienced a dead-space distance, within which there was no tunneling generation. Another gain stage was added to the positive feedback loop to accurately model the II current gain due to the tunneling-generated electrons. The extra stage was necessary due to the tunneling dead-space effect. Mixed tunneling and II current calculated by the MC simulations agreed well with the experimental data. II was less important for more heavily-doped diodes. This was due to the fact that, with higher doping concentration, the tunneling rate increased exponentially, and the depletion width available for II was smaller. The contribution of the diffusion current and its II was negligible compared to the total current in thin p-n diodes. This was the case in heavily-doped diodes even though the diffusion carriers experienced a longer effective distance for II than the tunneling-generated electrons.

II current was found to increase with increasing temperatures, and II current was observed even at bias lower than the bandgap energy. These paradoxical behaviors were explained by calculating the carrier distribution functions using MC simulations. There were always a small number of carriers in the energy distribution tail whose energy was higher than the supply bias. Those carriers were responsible for the existence of II current at very low bias. Besides, the energy distribution tail was Maxwellian and had the effective temperature of the lattice temperature. Therefore, when the lattice temperature increased, more high-energy carriers existed in the energy tail; consequently, the II current increased. Besides, at a higher lattice temperature, the bandgap energy was reduced. As a result, the effective threshold energy became smaller, and the II rate increased. An analytical formula was successfully derived for the II current in deep sub-micrometer devices operating at supply bias close to the bandgap energy. The analytical model correctly predicted the experimental II current gain in $0.1\mu\text{m}$ -pMOSFET. In short, in the sub-bandgap supply bias regime, II current needs to be derived from the tails of

the distribution functions. In this bias regime, II current gain depends exponentially on both the supply bias, $\exp(V_S)$, and the lattice temperature, $\exp(-1/T_L)$.

7.2 Suggestions for future studies

Thin SOI MOSFETs and FinFETs are among the most attractive technologies in downscaling devices due to their excellent short channel characteristics. The body thickness of SOI MOSFETs and FINFETs needs to be scaled with the scaling of gate lengths, and, consequently, the source, drain series resistances increase. Hence, a larger part of the supply bias is dropped across the source and drain regions, and the effective voltage across the channel becomes smaller. As shown in chapter 4, Franz-Keldysh effect is important for low bias operation, especially for short devices. This effect will play a more significant role in the Π current for shorter MOSFETs, FINFETs and 1T-DRAM. Therefore, a more systematic approach needs to be studied to include Franz-Keldysh effect for both electrons and holes in an Π process.

The anisotropy of Π scattering rate at low energy (see Appendix A) is significant. For example, it was calculated that, at $T=300\text{K}$, the electron Π threshold energy is 1.0eV and 1.4eV for the initial electron's wave vector in $\langle 100 \rangle$ and $\langle 110 \rangle$ directions, respectively. As devices are scaled down into deep sub-micrometer regime, the time for carriers to traverse the device length is compatible with the momentum relaxation time, and Π becomes more dependent on the directions of the electric field. This is especially true when the supply voltage is scaled down along the downscaling of the physical devices' lengths. More MC simulations need to be conducted to study the Π current as a function of the electric field direction in scaled devices.

Multiplication noise in avalanche photodiodes depends on the ratio of electron and hole Π . Based on the positive feedback model developed in chapter 3 and the dead-space results calculated in chapter 4, physical structures and dc bias of avalanche photodiodes can be

engineered to yield the desired noise performance. Besides, applying the technical findings in this work to designing more efficient MITATT diodes is also encouraged.

Appendix A

Calculations of impact ionization rates in silicon

A.1 Non-local empirical pseudopotential method of calculating Si band structures:

Cohen *et al.* [1] and Chelikowsky *et al.* [2] successfully calculated the entire conduction bands (CB) and valance bands (VB) of different semiconductors using the empirical pseudopotential method. The essential idea of this method is that tightly bound valance electrons can be modeled as nearly free electrons moving in a weak potential. The attractive potential between the nucleus and the valance electrons serves to cancel the repulsive potential between the core and valance electrons [3]. This work's calculation of Si band structures follows the procedures in Ref. [2], using 113 reciprocal lattice vectors and parameters listed in Ref. [4]. Figure A.1 shows the band structure of Si along high-symmetry axis. The lowest energy in the CB locates at $k = 0.85k_{\max}$ in the $\langle 100 \rangle$ direction, and its value is 1.08eV (bandgap energy). Along the $\langle 100 \rangle$ direction, at $k = 0.85k_{\max}$, the longitudinal and transverse masses are extracted to be $0.92m_0$ and $0.19m_0$, respectively. The energy difference between X-L valleys is 1.09eV. These extracted values are very close to the reported data in Ref. [5].

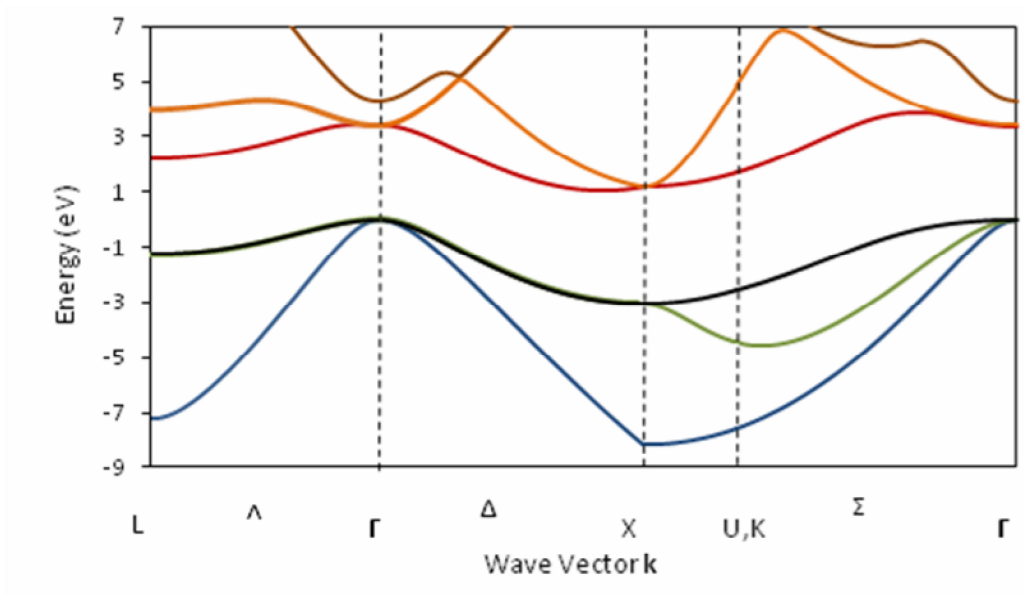


Figure A.1: Silicon band structures along different high-symmetry directions.

A.2 Integration over Brillouin zone

One approach of demonstrating procedures of integration over Brillouin zone (BZ) is to show the calculation procedure of DOS of electrons and holes. DOS is calculated by [2] using the formula:

$$N(E)\left(\frac{\text{states}}{eV \cdot cm^3}\right) = \frac{1}{N_{unitcell}} \sum_n \sum_{\mathbf{k}} \delta(E - E_n(\mathbf{k})) \quad (\text{A.1})$$

where $N_{unitcell}$: number of unit cells.

n : band index.

Converting the sum series in equation (A.1) into the continuous integration over $\mathbf{k}$ yields

$$N(E) = \frac{a^3}{4(2\pi)^3} \sum_n \int d^3k \delta(E - E_n(\mathbf{k})) \quad (\text{A.2})$$

where a is the lattice constant.

One of the most widely used and accurate methods for integration over BZ is the tetrahedron method [6-8]. Figure A.2 shows the first BZ modeled as a cuboid with

$k_x = [-\frac{2\pi}{a}; \frac{2\pi}{a}]$, $k_y = [-\frac{2\pi}{a}; \frac{2\pi}{a}]$, $k_z = [0; \frac{2\pi}{a}]$ by adding appropriate reciprocal lattice vectors to

transform the lower part of the BZ ($k_z < 0$) to the corresponding part in the upper part [8]:

$$\mathbf{k}' = \mathbf{k} + \begin{cases} \frac{2\pi}{a}(-1; -1; 1) & \text{if } k_x \geq 0 \text{ and } k_y \geq 0 \\ \frac{2\pi}{a}(-1; 1; 1) & \text{if } k_x \geq 0 \text{ and } k_y < 0 \\ \frac{2\pi}{a}(1; -1; 1) & \text{if } k_x < 0 \text{ and } k_y \geq 0 \\ \frac{2\pi}{a}(1; 1; 1) & \text{if } k_x < 0 \text{ and } k_y < 0 \end{cases} \quad (\text{A.3})$$

The modeled cuboid has the same volume as the first BZ. The cuboid is divided into small cubes, whose size is $\frac{2\pi/a}{N_{mesh}}$, where N_{mesh} is the resolution of the integration. The larger the value of N_{mesh} is, the more accurate the integration becomes.

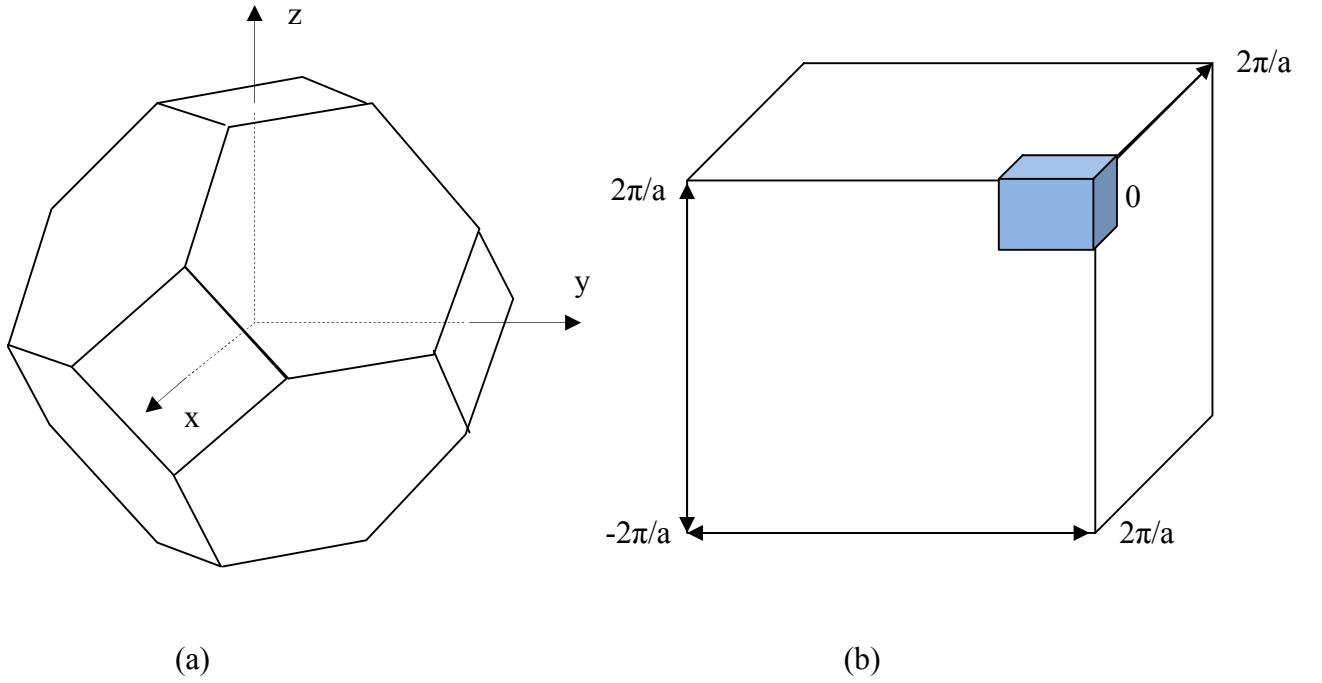
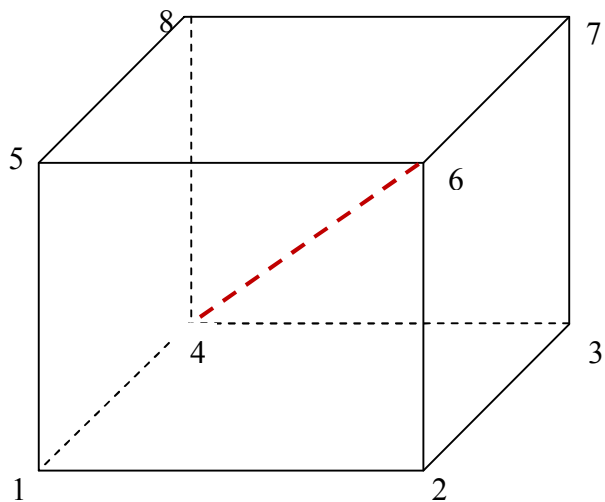
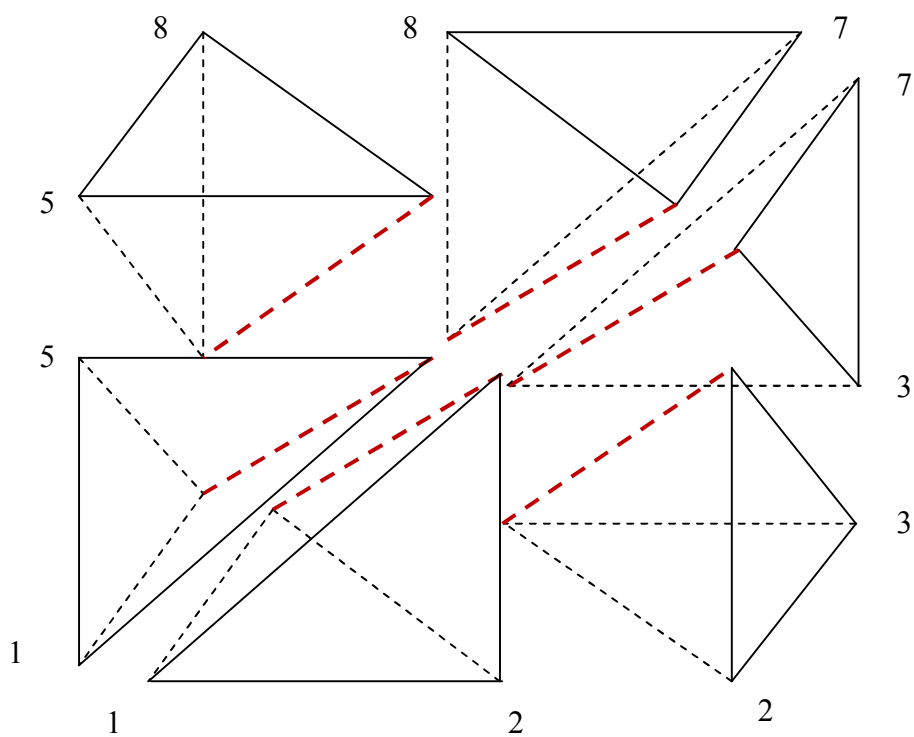


Figure A.2: The first BZ of silicon (a) and its cuboid model (b). The cuboid is divided into smaller cubes (shaded region).

Each cube of size $\frac{2\pi/a}{N_{mesh}}$ (figure A.2b) is then divided into six different tetrahedra of equal volume (figure A.3).



(a)



(b)

Figure A.3: (a) The magnified version of the shaded cube in figure A.2b. (b) Six equal-volume tetrahedral composition of a cube.

Within each tetrahedron, velocity is assumed to be constant, and energy is linearly interpolated; the equal-energy surface is a plane. The constant velocity in each tetrahedron is calculated as [9]:

$$\mathbf{v} = \frac{1}{\hbar} (\mathbf{k}_{node2} - \mathbf{k}_{node1}, \mathbf{k}_{node3} - \mathbf{k}_{node1}, \mathbf{k}_{node4} - \mathbf{k}_{node1})^{-1} \begin{pmatrix} E_{node2} - E_{node1} \\ E_{node3} - E_{node1} \\ E_{node4} - E_{node1} \end{pmatrix} \quad (\text{A.4})$$

The intersecting point between an equal-energy plane and a segment of a tetrahedron is determined by

$$\mathbf{k}_{i,j}(\varepsilon) = \mathbf{k}_{nodei} + \frac{E - E_{nodei}}{E_{nodej} - E_{nodei}} (\mathbf{k}_{nodej} - \mathbf{k}_{nodei}) \quad (\text{A.5})$$

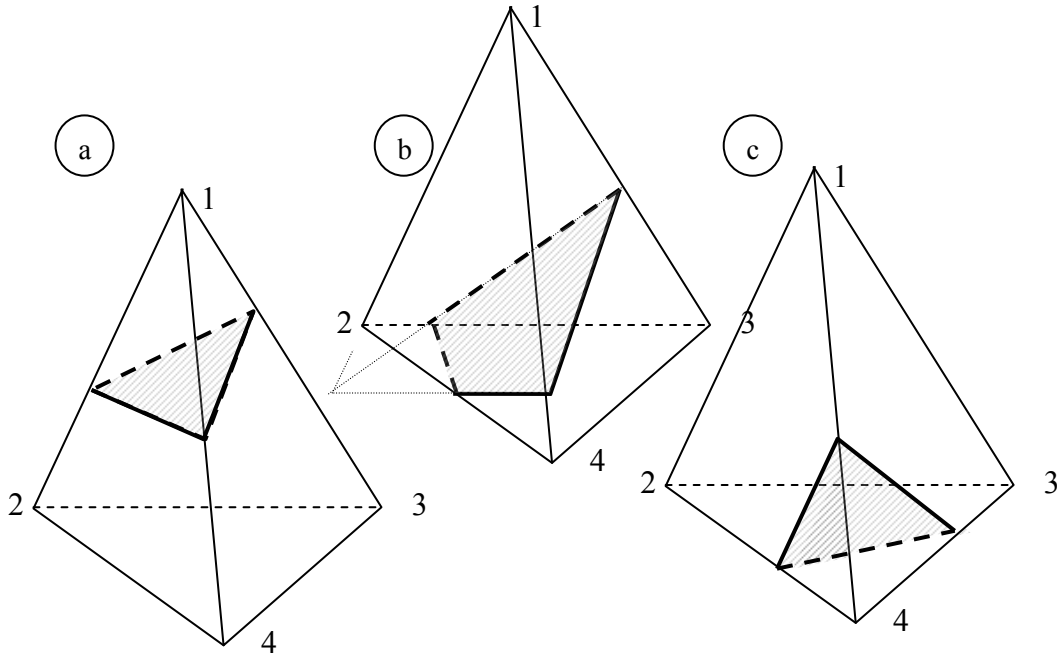


Figure A.4: Three cases of intersections between an equal-energy plane and a tetrahedron.

Each tetrahedron is programmed as a 4x4 array element, whose four rows represent four nodes and four columns represent energy. The four nodes are sorted in the order of increasing energy.

When energy is either smaller than the lowest node or larger than the highest node, the intersecting area between the energy plane and the tetrahedron is zero. Figure A.4 summarizes the cases of the non-zero intersecting area between an equal-energy plane and a tetrahedron.

- **Case a** $E_{node1} < E < E_{node2}$

$$Area = |(\mathbf{k}_{1,2} - \mathbf{k}_{1,4}) \times (\mathbf{k}_{1,3} - \mathbf{k}_{1,4})| \quad (A.6)$$

$$Area = \left| \left[\frac{E - E_{node1}}{E_{node2} - E_{node1}} (\mathbf{k}_2 - \mathbf{k}_1) - \frac{E - E_{node1}}{E_{node4} - E_{node1}} (\mathbf{k}_4 - \mathbf{k}_1) \right] \times \right. \\ \left. \times \left[\frac{E - E_{node1}}{E_{node3} - E_{node1}} (\mathbf{k}_3 - \mathbf{k}_1) - \frac{E - E_{node1}}{E_{node4} - E_{node1}} (\mathbf{k}_4 - \mathbf{k}_1) \right] \right| \quad (A.7)$$

$$Area = (E - E_{node1})^2 \left| \frac{(\mathbf{k}_2 - \mathbf{k}_1) \times (\mathbf{k}_3 - \mathbf{k}_1)}{E_{21}E_{31}} - \frac{(\mathbf{k}_2 - \mathbf{k}_1) \times (\mathbf{k}_4 - \mathbf{k}_1)}{E_{21}E_{41}} - \frac{(\mathbf{k}_4 - \mathbf{k}_1) \times (\mathbf{k}_3 - \mathbf{k}_1)}{E_{41}E_{31}} \right| \quad (A.8)$$

where $E_{ij} = E_{nodej} - E_{nodei}$

From equations (A.4) and (A.8), it is easy to derive the following equation:

$$Area = V_{tet} \frac{3\hbar(E - E_{node1})^2}{E_{21}E_{31}E_{41}} |\mathbf{v}^T| \quad (A.9)$$

where V_{tet} is the volume of the tetrahedron, and $V_{tet} = \frac{1}{6} \left(\frac{2\pi/a}{N_{mesh}} \right)^3$

- **Case b** $E_{node2} < E < E_{node3}$

The intersecting area is the subtraction of the two triangular areas, and it is trivial to derive the following formula:

$$Area = V_{tet} \frac{3\hbar |\mathbf{v}^T|}{E_{31}E_{41}} \left| E_{21} + 2(E - E_{node2}) - \frac{(E_{31} + E_{42})(E - E_{node2})^2}{E_{32}E_{42}} \right| \quad (A.10)$$

- **Case c** $E_{node3} < E < E_{node4}$

Similar to case (a), the area formula is:

$$Area = V_{tet} \frac{3\hbar(E_{node4} - E)^2}{E_{43}E_{42}E_{41}} |\mathbf{v}^T| \quad (A.11)$$

From equation (A.2), using the relationship of velocity and derivative of energy with respect to wave vector, deriving the following expression of DOS is trivial:

$$N(E) = \frac{a^3}{4\hbar(2\pi)^3} \sum_n \int_{A(E)} \frac{d^2k}{|\mathbf{v}(\mathbf{k}_2)|} \quad (A.12a)$$

where $A(E)$ is the equal-energy surface and $\mathbf{v}(\mathbf{k}_2)$ is the velocity.

$$N(E) = \frac{a^3}{4\hbar(2\pi)^3} \sum_n \sum_{tetrahedron} \frac{A(E)}{|\mathbf{v}(\mathbf{k}_2)|} \quad (A.12b)$$

Figure A.5 shows the calculated results of electron DOS contributed from the first four lowest CB using the mesh resolution N=14. Figure A.6 shows the hole DOS calculated using the mesh resolution N=14.

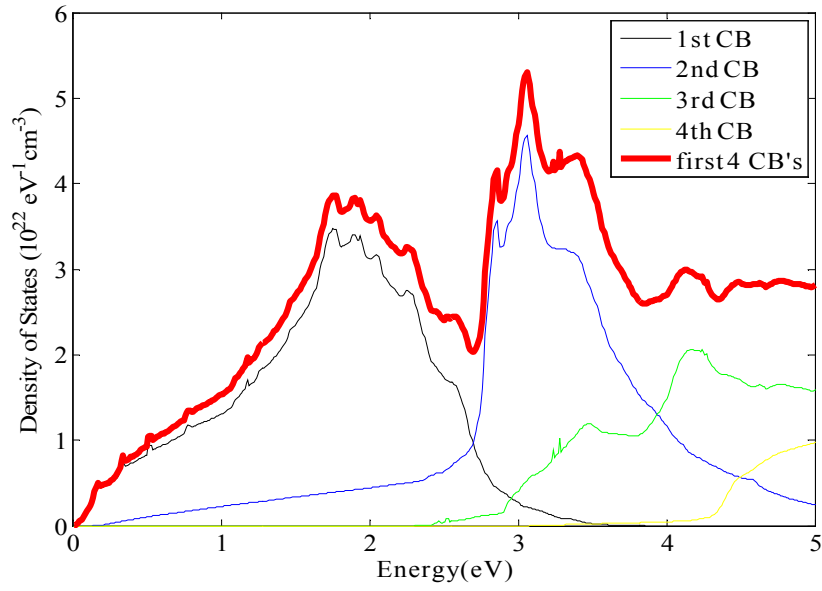


Figure A.5: Electron DOS of silicon contributed by the four lowest conduction bands. Energy equal to 0 corresponds to the lowest energy of the first conduction band. Contributions from each of the four lowest conduction bands are also plotted.

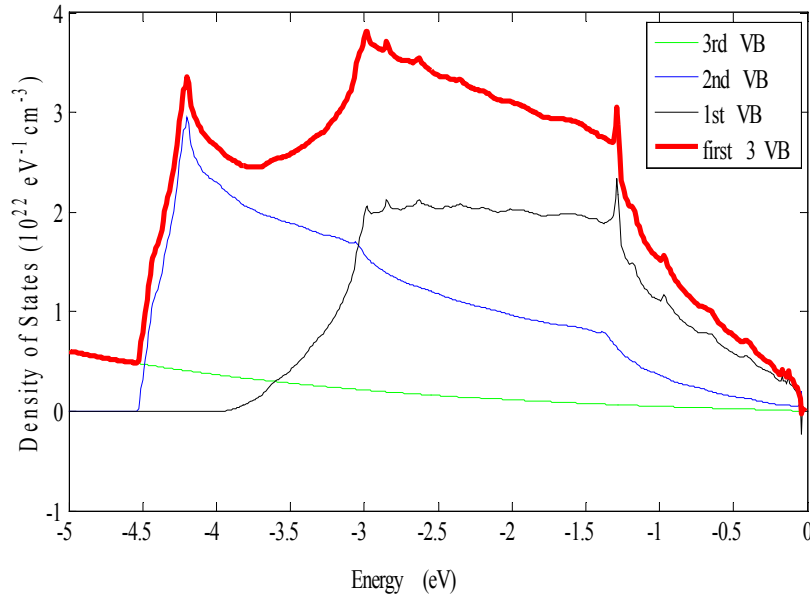


Figure A.6: Hole DOS of silicon from the four highest valence bands versus energy. Contributions from each of the four highest valence bands are also plotted.

A.3 Calculations of impact ionization rates:

Impact Ionization (II) is a process in which a high-energy carrier, let say electron in the CB, collides with an electron in the VB and creates a new electron hole pair (figure A.7).

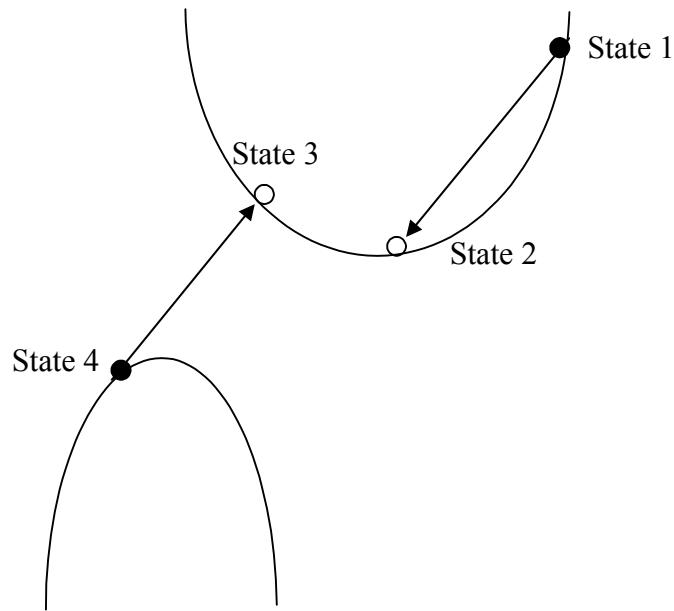


Figure A.7: A typical process of an II event. An electron in state 1 in the conduction band collides with an electron in state 4 in the valence band, resulting in two conduction band electrons in states 2 and 3 and one valence band hole in state 4.

In an II event, the total momentum and energy of electrons and holes are conserved:

$$\mathbf{k}_1 + \mathbf{k}_4 = \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{G} \quad (\text{A.13a})$$

$$E_1 + E_4 = E_2 + E_3 \quad (\text{A.13b})$$

where $\mathbf{G}$ is a reciprocal lattice vector that helps translate all $\mathbf{k}$'s back to the first BZ.

$\mathbf{k}_i$, E_i are wave vector and energy of carrier i , respectively.

II rate is calculated based on the first-order perturbation theory and Fermi's golden rule [10], and the probability of II process in figure A.7 can be described as:

$$P_{II}(1, 4; 2, 3) = \frac{4\pi}{\hbar} T_{II} \delta(E_1 + E_4 - E_2 - E_3) \quad (\text{A.14})$$

where T_{II} is the matrix element for the II process and is treated as a fitting constant [11].

So, the total probability of an electron in state 1 to experience an II event is

$$P_{II}(1) = \frac{1}{2} \sum_{n_2, n_3, n_4} \sum_{\mathbf{k}_2} \sum_{\mathbf{k}_3} \sum_{\mathbf{k}_4} P_{II}(1, 4; 2, 3) \quad (\text{A.15})$$

The factor $\frac{1}{2}$ in front of the right part of the equation comes from the fact that states 2 and 3 are interchangeable and counted twice in the calculation. Due to the conservation of momentum, the summation over $\mathbf{k}_3$ can be eliminated. Equation (A.15) becomes

$$P_{II}(1) = \frac{1}{2} \sum_{n_2, n_3, n_4} \sum_{\mathbf{k}_2} \sum_{\mathbf{k}_4} P_{II}(1, 4; 2, 3) \quad (\text{A.16})$$

Converting the sum series into continuous integrations and considering the spin degeneracy of the electron in valance band, the following equation can be easily derived [11].

$$P_{II}(n_1, \mathbf{k}_1) = \frac{2V^2}{(2\pi)^5 \hbar} \sum_{n_2, n_3, n_4} \int d^3 k_4 \int d^3 k_2 T_{II} \delta(E_1 + E_4 - E_2 - E_3) \quad (\text{A.17})$$

where V is the crystal volume.

The calculation of the integral $I = \int d^3 k_2 T_{II} \delta(E_1 + E_4 - E_2 - E_3)$ is similar to the DOS calculation.

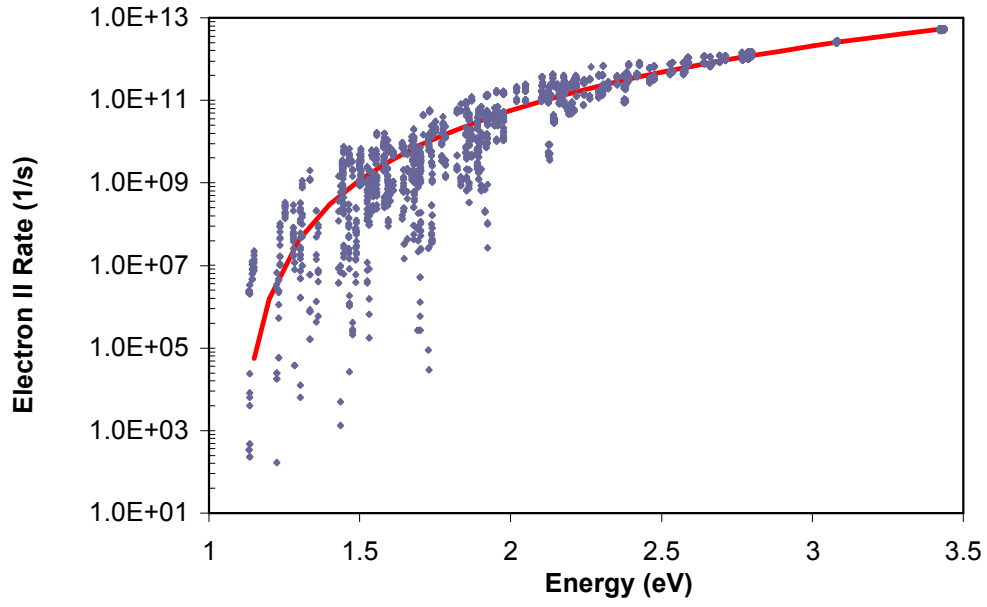
At the given $\mathbf{k}_1$ and $\mathbf{k}_4$, the program scans through all of the tetrahedra in the cuboid BZ. At each tetrahedron, $\mathbf{k}_2$ and $\mathbf{k}_4$ are taken to be the node of the lowest energy. There are no physical reasons for such choices of $\mathbf{k}_2$ and $\mathbf{k}_4$ other than their simplicity for the calculations. Other authors [9] choose the middle point of a cube instead. There is almost no inconsistency as the

result of choosing $\mathbf{k}_2$ and $\mathbf{k}_4$ at another node of each tetrahedron. This is mainly due to the smooth E- $\mathbf{k}$ relationship and the large number of the simulated meshes. With $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_4$ known, using the momentum conservation equation above, $\mathbf{k}_3$ can be figured out, and, consequently, E_3 can be determined. Then I is calculated by summing up the intersecting areas between the equal-energy plane of energy $(E_1 + E_4 - E_3)$ and each tetrahedron containing $\mathbf{k}_2$. The double integrals inside the summation of equation (A.17) are calculated by multiplying the integral I to the tetrahedron's volume containing $\mathbf{k}_4$, and the product is summed up for each $\mathbf{k}_4$ inside the cuboid BZ zone.

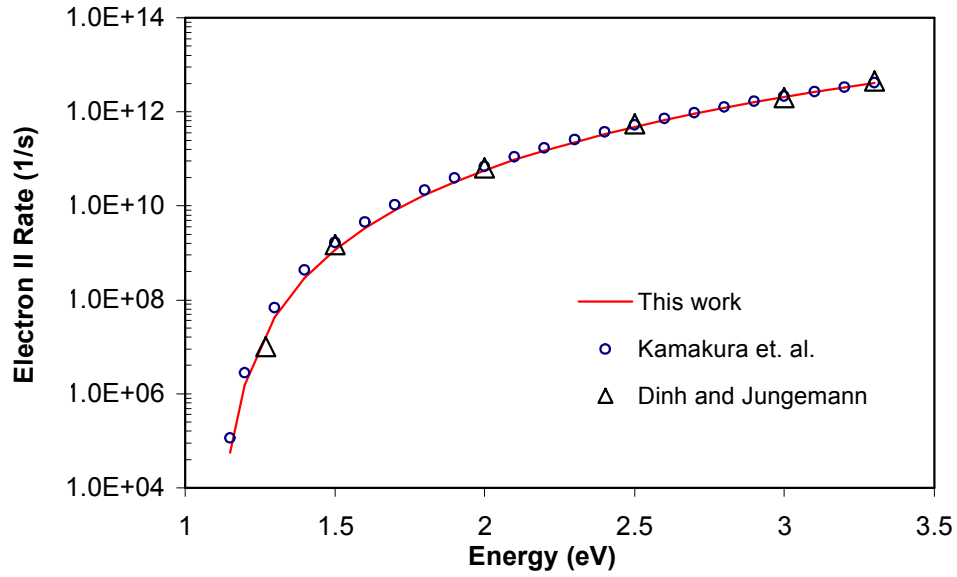
In this calculation, the three highest valance bands and the three lowest conduction bands are included. Figure A.8a shows the II rate of the initiating electron in the lowest CB. The mesh resolution is $N=9$. The cuboid BZ is divided into 17,496 tetrahedra. Energy equal to 0 corresponds to the lowest point in the first CB. The II rate is isotropic at high energy, independent of the direction of $\mathbf{k}$ of the initial electron (electron 1). As energy decreases, II rate is a strong function of $\mathbf{k}$. For example, at $\mathbf{k} = \frac{2\pi}{a}(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, the electron energy is 1.25eV, and the II rate is about 0. Meanwhile, at $\mathbf{k} = \frac{2\pi}{a}(\frac{1}{3}, 0, 0)$, the electron energy is 1.24eV, and the II rate is $4.3 \times 10^7 \text{ (s}^{-1}\text{)}$.

The best-fit curve for the electron II rate is:

$$P_{II}(\text{s}^{-1}) = 9.5 \times 10^{10} (E - 1.1\text{eV})^{4.79} \quad (\text{A.18})$$



(a)



(b)

Figure A.8 (a): Electron II rate versus the initial electron energy. Energy is set to 0 at the lowest point of the first CB. The solid line is the best-fit curve. (b): The best-fit electron II rate versus energy. Previous reports of II rates (Kamakura et. al.[7] and Dinh and Jungemann [8]) are shown for comparison.

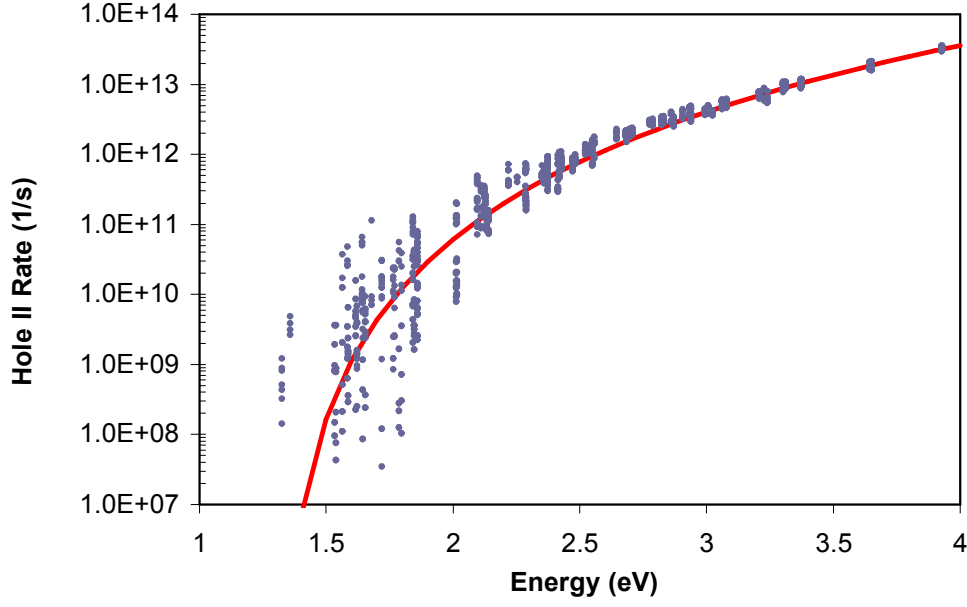


Figure A.9: Hole II rate as a function of the primary hole energy.

Figure A.8b shows the best-fit curve of the calculated II. The result agrees very closely with Ref. [7-8]. It is worth mentioning that the II rate in [7] is calculated with the matrix element calculated for each tetrahedron. This confirms the validity of the constant matrix element assumption in this work.

Hole II rate is shown in figure A.9. Energy equal to 0 corresponds to the highest point in the VB edge. The best-fit curve of the II rate yields

$$P_{II}(s^{-1}) = 3.3 \times 10^{11} (E - 1.3 eV)^{4.72} \quad (A.19)$$

Bibliography

- [1] Marvin L. Cohen and T. K. Bergstresser, "Band Structures and Pseudopotential Form Factors for Fourteen Semiconductors of the Diamond and Zinc-blende Structures," *Physical Review*, Vol. 141, No. 2, 1966.
- [2] James R. Chelikowsky and Marvin L. Cohen, "Electronic structure of silicon," *Physical Review B*, Vol. 10, No. 12, 1974.
- [3] James C. Phillips and Leonard Kleinman, "New Method for Calculating Wave Functions in Crystals and Molecules," *Physical Review*, Vol. 116, pp. 287, 1959.
- [4] Enzo Ungersboeck, Siddhartha Dhar, Gerhard Karlowatz, ViktorSverdlov, Hans Kosina, and Siegfried Selberherr, "The Effect of General Strain on the Band Structure and Electron Mobility of Silicon," *Transactions on Electron Devices*, Vol. 54, No. 9, 2007.
- [5] Simon M. Sze, "Physics of Semiconductor Devices," John Wiley & Sons, 1981.
- [6] Peter E. Blöchl, O. Jepsen and O. K. Andersen, "Improved Tetrahedron Method for Brillouin-Zone Integrations," *Physical Review B*, Vol. 49, No. 23, 1994.
- [7] Y. Kamakura, H. Mizuno, M. Yamaji, M. Morifuji, K. Taniguchi, C. Hamaguchi, T. Kunikiyo, and M. Takenaka, "Impact Ionization Model for Full Band Monte Carlo Simulation," *Journal of Applied Physics*, Vol. 75, pp. 3500, 1994.
- [8] Thanh Viet Dinh and Christoph Jungemann, "Impact Ionization Rates for Strained Si and SiGe," *Solid-States Electronics*, Vol. 53, pp. 1318, 2009.
- [9] Jungemann Christoph and Meinerzhagen Bernd, "Hierarchical device simulation, the Monte-Carlo Perspective," *Computational Microelectronics*. Wien, NY: Springer: 2003.
- [10] L. I. Schiff, "Quantum Mechanics," McGraw-Hill, New York, 1968.
- [11] Nobuyuki Sano and Akira Yoshii, "Impact-Ionization Theory Consistent With a Realistic Band Structure of Silicon," *Physical Review B*, Vol. 45, No. 8, 1992.